



# THE 31st INTERNATIONAL CONFERENCE ON COMPUTERS IN EDUCATION

# **CONFERENCE** PROCEEDINGS

VOLUME II



## **ORGANIZED BY**



The Asia-Pacific Society for Computers in Education (APSCE)



Learning and Educational

**Technologies Research Unit** 

## **HOST BY**



Kyoto University, Japan Research Council of Evidence-Driven Education

# CO-SPONSORED BY





Uchida Yoko Co.Ltd.

**Photron Limited** 

# SUPPORTED BY

Allied Telesis K.K. Digital Knowledge Co., Ltd. Information Processing Society of Japan (IPSJ) Institute of Electronics, Information and Communication Engineers (IEICE) Japanese Association for Education of Information Studies (JAEIS) Japanese Society for Artificial Intelligence (JSAI) Japanese Society for Educational Technology (JSET) Japanese Society for Information and Systems in Education (JSiSE) Japanese Society for Learning Analytics (JASLA) Shimane University 31<sup>st</sup> International Conference on Computers in Education Conference Proceedings Volume II

Editors:

Ju-Ling SHIH, Akihiro KASHIHARA, Weiqin CHEN, Hiroaki OGATA

Publisher:

Asia-Pacific Society for Computers in Education (APSCE)

Address: Center for Science and Technology for Learning,

National Central University, No. 300, Jhongda Road, Zongli District, Taoyuan City 32001, Taiwan

Telephone: +886-3-4227151 ext. 35407

Fax: +886-3-4227151 ext. 35407

Email: service@apsce.net

Website: http://www.apsce.net

Date of publication: December, 2023

ISBN: 978-626-968-902-6

Copyright 2023 Asia-Pacific Society for Computers in Education.

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, transmitted, in any forms or any means, without the prior permission of the Asia-Pacific Society for Computers in Education. Individual papers may be uploaded on to institutional repositories or other academic sites for self-archival purposes.

# **EDITORS**

Ju-Ling SHIH, National Central University, Taiwan Akihiro KASHIHARA, University of Electro-Communications, Japan Weiqin CHEN, Oslo Metropolitan University and University of Bergen, Norway Hiroaki OGATA, Kyoto University, Japan



# ASSOCIATE EDITORS (in alphabetical order)

Ryan BAKER University of Pennsylvania, USA Ben CHANG National Central University, Taiwan Seb DIANATI Charles Darwin University, Australia Jayakrishnan MADATHIL Indian Institute of Technology Madras, India Ahmed Mohamed Fahmy YOUSEF Fayoum University, Egypt Yuqin YANG Central China Normal University, China Hafed ZARZOUR Souk Ahras University, Algeria

# **MESSAGE FROM THE CONFERENCE CHAIR**

Weiqin CHEN Conference Chair Oslo Metropolitan University and University of Bergen, Norway

On behalf of the organizing committee, I would like to welcome all participants of the 31st International Conference on Computers in Education (ICCE) 2023, the flagship conference series of the Asia-Pacific Society for Computers in Education (APSCE). After the virtual conferences in 2020 and 2021 and the hybrid conference in 2022, we have come back to in-person conference this year.



ICCE is no stranger to Japan. After having successfully hosted ICCE 1999, 2007 and 2014, Japan is once again the host for ICCE, this time in Matsue, the "City of Water". Matsue, with magnificent castle, beautiful gardens, and breath-taking sunset over Lake Shinji, will undoubtedly give participants a unique experience. The conference theme of ICCE 2023, "Designing new technologies for education in a big social change world" signifies the role of technological innovation and adoption in transforming education and addressing societal challenges.

Four outstanding keynote speakers will share their insights across varying areas in the field of computers in education. Curtis J. Bonk from Indiana University, USA, will focus on achieving smarter and more innovative forms of learning where digital technologies provide learners of all ages with open, informal, adaptive, nontraditional, and self-directed learning opportunities. Tak-Wai Chan from National Central University, Taiwan, will share with us his vision of "Global Harwell" as the ultimate educational goal and how Seamless Interest-Driven Co-Creator Theory (SIDC) can contribute to achieving this goal. Davinia Hernández-Leo from Universitat Pompeu Fabra, Barcelona, Spain, will explicate how technology can support learning design and the orchestration of complex learning scenarios and thus improve the efficiency and effectiveness of teachers' tasks. Masaru Kitsuregawa from the Research Organization of Information and Systems, Japan will talk about the research data management platform, GakuNin RDM, that provides support for publishing and sharing of big data including educational data, which has strong implications for learning analytics and educational datamining. There will also be three equally inspiring theme-based invited speeches. Kaushal Kumar Bhagat from Indian Institute of Technology Kharagpur, India, will present the potential benefits of game-based learning and how it can be used to create engaging and effective learning experiences. Brendan Flanagan from Kyoto University, Japan, will discuss challenges and opportunities of educational data science focusing on reading systems. Daner Sun from Education University of Hong Kong will talk about the evolution of mobile learning

environments and share insights gained from her experiences in research and teaching. These speeches connect with the essence of the conference theme in different ways and will stimulate reflections and inspire us to rethink the design of digital technologies and their impacts on education and the society.

Indeed, organizing such a large-scale conference requires the orchestrated efforts and unwavering support from the conference organizing committee members and conference paper reviewers. I would like to express my sincere appreciations to all the individuals who have rendered their help in every possible way to make this conference a reality. The names of the hard-working Local Organizing Committee (LOC) chair and team members, International Program Coordination (IPC) chairs, Sub-conference chairs, Program Committee (PC) members and reviewers, chairs and organizers of Workshops, Interactive Events, Tutorials, Panels, Work-In-Progress Posters (WIPP), Doctoral Student Consortium (DSC), Early Career Workshops (ECW), Executive Summary (ES), APSCE Merit Scholarship, and Showcase of Advancements in Technology-Enhanced Learning in Underrepresented Countries (SATELUC) are enlisted in the proceedings. I am also grateful to all the paper authors and registered participants for their exciting academic contributions to the fruitful intellectual exchange in this conference.

Last but not list, I would like to express my heartfelt appreciation to the Managing Secretary of APSCE Pham-Duc Tho for his support, the standing committee for being flexible and proactive, and the consultants for sharing their experiences and wisdom and advising us along the way.

I hope all participants will have opportunities to renew friendships, forge new friendships and professional collaborations. I trust that you will have a productive and fun-filled time at this very special conference and leave Matsue —a picturesque city with rich and remarkable heritage—with beautiful, affectionate memories.

Thank you!



# MESSAGE FROM THE INT'L PROGRAM COORDINATION CHAIRS



Ju-Ling SHIH Program Coordination Chair National Central University, Taiwan



Akihiro KASHIHARA Program Coordination Co-Chair University of Electro-Communications, Japan

The International Conference on Computers in Education (ICCE) is an annual conference series encompassing a broad range of issues related to using Information and Communication Technology (ICT) for education, organized by the Asia-Pacific Society for Computers in Education (APSCE). ICCE 2023 takes place at Matsue, Shimane prefecture, Japan from December 4-8, 2023. It aims to bring together researchers from all over the world to share and exchange research and to develop and deploy new ideas that span the field of Computers in Education.

Following the tradition of previous conferences in this series, ICCE 2023 is organized as a metaconference, where there are seven Sub-Conferences, each of which focuses on specialized themes. Each Sub-Conference is organized by a program committee appointed by the respective Special Interest Group (SIG – see https://apsce.net/sigs). These Sub-Conferences are:

- C1: ICCE Sub-Conference on Artificial Intelligence in Education/Intelligent Tutoring System (AIED/ITS)
- C2: ICCE Sub-Conference on Computer-supported Collaborative Learning (CSCL) and Learning Sciences (LS)
- C3: ICCE Sub-Conference on Advanced Learning Technologies (ALT), Learning Analytics, Platforms and Infrastructure
- C4: ICCE Sub-Conference on Classroom, Ubiquitous, and Mobile Technologies Enhanced Learning (CUMTEL)
- C5: ICCE Sub-Conference on Educational Gamification and Game-based Learning (EGG)

• C6: ICCE Sub-Conference on Technology Enhanced Language Learning (TELL)

# • C7: ICCE Sub-Conference on Practice-driven Research, Teacher Professional Development and Policy of ICT in Education (PTP)

The International Program Committee is led by a strong and dedicated team, which includes the Conference Chair, the Program Coordination Chair and Co-Chair, Sub-Conference Chairs and Co-Chairs and experts in the field of Computers in Education from many different countries or economies. Former ICCE local organizing and program coordination chairs have played important roles as consultants in overseeing the organization process of this conference.

The conference received a total of 256 papers (192 full, 44 short, and 20 posters) from 26 different countries or economies. Table 1 provides the submissions by the country of the first author in each paper.

Countries or Economies					
Japan	67	Thailand	5	France	1
Taiwan	44	Poland	4	Germany	1
China	41	Australia	2	Italy	1
India	19	Indonesia	2	New Zealand	1
Hong Kong	17	Spain	2	Nigeria	1
Philippines	15	Tunisia	2	Norway	1
Singapore	13	Canada	1	Turkey	1
Malaysia	6	Croatia	1	Viet Nam	1
United States	6	Ecuador	1		

#### Table 1. Submission Statistics (based on first author's country)

All papers were subjected to a rigorous review process by 3 to 5 reviewers from the respective Sub-Conference program committees. After the reviews were completed, a meta-review was provided for each paper. In total, 740 reviews and meta-reviews were received. After a discussion period within the individual program committees led by the Sub-Conference Executive Chairs and Co-Chairs, recommendations were made to the Program Coordination Committee Chair and Co-Chair, who oversaw the review process and quality for all Sub-Conferences.

This resulted in 44 full, 67 short, and 48 poster acceptances across all of the seven Sub-Conferences. The overall acceptance rate for full papers is 22.92%. The acceptance rate for the full papers in the individual Sub-Conference closely mirrored the overall acceptance rate. This is a testimony to the continued maintenance of the quality of presentations in our conference. The complete statistics of paper acceptance is shown in Table 2.

	Total Submissions	Submitted as Full Only	Accepted as Full	Full Only(%)	Accepted as Short	Accepted as Poster	Overall Accepted(%)
C1 - AIED/ITS	46	37	8	21.62%	5	12	54.35%
C2 - CSCL/LS	36	26	6	23.08%	11	8	69.44%
C3 - ALT/LA/PI	49	39	10	25.64%	13	8	63.27%
C4 - CUMTEL	19	14	2	14.29%	5	3	52.63%
C5 - EGG	33	29	7	24.14%	7	7	63.64%
C6 - TELL	27	18	4	22.22%	12	4	74.07%
С7 - РТР	46	29	7	24.14%	14	6	58.70%
ICCE 2023	256	192	44	22.92%	67	48	62.11%

Table 2. Paper Acceptance Statistics

In addition to full papers, short papers and posters, ICCE 2023 includes various program components, such as Keynote Speeches, Theme-based Invited Speeches, Workshops, Interactive Events, Panels, Work-in-Progress Posters (WIPP), Extended Summary (ES), Doctoral Student Consortia (DSC), and Early Career Workshop (ECW). All the papers in these program components are published in separate proceedings with their own ISBN numbers. Pre-conference events are held on the first two days of the conference, including 13 workshops, 4 Interactive Events, DSC, ECW, and APSCE Student Wing Workshop.

We would like to thank all who have contributed to making ICCE 2023 a successful conference. First of all, we would like to thank all paper authors for your contributions and for choosing ICCE 2023 as an outlet to present your research. We would also like to thank the IPC Executive Chairs/Co-Chairs and members, who undertook the responsibility of reviewing and selecting papers that represent research of high quality. Specially thanks to our Keynote and Invited Speakers for accepting our invitations and bring inspiring research to ICCE 2023 participants. The Local Organizing Committee deserves a big thank you for their hard work under the tremendous time pressure.

We hope that all participants will find the activities in ICCE 2023 interesting and inspiring, and have opportunities to meet old friends and establish new professional collaborations. Furthermore, we hope that participants will enjoy not only the academic activities, but also the vibrant and exciting culture experience at Matsue, Shimane prefecture, Japan.



# MESSAGE FROM THE LOCAL ORGANIZING COMMITTEE CHAIR



Hiroaki OGATA LOC Chair Kyoto University, Japan

Kon'nichiwa 😳 (Hello!)

On behalf of the local organising committee, I would like to extend my warm welcome to all delegates of the 31<sup>st</sup> International Conference on Computers in Education (ICCE 2023), held for the fourth time in Japan (the first time in Chiba in 1999, the second time in Hiroshima in 2007, and the third time in Nara in 2014). It is my great pleasure and honor to host ICCE 2023 as it takes place fully in-person this year. The theme of the conference, "Designing new technologies for education in a big social change world," aptly reflects what we had experienced during COVID-19, and digital technologies adoption including AI and big data has taken an exponential leap for transforming education in the new norm.

It is a great privilege to share our beautiful country with you. I hope you will be able to enjoy Matsue city, which is often referred to as Little Kyoto. Known as the "City of Water", Matsue stands by the Sea of Japan where Lake Shinji and Nakaumi meet, in the middle of Shimane Peninsula. A former feudal stronghold, Matsue is a true castle town crossed with many canals and boasts one of the twelve remaining original castles in Japan, and famous for its beautiful sunsets over Lake Shinji. In addition, Matsue is the birthplace of Japanese culture and origin of conventions in Japanese Mythology.

I would like to thank the APSCE Executive Committee for giving us this wonderful opportunity. Our sincere thanks to the standing committee, the International Program Committee, reviewers, authors, participants and student volunteers. Especially, I would like to express my gratitude to all the local organizing committee members and sponsors: Uchida Yoko Co. Ltd., Photron Limited, Research Council of Evidence-Driven Education, and Learning and Educational Technologies Research Unit, Kyoto University, Japan. Also, this event is supported by Allied Telesis K.K., Digital Knowledge Co., Ltd., IPSJ, IEICE, JAEIS, JSAI, JSET, JSiSE, JASLA, and Shimane University, Japan. We trust all of you will enjoy the conference, and take home a lot of great memories from Matsue city in Shimane prefecture, Japan.

Arigato! 🕲 (Thank you!)

# **ORGANIZATION**

## **Standing committee**

Conference Chair Weiqin CHEN,	Oslo Metropolitan University and University of Bergen, Norway
International Program Coo Ju-Ling SHIH,	<i>rdination Chair</i> National Central University, Taiwan
International Program Coo Akihiro KASHIHARA,	<i>rdination Co-Chair</i> University of Electro-Communications, Japan
Local Organizing Committe	e Chair
Hiroaki OGATA,	Kyoto University, Japan
Consultants	
Sridhar IYER,	Indian Institute of Technology, India
Mas Nida MD KHAMBARI,	Universiti Putra Malaysia, Malaysia
Lung Hsiang WONG,	Nanyang Technological University, Singapore

## **Sub Conferences**

# C1: ICCE Sub-Conference on Artificial Intelligence in Education/Intelligent Tutoring System (AIED/ITS) and Adaptive

Learning	
PC Executive Chair	
Ryan BAKER,	University of Pennsylvania, USA

**PC Co-Chair** Sébastien LALLÉ,

Sorbonne University, France

#### SIG Chair

May Marie TALANDRON-FELIPE, University of Science and Technology, Phillipines

# C2: ICCE Sub-Conference on Computer-supported Collaborative Learning (CSCL) and Learning Sciences

Ben CHANG,	National Central University, Taiwan
PC Co-Chair Daniel BODEMER,	University of Duisburg Essen, Germany
<b>PC Co-Chair</b> Gaoxia ZHU,	National Institute of Education, Singapore
<i>SIG Chair</i> Elizabeth KOH,	National Institute of Education, Singapore ix

# C3: ICCE Sub-Conference on Advanced Learning Technologies (ALT), Learning Analytics and Digital Infrastructure

Seb DIANATI,	Charles Darwin University, Australia
<b>PC Co-Chair</b> Zablon PINGO,	Charles Darwin University, Australia
<b>PC Co-Chair</b> Suman LAUDARI,	University of Technology Sydney, Australia
<b>PC Co-Chair</b> Ashwin DIXIT,	Indian Institute of Technology Bombay, India
<i>SIG Chair</i> Ramkumar RAJENDRAN,	Indian Institute of Technology Bombay, India

## C4: ICCE Sub-Conference on Classroom, Ubiquitous, and Mobile Technologies Enhanced Learning (CUMTEL)

----

Yuqin YANG,	Central China Normal University, China
<b>PC Co-Chair</b> Daner SUN,	Education University of Hong Kong, Hongkong
<b>PC Co-Chair</b> Gaoxia ZHU,	Nanyang Technological University, Singapore
<b>PC Co-Chair</b> Lu-Ho HSIA,	National Chin-Yi University of Technology, Taiwan
<i>SIG Chair</i> Daner SUN,	Education University of Hong Kong, Hongkong

## C5: ICCE Sub-Conference on Educational Gamification and Gamebased Learning (EGG)

PC Executive Chair Hafed ZARZOUR,	Souk Ahras University, Algeria
<b>PC Co-Chair</b> Jewoong MOON,	The University of Alabama, USA
<i>PC Co-Chair</i> Junfeng YANG,	Hangzhou Normal University, China
<i>PC Co-Chair</i> Luiz RODRIGUES	Center for Excellence in Social Technologies – NEES, UFAL, Brazil

**SIG Chair** Ahmed TLILI,

Beijing Normal University, China

## C6: ICCE Sub-Conference on Technology Enhanced Language Learning (TELL)

#### **PC Executive Chair**

Ahmed Mohamed Fahmy YOUSEF, Fayoum University, Egypt

#### PC Co-Chair

Ahmed Hosny Saleh METWALLY, Beijing Normal University, China

#### SIG Chair

Vivian WU,

Asia University, Taiwan

# C7: ICCE Sub-Conference on Practice-driven Research, Teacher Professional Development and Policy of ICT in Education (PTP)

#### PC Executive Chair

Jayakrishnan MADATHIL,	Indian Institute of Technology Madras, India
<i>PC Co-Chair</i> Rotem ISREL-FISHELSON,	University of Maryland, USA
<b>PC Co-Chair</b> Shitanshu MISHRA,	UNESCO MGIEP, India
<i>SIG Chair</i> Mas Nida KHAMBARI,	Universiti Putra Malaysia, Malaysia

## Workshop/Tutorial/Interactive Event

WIPP	
<b>PC Co-Chair</b> Zi YANG,	Xiamen University
<b>PC Co-Chair</b> Rui LI,	Hunan university
<b>PC Co-Chair</b> Junjie Gavin WU,	Shenzhen Technology University
<i>PC Executive Chair</i> Di ZOU,	The Education University of Hong Kong

**PC Executive Chair** Wen YUN,

PC Co-Chair

Nanyang Technological University, Singapore

Zhang SI,

Central China Normal University, China

## DSC

<b>PC Executive Chair</b> Vwen Yen Alwyn LEE,	Nanyang Technological University, Singapore
<b>PC Co-Chair</b> Yanjie SONG,	The Education University of Hong Kong
<b>PC Co-Chair</b> Jayakrishnan MADATHIL,	Indian Institute of Technology Madras, India
<i>PC Co-Chair</i> Bo JIANG,	East China Normal University, China

## **ECW**

PC Executive Chair	
Hui-Chun CHU,	Soochow University, Taiwan

**PC Co-Chair** Chengjiu YIN,

Kyushu University, Japan

PC Co-Chair Kai-Hsiang YANG,

National Taipei University of Education (NTUE)

## ES

PC Executive Chair	
Juan ZHOU,	Tokyo Institute of Technology, Japan
PC Co-Chair	
Ruining YANG,	Hunan University, China

## **Merit Scholarships**

PC Executive Chair	
Nur Aira ABD RAHIM,	Universiti Putra Malaysia, Malaysia
PC Co-Chair	
Ting-Chia HSU,	National Taiwan Normal University, Taiwan

PC Co-Chair Niwat SRISAWASDI,

Khon Kaen University, Thailand

## SATELUC

**PC Executive Chair** Lung-Hsiang WONG,

Nanyang Technological University, Singapore

PC Co-Chair	
Mas Nida KHAMBARI,	Universiti Putra Malaysia, Malaysia
PC Co-Chair	
Ivica BOTICKI,	University of Zagreb, Crotia
PC Co-Chair	
May-Marie TALANDRON-FELIPE,	University of Science and Technology of Southern
Philippines, Philippines	
PC Co-Chair	
Saida ULFA,	State University of Malang, Indonesia



# Local Organizing Committee

Local Organizing Chair Hiroaki OGATA,	ACCMS, Kyoto University
Local Coordinator Tatsunori MATSUI,	Waseda University
<b>Registration Coordinator</b> Yusuke HAYASHI,	Hiroshima University
<b>Social Event Coordinator</b> Tomoko KOJIRI,	Kansai University
<b>Printing / Publication Coord</b> Chengjiu YIN,	<b>inator</b> Kyushu University
Web Site Management/ Soc Hiroyuki MITSUHARA,	<b>ial Media Coordinator</b> Tokushima University
Technical Support (Physical) Naka GOTODA,	<b>'Virtual) Coordinator</b> Kagawa University
<b>Sponsorship Coordinator</b> Noboru OKUBO,	UCHIDA YOKO Co., Ltd.
Financial Coordinator Izumi HORIKOSHI,	Kyoto University
Secretary Noriko UOSAKI,	Kyoto University

## **APSCE Advisors**

Riichiro MIZOGUCHI,	Japan Advanced Institute of Science and Technology (JAIST)
Akihiro KASHIHARA,	The University of Electro-Communications
Masanori SUGIMOTO,	Hokkaido University
Tomoko KOJIRI,	Kansai University

## Members

Emmanuel Ayedoun	Kansai University
Yiling DAI	Kyoto University
Brendan John FLANAGAN	Kyoto University

Toshitaka FUKAMI	Shimane University
Yoshiko GODA	Kumamoto University
Shinobu HASEGAWA	Japan Advanced Institute of Science and Technology (JAIST)
Toshihiro HAYASHI	Kagawa University
Yugo HAYASHI	Ritsumeikan University
Yuki HAYASHI	Osaka Metropolitan University
Tsukasa HIRASHIMA	Hiroshima University
Tomoya HORIGUCHI	Kobe University
Tatsuya HORITA	Tohoku University
Rieko INABA	Tsuda University
Tomoo INOUE	University of Tsukuba
Kazuhide KANENISHI	Tokushima University
Yasuhisa KATO	Tokyo Online University
Toshihiro KITA	Kumamoto University
Kento KOIKE	Kyoto University
Yusuke KOMETANI	Kagawa University
Tatsuhiro KONISHI	Shizuoka University
Hidenobu KUNICHIKA	Kyushu Institute of Technology
Fusako KUSUNOKI	Tama Art University
Yukie MAJIMA	Osaka Metropolitan University
Rwitajit MAJUMDAR	Kumamoto University
Hiroyuki MASUKAWA	University of the Sacred Heart, Tokyo
Hideya MATSUKAWA	Tohoku University
Takahiko MENDORI	Kochi University of Technology
Tadashi MISONO	Shimane University
Takashi MITSUISHI	Tohoku University
Makoto MIYAZAKI	Teikyo University
Toshio MOCHIZUKI	Senshu University
Yusuke MORITA	Waseda University
Kousuke MOURI	Hiroshima City University
Masayuki MURAKAMI	Osaka University
Keiichi MURAMATSU	Waseda University
Toru NAGAHAMA	Tohoku University
Kae Nakaya	Tokyo Woman's Christian University
Junko NEMOTO	Meiji Gakuin University
Chifumi NISHIOKA	National Institute of Informatics

Youji OCHI	Kindai University
Yuichi OHKAWA	Tohoku University
Tomoya OHYAMA	Tohoku University
Fumiya OKUBO	Kyushu University
Kiichiro OKUBO	Shimane University
Jun OSHIMA	Shizuoka University
Kazuhisa SETA	Osaka Metropolitan University
Katsusuke SHIGETA	Hokkaido University
Takanori SHIGI	UCHIDA YOKO Co., Ltd.
Atsushi SHIMADA	Kyushu University
Masato SOGA	Wakayama University
Takahiro SUMIYA	Hiroshima University
Katsuaki SUZUKI	Kumamoto University
Hideyuki TAKADA	Ritsumeikan University
Kyosuke TAKAMI	Kyoto University
Ryo ΤΑΚΑΟΚΑ	Yamaguchi University
Yasuhisa TAMURA	Sophia University
Yoshimasa TAWATSUJI	Waseda University
Takahito TOMOTO	Tokyo Polytechnic University
Hiroshi UEDA	Hosei University
Maomi UENO	The University of Electro-Communications
Kenzi WATANABE	Hiroshima University
Yuki WATANABE	Tokyo University of Science
Masanori YAMADA	Kyushu University
Yasuhiro YAMADA	Shimane University
Sho YAMAMOTO	Kindai University
Kazuaki YOSHIHARA	Kindai University

# Special Interest Groups(SIG) 2022-2023

SIG 1: Artificial Intelligence in Education/Intelligent Tutoring Systems/Adaptive Learning (AIED/ITS/AL)

May Marie TALANDRON-FELIPE, University of Science and Technology, Phillipines

SIG 2: Computer-supported Collaborative Learning and Learning Sciences (CSCL) Elizabeth KOH, Nanyang Technological University, Singapore

SIG 3: Advanced Learning Technologies, Platforms and Infrastructure (ALT)

Eunice SARI, UX, Indonesia

SIG 4: Classroom, Ubiquitous and Mobile Technologies Enhanced Learning

Daner SUN, The Education University of Hong Kong, Hong Kong

SIG 5: Educational Gamification and Game-based Learning (EGG)

TLILI, Beijing Normal University, China

SIG 6: Technology Enhanced Language Learning (TELL)

Vivian WU, Asia University, Taiwan

SIG 7: Practice-driven Research, Teacher Professional Development and Policy of ICT in Education (PTP)

Mas NidaMDKHAMBARI, Universiti Putra Malaysia, Malaysia

SIG 8: Development of Information and Communication Technology in the Asia-Pacific Neighborhood (DICTAP)

Patcharin PANJABUREE, Mahidol University, Thailand

SIG 9: Educational Use of Problems/Questions in Technology-Enhanced Learning (EUPQ)

Takahito TOMOTO, Tokyo PolytechnicUniversity, Japan

SIG 10: Learning Analytics and Educational Data Mining (LAEDM)

Ramkumar RAJENDRAN, Indian Institute of Technology Bombay, India

SIG 11: Computational Thinking Education & STEM Education (CTE&STEM)

Chee Kit LOOI, Nanyang Technological University, Singapore

xvii

#### **C1: AIED/ITS PC Member**

Ange Adrienne Nyamen Tato, École de Technologie Supérieure, Canada Benedict du Boulay, University of Sussex, United Kingdom Kritya Bunchongchit, Mahidol University International College, Thailand Michelle Banawan, Asian Institute of Management, Philippines Ryan Baker, University of Pennsylvania, United States April Clarke, University of Canterbury, New Zealand Jose De-La-Cruz, Universidad de Malaga, Spain Brendan Flanagan, Kyoto University, Japan Claude Frasson, University of Montreal, Canada Bastiaan Heeren, Open University, The Netherlands Erik Harpstead, Carnegie Mellon University, United States Tsukasa Hirashima, Hiroshima University, Japan Yuki Hayashi, Osaka Prefecture University, Japan Tomoko Kojiri, Kansai University, Japan Cristina Maier, McGraw Hill Education, United States Tanja Mitrovic, University of Canterbury, Christchurch, New Zealand Tatsunori Matsui, Waseda University, Japan Elvira Popescu, University of Craiova, Romania Valery Psyche, Université TÉLUQ, Canada Maria Mercedes T. Rodrigo, Ateneo de Manila University, Philippines John Stamper, Carnegie Mellon University, United States Kazuhisa Seta, Osaka Metropolitan University, Japan Stefan Slater, Teachers College, United States Masaki Uto, The University of Electro-Communications, Japan Jill-Jênn Vie, Inria Lille, France



#### **C2: CSCL PC Member**

Jürgen Buder, Leibniz-Institut für Wissensmedien, Germany Sanjay Chandrasekharan, Homi Bhabha Centre for Science Education, India Ben Chang, National Central University, Taiwan Ching-Yi Chang, Taipei Medical University, Taiwan Cheng-Huan Chen, Asia University, Taiwan Sherry Chen, National Central University, Taiwan Andrew Clayphan, The University of Sydney, Australia Pryce Davis, University of Nottingham, United Kingdom Xueqi Feng, Southern University of Science and Technology, China Lydia Harbarth, University of Duisburg-Essen, Germany Jo-Chi Hsiao, National Yang Ming Chiao Tung University, Taiwan Lingyun Huang, Education University of Hong Kong, China Gwo-Haur Hwang, National Yunlin University of Science and Technology, Taiwan Heisawn Jeong, Hallym University, Korea Vwen Yen Alwyn, Lee, Nanyang Technological University, Singapore Rose Liang, National University of Singapore, Singapore Feng Lin, Singapore University of Social Sciences, Singapore Pei-Yi Lin, National Tsing-Hua University, Taiwan Sunny Lin, National Yang Ming Chiao Tung University, Taiwan Kian Keong Aloysiu Ong, National Institute of Education, Singapore Johanna Pöysä-Tarhonen, University of Jyvaskyla, Finland Nikol Rummel, Ruhr-Universität Bochum, Germany Lenka Schnaubert, University of Nottingham, United Kingdom Osman Tasdelen, Universität Duisburg-Essen de, Germany Yuyao Tong, University of Hong Kong, Hong Kong Daiyi Wang, Providence University, Taiwan Hongzhi Yang, The University of Sydney, Australia Gaoxia Zhu, National Institute of Education, Nanyang Technological University, Singapore



#### **C3: ALT PC Member**

Vishwas Punjaji Badhe, Indian Institute of Technology Bombay, India Oluwafemi Samson Balogun, University of Eastern Finland, Finland Gaurang Belekar, Indian Institute of Technology, Bombay, India Manuel Caeiro Rodríguez, University of Vigo, Spain Liang Changhao, Kyoto University, Japan Weiqin Chen, Oslo Metropolitan University, Norway Yiling Dai, Kyoto University, Japan Syaamantak Das, Indian Institute of Technology Bombay, India Seb Dianati, Charles Darwin University, Australia Yoshiko Goda, Kumamoto University, Japan Carl Haas, University of Waterloo, Canada Marc Jansen, University of Applied Sciences Ruhr West, Germany Nisumba Soodhani K, Indian Institute of Technology Bombay, India Ulfa Khwaja, Indian Institute of Technology Bombay, India Mehmet Kokoç, Karadeniz Technical University, Turkey Siu Cheung Kong, The Education University of Hong Kong, Hong Kong Suman Laudari, Charles Darwin University, Australia Huiyong Li, Kyoto University, Japan Sonsoles López-Pernas, University of Eastern Finland, Finland Rwitajit Majumdar, Kyoto University, Japan Jon Mason, Charles Darwin University, Australia Ritayan Mitra, Indian Institute of Technology Bombay, India Regina Motz, Universidad de la República, Uruguay Shivsevak Negi, Don Bosco Institute of technology, Mumbai, India Indrayani Nishane, Indian Institute of Technology Bombay, India Rumana Pathan, Indian Institute of Technology Bombay, India Herold Pc, Indian Institute of Technology Bombay, India Zablon Pingo, university Of Technology, Sydney, Australia Antony Prakash, Indian Institute of Technology Bombay, India Ramkumar Rajendran, Indian Institute of Technology Bombay, India Nihar Sanda, Indian Institute of Technology Bombay, India Ismaila Temitayo Sanusi, University of Eastern Finland, Finland Bill Searle, Charles Darwin University, Australia Niusha Shafiabady, Charles Darwin University, Australia Jyoti Shaha, Indian Institute of Technology Bomaby, India Atsushi Shimada, Kyushu University, Japan Daevesh Singh, Indian Institute of Technology Bombay, India María Estrella Sousa Vieira, University of Vigo, Spain Ashwin T S, Indian Institute of Technology Bombay, India Kyosuke Takami, Education Data Science Center, National Institute for Educational Policy Research (NIER), Japan Manjunath K Vanahalli, National Institute of Technology, India Samarth Yadannavar, Indian Institute of Technology Bombay, India



#### **C4: CUMTEL PC Member**

Kaushal Bhagat, Centre for Educational Technology, Indian Institute of Technology, Kharagpur, India Ivana Bosnic, University of Zagreb, Croatia Ivica Boticki, Fakultet elektrotehnike i računarstva, Croatia Huiying Cai, Jiangnan University, China Ben Chang, National Central University, Taiwan Guang Chen, Beijing Normal University, China Feng-Kuang Chiang, Shanghai Jiao Tong University, China Chi-Ming Chu, National Ilan University, Taiwan Haiguang Fang, Capital Normal University, China Xueqi Feng, Southern University of Science and Technology, China Maja Gligora Marković, University of Rijeka, Croatia Martina Holenko Dlab, University of Rijeka, Croatia Yih-Ruey Juang, Jinwen University of Science and Technology, Taiwan Tai-Chien Kao, National Dong Hwa University, Taiwan Chiu-Lin Lai, National Taipei University of Education, Taiwan Chen-Yu Lee, Ling Tung University, Taiwan Jing Leng, East China Normal University, China Xiuhan Li, Central China Normal University, China Ma Luo, East China Normal University, China Igor Mekterović, Fakultet elektrotehnike i računarstva, Croatia Kuo-Liang Ou, National Tsing Hua University, Taiwan Yanjie Song, The Education University of Hong Kong, Hong Kong Daner Sun, The Education University of Hong Kong, China Yuyao Tong, University of Hong Kong, China Zhihong Wan, The Education University of Hong Kong, China Xuefeng Wei, Ludong University, China Longkai Wu, National Institute of Education, Singapore Kai-Hsiang Yang, National Taipei University of Education, Taiwan Xianmin Yang, Jiangsu Normal University, China Yuqin Yang, Central China Normal University, China Ying Zhan, The Education University of Hong Kong, Hong Kong

#### C5: EGG PC Member

Ahmed Ahmim, Faculty of Exact Sciences and Sciences of Nature and Life University of Larbi Tebessi, Algeria Alex Barrett, Florida State University, United States Abdelmalek Bouguettaya, CRTI, Algeria Chefrour, université badji mokhtar annaba, Algeria Chih-Pu Dai, Florida State University, United States Zhaihuan Dai, University of South Florida, United States Samia Drissi, univeristé de souk ahras, Algeria Maazouzi Faiz, Univ annaba, Algeria Zakaria Gheid, University of Souk Ahras, Algeria, Algeria Kamel Eddine Heraguemi, M'sila University, Algeria Hyangeun Ji, Temple University, United States Amine Khaldi, kasdi merbah university, Algeria Lukas Liu, The University of Hong Kong, Hong Kong Soltani Mohamed, Souk Ahras University, Algeria Yanjun Pan, Florida STATE UNIVERSITY, United States Marcela Sávia Pessoa, Universidade do Estado do Amazonas, Brazil Khedairia Soufiane, souk ahras university, Algeria Luke West, Florida State University, United States

#### **C6: TELL PC Member**

Michael Adarkwah, Southwest University, China Ting Da, Beijing Normal University, China Reza Hadi Mogavi, Sharif University of Technology, Iran Feifei Han, Australian Catholic University, Australia Ahmed Hosny, Beijing Normal University, Egypt Muhammad Yasir Mostafa, Beijing Normal University, China Stylianos Mystakidis, University of Patras, Greece Michelle Siao-Cing Guo, National Taipei University of Business, Taiwan Wanwisa Wannapipat, Khon Kaen University, Thailand Vivian Wu, Asia University, Taiwan Dong Yang, Beijing Normal University, China Ahmed Mohamed Fahmy Yousef, Fayoum University, Egypt

#### **C7: PTP PC Member**

Vishwas Badhe, Indian Institute of Technology Bombay, India Aparajita Biswal, B H Gardi College of engineering and technology, India Ivica Boticki, Fakultet elektrotehnike i računarstva, Croatia Chatterjee, Indian Institute of Technology Madras, India Arup Rohan Dasgupta, Anjuman-I-Islam's Kalsekar Technical Campus, India Ajita Deshmukh, MIT-ADT University, Pune, India Anita Diwakar, Indian Institute of Technology Bombay, India Lakshmi Ganesh, Indian Institute of Technology Bombay, India Anchal Garg, University of Bolton, United Kingdom Arnon Hershkovitz, Tel Aviv University, Israel Martina Holenko Dlab, University of Rijeka, Croatia Sajna Jaleel, Mahatma Gandhi University, India Kapil Kadam, Indian Institute of Technology Bombay, India Navneet Kaur, Indian Institute of Technology Bombay, India Najwan Khambari, Universiti Teknikal Malaysia Melaka, Malaysia Chen-Yu Lee, Ling Tung University, Taiwan Ganesh Lokhande, Symbiosis International (Deemed) University, India Jayakrishnan Madathil, Indian Institute of Technology–Madras, India Hagit Meishar Tal, Holon Institute of Technology (HIT), Israel Shitanshu Mishra, UNESCO Mahatma Gandhi Institute of Education for Peace and Sustainable Development, India Priscilla Moses , Universiti Tunku Abdul Rahman, Malaysia Soumya Narayana, Indian Institute of Technology Bombay, India Lucian Ngeze, Indian institute of Technology Bombay, India Yogendra Pal, NIIT University, India Mrinal Patwardhan, Indian Institute of Technology Bombay, India Prajish Prasad, FLAME University, India Rajashri Priyadarshini, Indian Institute of Technology Bombay, India Ashutosh Raina, Indian Institute of Technology Bombay, India Rekha Ramesh, Mumbai University, India Vivek Sabanwar, Indian Institute of Technology Bombay, India Sameer Sahasrabudhe, Indian Institute of Technology Bombay, India Charu Saini, UNESCO Mahatma Gandhi Institute of Education for Peace and Sustainable Development, India Madhuri Srinivas, SMIORE – Education, India Narasimha Swamy, Indian Institute of Technology Bombay, Mumbai, India Briju Thankachan, Indian Educational Technology Association, India Bindu Thirumalai, Tata Institute of Social Sciences in, India Vikram Vincent, Indian Institute of Technology Bombay, India

Ying Zhan, The Education University of Hong Kong, Hong Kong

# EARLY CAREER RESERCHER AWARD WINNER (2023)



**Dr. Rwitajit MAJUMDAR** is an Associate Professor at the Research and Educational Institute for Semiconductors and Informatics at Kumamoto University. He is attached to the Graduate School of Social and Cultural Sciences in the Division of Instructional System Studies. Before joining Kumamoto University in September 2023, he was a senior lecturer at the Academic Center for Computing and Media Studies at Kyoto University since 2021. He joined as a post-doc researcher in Prof. Hiroaki Ogata's lab and moved to Japan in 2018.

Rwitajit graduated from the Inter-disciplinary program (IDP) in Educational Technology at the Indian Institute of Technology Bombay in India, co-advised by Prof. Sridhar

Iver of the Computer Science and Engineering department and Prof. Aniruddha Joshi from the Design School. He did his undergraduate studies and master's from BITS Pilani, India, in Engineering Technology and Design Engineering. He attended doctoral coursework at the Indian Institute of Science in Bangalore before moving to Mumbai for doctoral research.

Rwitajit's research interests include Learning Analytics, designing data-driven services, and studying human-data interactions in the context of education. In the last five years, he has received 3 national grants from JSPS as PI and 3 as co-PI for research related to the GOAL project for designing data-driven platforms to develop learners' self-direction skills and build knowledge model-based learning infrastructure. At Ogata lab, over the years, Rwitajit worked directly with 7 Ph.D. and 7 master's students and other research members in various learning analytics research projects and has co-authored more than 100 international conference papers and 30 journal publications. He continues to bridge researchers from the East in Japan, India, Taiwan, and the West to share expertise and perspectives in different collaborative research projects. In 2023

He has actively participated in the APSCE events, being co-chair for the Advanced Learning Technologies (ALT), Learning Analytics and Digital Infrastructure sub-conference track of ICCE, organizing workshops related to Learning Analytics as well as Embodied Learning in ICCE and contributing to RPTEL journal as an author as well as reviewer. Rwitajit was also awarded the IEEE TCLT Early Career Researcher Award in Learning Technologies during ICALT 2023.

Along with research, Rwitajit likes to travel and experience local cultures and practices. His hobbies include photography, music, and mending broken potteries with the art of Kintsugi.

# LAST TEN YEARS' DISTINGUISHED RESEARCHER AWARD WINNERS

**2022 - APSCE Distinguished Researcher Award** 

Maiga Chang, Athabasca University, Canada 2021 - APSCE Distinguished Researcher Award

Maria Mercedes T. Rodrigo, Ateneo de Manila University, Philippines

2020 - APSCE Distinguished Researcher Award

Wenli CHEN, Nanyang Technological University, Singapore

2015 - APSCE Distinguished Researcher Award

Lung-Hsiang WONG, Nanyang Technological University, Singapore

2014 - APSCE Distinguished Researcher Award

Hiroaki OGATA, Kyushu University, Japan

2011 - APSCE Distinguished Researcher Award

Antonija MITROVIC, University of Canterbury, New Zealand

Chen-Chung LIU, National Central University, Taiwan



# LAST TEN YEARS' EARLY CAREER RESEARCHER AWARD WINNERS

2022 - APSCE Early Career Researcher Award

Daner Sun, The Education University of HongKong, HongKong

2021 - APSCE Early Career Researcher Award

Bo Jiang, East China Normal University, China

2020 - APSCE Early Career Researcher Award

Kaushal Kumar BHAGAT, Indian Institute of Technology, Kharagpur, India

2019 - APSCE Early Career Researcher Award

Cheng-Jiu YIN, Kobe University, Japan

2018 - APSCE Early Career Researcher Award

Ting-Chia HSU, National Taiwan Normal University, Taiwan

2017 - APSCE Early Career Researcher Award

Jon MASON, Charles Darwin University, Australia

2015 - APSCE Early Career Researcher Award

Morris Siu-Yung JONG, The Chinese University of Hong Kong, Hong Kong

# SPEAKERS OF APSCE WEBINAR SERIES (December 2022 – November 2023)

APSCE Webinar #30: The Metaverse and Language Learning

Date: 16 December 2022 (Friday)

Speaker: Prof. Yu-Ju LAN (National Taiwan Normal University, Taiwan) Moderator: Prof. Vivian Wen-Chi WU (Asia University, Taiwan) Curated by: APSCE Technology-Enhanced Language Learning (TELL) SIG

#### APSCE Webinar #31 : Leveraging Deep NLP and Generative AI in Education

Date: 10 March 2023 (Friday) Speaker: Dr. Michelle Banawan, Asian Institute of Management, Philippines Moderator: Dr. May Marie P. Talandron-Felipe, University of Science and Technology of Southern Philippines, Philippines Curated by: APSCE Artificial Intelligence in Education / Intelligent Tutoring Systems / Adaptive Learning (AI-Ed) SIG

#### APSCE Webinar #32 : Graphical organizer-based in-field mobile learning

Date: 21 April 2023 (Friday)

Speaker: Prof. Hui-Chun CHU, Soochow University, Taiwan Moderator: Prof. Jerry Chih-Yuan SUN, National Yang Ming Chiao Tung University, Taiwan Curated by: APSCE Advanced Learning Technologies, Platforms & Infrastructures (ALT) SiG

# APSCE Webinar #33 : Institutional and Psychological Factors Affecting Online Distant

**Foreign Language Learning Behaviors** Date: 5 May 2023 (Friday) Speaker: Prof. Yuichi ONO, University of Tsukuba, Japan Moderator: Prof. Vivian Wen-Chi WU (Asia University, Taiwan) Curated by: APSCE Technology-Enhanced Language Learning (TELL) SIG

### APSCE Webinar #34: Transforming Education with AI and Computational Action

Date: 18 May 2023 PANELIST: Natalie LAO, Massachusetts Institute of Technology, USA Mark FRIEDMAN, App Inventor Foundation, USA Keertan KINI, Stanford University, USA Chair: Ting-Chia HSU, National Taiwan Normal University, Taiwan Curated by: APSCE Computational Thinking in Education/STEM (CTE/STEM) SIG

# APSCE Webinar #35 : Three Challenges in Implementing Multimodal Learning Analytics in Real Learning Environments

Date: 31 May 2023 (Wednesday) Speaker: Assoc. Prof. Bertrand Schneider, Harvard Graduate School of Education, USA Moderator: Dr. Elizabeth Koh, Nanyang Technological University, Singapore Curated by: APSE Computer-Supported Collaborative Learning / Learning Sciences (CSCL/LS) SIG

xxviii

# APSCE Webinar #36 (Postponed): Smart Technologies in Education: Policies for Effective and Ethical Use

Date: 16 June 2023 PANELIST: Habibah Ab Jalil, Universiti Putra Malaysia, Malaysia Vikas Kanungo, World Bank Lung Hsiang WONG, Nanyang Technological University, Singapore Moderator: Khaizer Omar Universti Putra Malaysia, Malaysia Curated by: APSCE Practice-Driven Research, Teachers' Professional Development & ICT Policies (PTP) SIG

# **APSCE** Webinar #37: Moving Toward a Mobile Learning Landscape: Effective Device Integration

Date: 3 July 2023 Speaker: Dr. Helen Crompton, Old Dominion University, USA Moderator: Dr. Daner Sun, Education University of Hong Kong Curated by: APSCE Classroom, Mobile and Ubiquitous Technology Enhanced Learning (CUMTEL) SIG

# APSCE Webinar #36 (Re-scheduled): Smart Technologies in Education: Policies for Effective and Ethical Use

Date: 31 July 2023 PANELIST: Habibah Ab Jalil, Universiti Putra Malaysia, Malaysia Maiga Chang, Athabasca University, Canada Lung-Hsiang Wong, Nanyang Technological University, Singapore Moderator: Muhd Khaizer Omar, Universiti Putra Malaysia, Malaysia Curated by: SIG 7 – Practice-driven Research, Teacher Professional Development and Policy of ICT in Education (PTP)

#### **APSCE Webinar #38: Human-Centered Learning Technologies and Multimodal Data**

Date: 21 July 2023 Speaker: Prof. Michail (Michalis) Giannakos Norwegian University of Science and Technology (NTNU), Norway Moderator:Dr. Ramkumar Rajendran, IIT Bombay, India Curated by: APSCE Learning Analytics and Educational Data Mining (LAEDM) SIG

# APSCE Webinar #39: Unlocking Potential: Leveraging Multimodal Learning Analytics for Collaborative Learning

Date: 11 August 2023 Speaker: Prof. Dragan Gašević, Monash University, Australia Moderator: Dr. Ramkumar Rajendran, IIT Bombay, India Curated by: APSCE Learning Analytics and Educational Data Mining (LAEDM) SIG

APSCE Webinar #40: Where now for 'Smart'? Consequent questions and the co-production of knowledge

Date: 6 September 2023 (Wednesday)

Speaker: Assoc. Prof. Jon Mason, Charles Darwin University, Australia Moderator: Prof. Takahito Tomoto, Chiba Institute of Technology, Japan Curated by: APSCE Education Use of Problems/Questions in Technology-Enhanced Learning (EUPQ) SIG

# APSCE Webinar #41: Building In-Context Understanding of Learning Behaviors for Designing Game-Based Assessments

Date: 3 October 2023 (Tuesday)

Speaker: Assist. Prof. Zhichun "Lukas" Liu, The University of Hong Kong, Hong Kong SAR Moderator: Assoc. Prof. Ahmed Tlili, Beijing Normal University, China Curated by: APSCE Educational Gamification and Game-based Learning (EGG) SIG

#### **APSCE Webinar #42: Digital Transformation of Higher Education: Challenges and Insights** Date: 25 October 2023 (Wednesday)

Speaker: Prof. Merlin Teodosia Suarez, De La Salle University, The Philippines Moderator: Assoc. Prof. Patcharin Panjaburee, Khon Kaen University, Thailand Curated by: APSCE Development of Information and Communication Technology in the Asia-Pacific Neighborhood (DIC TAP) SIG

# **KEYNOTE SPEAKERS**

# Curtis J. BONK Indiana University, USA



## Time to Wake Up from Our Innovative Learning Dreams and Make Smarter Learning a Reality

**Abstract:** For a half century, educators, psychologists, and researchers have been predicting that highly intensive, innovative, and individualized learning formats are only a few years away. Learners of all ages would enter enticing microworlds, highly engaging learning experience holodecks, fully immersive hands-on scenarios, high fidelity simulations and games, Al-based adaptive microlearning snippets, and completely free and open educational resources and courses on any topic. Massive open online classes were promised one day and then on demand microlearning snippets were delivered in the next. The learning related dreams we had in past decades were quickly forgotten as the next wave of learning technology came along. But all those dreams will prove pointless if they fail to address true problems or issues that some aspect of society is struggling with. It is time to wake up from such dreams of a glistening technological future and have our dream machines help us envision a world filled with open, informal, adaptive, nontraditional, and self-directed learning opportunities. When that happens, we will truly have arrived in the age of smarter and more innovative forms of learning where the learner is finally in charge of the dreams.

**Biography:** Curtis J. Bonk is Professor in the School of Education at Indiana University (IU) teaching psychology and technology courses and Adjunct in the School of Informatics at IU. He is a former software entrepreneur, certified public accountant, corporate controller, and educational psychologist who presently is an educational technologist, award-winning writer, highly published researcher, statewide and national awardee in innovative teaching with technology, and internationally acclaimed presenter. Curt is the author of over 400 publications including 20 books such as the Handbook of Blended Learning: Global Perspectives, Local Designs, MOOCs and Open Education in the Global South, The World is Open: How Web Technology is



Revolutionizing Education, and Transformative Teaching Around the World. He has given close to 2,000 talks around the world, including over 300 keynote and plenary talks. In 2020, Curt was awarded the IU President's Award for Excellence in Teaching and Learning Technology and in 2021, he received the David H. Jonassen Excellence in Research Award. Recently, the American Educational Research Association named him a 2022 AERA Fellow for his exceptional contributions to, and excellence in, education research, and the following week, he was honored with the International Engagement award from the IU School of Education. In 2022, he was also listed in the top 2% of scientists in the world based on publication citations for career. In 2023, AERA awarded Curt and his colleague Dr. Min Young Doo from Kangwon National University in Korea with the Outstanding International Research Collaboration Award. Curt Bonk co-hosts the weekly award-winning podcast show, Silver Lining for Learning (https://silverliningforlearning.org/). He can be contacted at cjbonk@indiana.edu and his homepage is <u>http://curtbonk.com/</u>.

# **KEYNOTE SPEAKERS**



Tak-Wai CHAN National Central University, Taiwan

## Global Harmony and Wellbeing —Should it be our Ultimate Educational Goal Worldwide?

**Abstract:** More than 160 years ago, Dickens wrote in the first sentence of his A Tale of Two Cities: "It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of Light, it was the season of Darkness, it was the spring of hope, it was the winter of despair...". Today, the human lifespan has extended, but millions of people have deceased in a short time before our eyes due to COVID-19. Digitization enables all of us to connect and communicate, but we constantly quarrel over different beliefs; online games offer immersive experience, but the specter of children's addiction looms large; artificial intelligence promises to enhance our lives, yet it harbors the potential of great harm to human beings; the advent of metaverse could herald a luminous future, but it may also signal a dark abyss to come; we cheer for technological advancement, but are simultaneously beset by concerns over environmental pollution, wealth disparity, and other daunting challenges.

During my keynote address at AIED2007, I put forward my observation: our research community has been experiencing three orientations of research: dream-oriented, adoption-oriented, and humanity-oriented. Humanity-oriented research, which was emerging at the time, addresses that learning should go beyond knowledge acquisition, and hence it should cover cognitive, affective, social and attitudinal domains. For the sake of humanity, we should strive to lay the foundations for the future world by bettering yourself, nurturing a caring family, incubating a humane society, fostering a peaceful and collaborative world.

xxxiii
Furthermore, I posed 4 grand challenge problems. The first three were informed by my observation of research in artificial intelligence in education (AIED), computersupported collaborative learning, mobile learning, and game-based learning, as well as by the challenges of transforming education at that time. The fourth problem, the 'global educational goal problem'—rethinking the educational goal from the global perspective—was due to the threats such as nuclear holocaust, earth resource exhaustion, climate change, societal polarization, and mass extinction of species. These issues present profound risks to humanity and the planet's future. There is an increasing concern about whether our descendants will be able to survive on Earth, let alone live enjoyably. Recognizing the pressing need to address the global educational goal problem, I reordered the four problems in my last slide, placing it as the first grand challenge problem.

I acknowledge that when I first posed the global educational goal problem, I didn't have any idea about the answer, and I believed that this was the case for most other researchers in our field as well. However, the problem is so fundamental that it underpins why we do what we do.

Given the recent regrettable conflicts in the world, leaders in a prominent society in our field have recently raised questions and made statements: "Can we develop partnerships, processes, and learning environments that can reduce divisiveness and engender abilities to talk, work, and learn across differences? Can we design to disrupt othering and hate and to promote empathy and care? There is growing urgency to continue developing theory and practice to address this aim. We have expertise in our membership, we need to learn from each other, especially from those who have been thinking about this for a while now. This is where we can make a difference and take action as a Society." The world truly stands on the brink of peril. Can we coexist harmoniously with our environment? What is the future of humankind? What role do we, as educators, play in shaping the world of tomorrow?

What is education? It is said that at the end of World War II, a letter was found in a Nazi concentration camp. It is addressed to Teachers (Note 1).

#### Dear Teachers:

I am a survivor of a concentration camp. My eyes saw what no man should witness: gas chambers built by learned engineers, children poisoned by educated physicians, infants killed by trained nurses, women and babies shot and burned by high school and college graduates.

So, I am suspicious of education.

My request is: help your students become human. Your efforts must never produce learned monsters, skilled psychopaths, educated Eichmanns. Reading, writing, arithmetic are important only if they serve to make our children more human.

If knowledge and human values are disconnected, can education lead to a more civilized world? If global society is not harmonious, can wellbeing be achievable?

Mandela once said, "Education is the most powerful weapon which you can use to change the world." Indeed, education stands as a beacon of hope for the future. The education we impart today will shape the destiny of all humans in the years to come.



In two decades or so, today's students in schools and universities will become the pillars of our society.

Designing the future of education is equivalent to designing the future world. Educational researchers—particularly those in our community who are engaged with technology—bear an even greater responsibility to lead change through global collaboration.

In this talk, I will share some thoughts based on years of ongoing discussions among a group of international researchers (Note 2). First, we may define the 'ultimate educational goal worldwide' as the final, most far-reaching aspiration that people envision, plan, and commit to achieving in education. It is designed to serve everyone across all societies and cultures, transcending individual and societal objectives to embrace humanity as a whole. For the sake of brevity, we will refer to the 'ultimate educational goal worldwide' simply as the 'ultimate educational goal' henceforth.

Second, assuming the 'ultimate educational goal' exists, its realization would necessitate a form of 'ultimate education.' Perhaps we can formulate this ultimate education as follows:

ultimate education = (ultimate educational goal, design theory, future digital world)

In this formulation, the 'ultimate educational goal' represents 'why' and 'what' to learn; design theory informs 'how' to learn; and 'future digital world' addresses 'who' and 'where' to learn, as well as how the ultimate education can be supported by and integrated into the digital future world, propelling the transformation of education toward that overarching goal.

We propose 'Global Harmony and Wellbeing' (abbreviated as 'Global Harwell') as a candidate for the ultimate educational goal in this formulation (Note 3). We use 'candidate' to acknowledge that, after thorough discussion, other superior options for the ultimate educational goal may emerge. Additionally, we introduce the Interest-Driven Creator (IDC) Theory as a candidate for the design theory—a theory planned to be a revised version of its predecessor. Moreover, we suggest the term 'Seamless Al World' is a candidate concept (or notion) that succinctly and accurately captures the essence of the future digital world (Note 4).

Once the ultimate educational goal is identified, to impact education significantly and accelerate its transformation, we may:

1. Build global and local awareness of the goal,

2. Establish experimental and model educational sites in various countries and regions,

3. Create a 'global ultimate educational park,' and

4. Disseminate the principles of ultimate education worldwide."

Given its potential far-reaching impact on education, along with its inherent intricacy and the fast-paced nature of the digital world, establishing the ultimate education worldwide will require concerted efforts of several generations. There's an adage: "If you want to go fast, go alone; if you want to go far, go together." In fact, if you go alone, you won't get anywhere. However, given the pressing urgency humanity faces and our planet teetering on the edge of catastrophe, we must unite to achieve swift and substantial progress. We cordially invite you to join us in this fruitful dialogue. To facilitate this, we plan to host a series of forums, either online or in person, to deliberate the aforementioned issues at the website: globalharwellgoal.org. We have chosen 'Global Harwell Goal' as the name for our website because the proposed goal can, at the very least, serve as a reference point and stimulate discussions until a more suitable name for the ultimate educational goal is agreed upon.

Note 1: Chee-Kit Looi forwarded the story to me. The letter, often attributed to Haim Ginott, was published in his book 'Teacher and Child.'

Note 2: Tak-Wai Chan, Chee-Kit Looi, Siu-Cheung Kong, Wenli Chen, Lung-Hsiang Wong, Su Luan Wong, Ben Chang, Ju-Ling Shih, Ying-Tien Wu, Fu-Yun Yu.

Note 3: It is worth noting that most of us are researchers, not experts in philosophy, politics, or religion. Nevertheless, we can still propose possible candidates for the ultimate educational goal based on our knowledge and life experiences, without deliberately considering religious doctrines or ideological beliefs. Furthermore, if Global Harwell is our ultimate educational goal, its fulfillment would require what we might call Global Harwell Education.

Note 4: If we accept Global Harwell as our ultimate educational goal, addressing the 'why' and 'what'; IDC Theory as our activity design framework, explaining the 'how'; and Seamless AI World as the concept describing the digital future, outlining the 'where' and 'who'; then 'Seamless IDC Theory' could be a theory for designing Global Harwell Education.

**Biography:** Professor Tak-Wai Chan is a trailblazer in digital learning and a global leader in the field. Almost 40 years ago, at a time when computers and the internet were not yet mainstream in the mid-eighties, he began researching on AI supported learning for his doctoral dissertation, proposing a new genus of AI in education system called learning companion system in 1988. This virtual companion system, called Integration-Kid, was the first artificial companion in the world. In 1989, he and his students started to build the world's first dedicated networked learning system for collaborative learning and learning through competition games, called Distributed West (1992). In early 2000s, he and his colleagues built the largest online learning community called EduCity (1.5 million learners with 1,700 schools involved in 2003), which was also referred as the first learning society in the world. In the same time period, his team conducted frontier research on mobile learning, intelligent classroom, future classroom, interactive clicker, e-schoolbag, one-to-one technology enhanced learning, and so forth. After this series of research, in 2006, working together with a large group of international researchers mainly from the Western countries, he proposed the concept of Seamless Learning. In 2010s, after some long-term experiments on reading (MSSR) and writing in one-to-one technology enhanced classroom, in collaboration with a group of Asian scholars, he proposed the Interest-

xxxvi

Driven Creator (IDC) Theory. Again, he and his colleagues are now calling for building Seamless Interest-Driven Creator (SIDC) Theory with interested researchers.

In addition to his research, Professor Chan has also been a major founder of two societies: the Asia-Pacific Society for Computers in Education (APSCE) and the Global Chinese Society for Computers in Education (GCSCE). These two societies respectively host annual conference series ICCEs and GCCCEs, as well as the journals RPTEL and JLCE. Moreover, to cope with the expanding research community of the field, he has been assisting the establishment of APSCE Theme-Based International Conference Series (TBICS), including CTE-STEM, ICFULL, MetaACES.



xxxvii

## **KEYNOTE SPEAKERS**



Davinia HERNÁNDEZ-LEO Universitat Pompeu Fabra, Barcelona, Spain

#### **Computers in Education: how can we support teachers?**

**Abstract:** While it is widely agreed that the role of teachers is key to achieve students' learning, research on how technology can support teachers' tasks is often underemphasized. In this talk I will summarise research results leading to practical implications in the design of technologies that improve the efficiency and effectiveness of teachers' tasks, caring also for their wellbeing. In particular, I will focus on how technology can support learning design and the orchestration of complex learning scenarios, such as computer-supported collaborative learning in large classrooms. The technology presented will include authoring tools, teaching community platforms, enactment systems, orchestration dashboards and data-driven interventions based on learning analytics. I will also discuss synergies between technological solutions emphasizing human-in-control and machine-in-control perspectives. During the talk, participants will be able to experience some notions covered by interacting using the PyramidApp tool and the Integrated Learning Design Environment (ILDE).

**Biography:** Davinia Hernandez-Leo is Full Professor, Serra Hunter and ICREA Academia Fellow at the Department of Information and Communications Technologies Department (DTIC) at Universitat Pompeu Fabra, Barcelona (Spain), where she is the director of the Interactive and Distributed Technologies for Education research group (TIDE). She obtained a Ph.D. at University of Valladolid, Spain, and has been visiting researcher at Open University of the Netherlands, Fulbright Scholar at Virginia Tech and visiting academic at the University of Sydney. She has published extensively and received several awards, including best and most cited scientific paper awards and

xxxviii

recognitions for technology contributions. Prof. Hernández-Leo has been Vice-President of the European Association for Technology-Enhanced Learning, a Associate Editor of the IEEE Transactions of Learning Technologies, and is currently an elected member of the CSCL Committee within the International Society of the Learning Sciences and member of the Steering Committee of the European Conference on Technology-Enhanced Learning. She is very active in research supervision (PhD students, visitors, etc.), in participation and lead of European and national projects, and in collaborations with companies, non-profit organizations, policy makers and private foundations. Her research activity is broadly centered on the domain of learning technologies, spanning fields such as learning design technology, computer-supported collaborative learning (CSCL), community platforms and learning analytics.

## **KEYNOTE SPEAKERS**



Masaru KITSUREGAWA Research Organization of Information and Systems, Japan

#### **Building a Research Data Platform and Education**

Abstract: We have entered a highly uncertain, unpredictable age beset by natural disasters and wars around the world as well global-scale pandemics. However, we must not despair at this state of affairs and simply wait in hope of better circumstances. Rather, we must move forward with an eye to the future. The Research Organization of Information and Systems (ROIS), consisting of four distinguished research institutes, aims to solve complex phenomena and issues relating to life, the earth, the natural environment, and human society by reframing these issues from the perspective of information and systems while advancing data science to conduct integrated research that transcends disciplinary boundaries. In line with its mission to support resourcesharing and joint research among all universities, ROIS promotes cutting-edge research in specialized fields through joint research that transcends university boundaries by providing researchers nationwide with access to large-scale, state-of-the-art equipment and facilities, big data, valuable materials, and analytical methods. Especially, the National Institute of Informatics replaced the previous Science Information NETwork (SINET) with the world's fastest ultra-high-speed network infrastructure, SINET6, which provides transmission speeds of up to 400 Gbps. In addition to the over 1,000 institutions and universities currently being served, the network will soon be offered to elementary, junior high, and high schools as well. SINET is also expected to make substantial contributions to industry and continuing education. The full rollout of the GakuNin RDM research data management platform not only provides data management support for individual researchers but also

supports the development of open science by providing a platform for the proper release of research data including educational big data.

**Biography:** Masaru Kitsuregawa graduated from the Electronics Engineering Department, Faculty of Engineering, the University of Tokyo in March 1978, completed his doctorate in information engineering at the same university and received a Ph.D. in 1983. He became a lecturer at the Institute of Industrial Science in April 1983, an associate professor in 1984, and a professor in 1997, all at the same university. He has been director general of the National Institute of Informatics since 2013. Currently he is a president of Research Organization of Information and Systems in Japan. Dr. Kitsuregawa has made numerous distinguished achievements in the database field over a long period. He was a leading researcher on the high-speed operation of a hashed relational database. With a conventional simple method, the relational database operation cost is the square of the number of records. To solve this problem, he developed the GRACE hash method, which operates a database at a linear cost by combining a dynamic destaging method, bucket adjustment and different implementation methods. This method is referred to in Wikipedia as a basic method of operating a relational database. Today, all major database software programs use a hash algorithm. Dr. Kitsuregawa's research established the foundation of this algorithm. In recognition of his achievements in enhancing database performance, including those mentioned above, he received the ACM SIGMOD E. F Codd Innovations Award, which is the most prestigious award in database system research. He was the first recipient from Asia. Also, he was designated a fellow by IEICE, IEEE, and ACM, and also received many awards, including Achievement Award from IEICE, Medal with Purple Ribbon, and Legion d'Honneur, Chevalier.



## **THEME-BASED INVITED SPEAKERS**



Kaushal Kumar BHAGAT Indian Institute of Technology Kharagpur, India

# Game On! Leveraging the Benefits of Game-Based Learning in the Digital Age

Abstract: In today's digital age, game-based learning has become an increasingly popular way to engage students and enhance their learning experiences. Game-based learning leverages the engaging and immersive nature of games to create a fun and interactive learning environment, which can help students to develop critical thinking, problem-solving, and collaboration skills. In this keynote presentation, we will explore the benefits of game-based learning and discuss how it can be used to meet the needs of today's learners. We will discuss the importance of incorporating game-based learning into the classroom and explore some of the latest research on the effectiveness of this approach. We will also explore some fundamental design principles of successful game-based learning and highlight some of the best practices that educators can use to create engaging and effective games for their students. Finally, we will examine some of the challenges and limitations of game-based learning and discuss how educators can work to overcome these obstacles. Overall, this keynote presentation will provide attendees with a comprehensive overview of game-based learning and its potential to transform education in the digital age. Whether you are an educator, a curriculum developer, or a game designer, this presentation will provide valuable insights into how you can leverage the benefits of game-based learning to create engaging and effective learning experiences.

**Biography:** Dr. Kaushal Kumar Bhagat is currently working as an assistant professor in the Advanced Technology Development Centre at the Indian Institute of Technology (IIT), Kharagpur, India. He received his Ph.D. from the National Taiwan Normal

University in September 2016. He then served a two-year postdoctoral position at the Smart Learning Institute at Beijing Normal University. In 2015, Dr. Bhagat received NTNU International Outstanding Achievement Award. He was also awarded the 2017 IEEE TCLT Young Researcher award. In 2020, he received APSCE Early Career Researcher Award (ECRA) from the Asia-Pacific Society for Computers in Education. He was also awarded the 2022 Excellence in Distance Education Award (EDEA) by the Commonwealth of Learning (COL), Canada. He is an associate editor of the British Journal of Educational Technology (BJET), Regional Associate Editor (Asia) of the Journal of Learning for Development (JL4D), and editor-in-chief of Contemporary Educational Technology (CET). He is also an editorial board member of several reputed international journals. He is a consultant for international organizations like the Commonwealth of Learning, UNESCO, etc. His research area of interest includes augmented reality, virtual reality, game-based learning, online learning, and technology-enhanced learning.

## **THEME-BASED INVITED SPEAKERS**



Brendan FLANAGAN Kyoto University, Japan

#### **Challenges and Opportunities of Educational Data Science for Reading Systems**

**Abstract:** As educational systems are collecting an increasing amount of data on the learning behavior of students, its analysis has given rise to the fields of Educational Data Mining, and more recently Learning Analytics. As a result, educational AI that is constructed from and consumes learning behavior data has become more prevalent in learning systems and is fueling increased research attention in the field. While many datasets have been made public to promote research, important issues such as information privacy have also limited broader analysis and have resulted in data silos and hindered replication studies within the community. This talk will give an overview of educational data science focusing on reading systems and discuss important ongoing challenges including data analysis for niche learning contexts, data divide, and insights into methods for promoting collaboration through synthetic data and their possible limitations.

**Biography:** Brendan Flanagan is an Associate Professor at the Center for Innovative Research and Education in Data Science, Institute for Liberal Arts and Sciences, and the Data Science Department at the Graduate School of Informatics, Kyoto University. His research interests include Learning Analytics, Educational Data Science, Computer

xliv

Assisted Language Learning, and the Application of Blockchain in Education. He has also hosted educational data challenges at prominent international conferences for more than 5 years to promote educational data science research. He is currently the Principle Investigator of several government-funded research projects that investigate knowledge and learning process analysis, and explainable predictions by learning systems, and is also part of a Japanese Cabinet Office (NEDO) funded large research project into educational symbiotic AI systems.

## **THEME-BASED INVITED SPEAKERS**



Daner SUN Education University of Hong Kong, Hong Kong

#### **Exploring the Evolution of Mobile Learning Environments**

**Abstract:** The rapid advancement of technology and the changing landscape of education have led to significant changes in technology-enabled learning environments. This presentation will explore the impact of changing situations on mobile technology-enabled learning environments, with the speaker sharing insights as both a researcher and an instructor. The talk will cover the evolving distribution and adjustment of components in these environments, as well as changes in pedagogy before, during, and after the Covid-19 pandemic. Additionally, the speaker will highlight the emerging dominance of new technologies in Hong Kong and worldwide, and propose future research directions for mobile learning.

**Biography:** Dr Daner Sun is an assistant professor at the Department of Mathematics and Information Technology, the Education University of Hong Kong (EdUHK), Hong Kong. Her research interests are mobile learning, STEM education, and higher-order thinking in technology-supported learning. So far, Dr Sun has published more than 30 SSCI journal papers. She serves as the editor/co-editor for conference proceedings and journal special issues and acts as a reviewer in the community. Besides being the awardee of the APSCE Early Career Researcher Award (ECRA) 2022, she is also the awardee of Outstanding Performance in Research 2023, Outstanding Performance in Knowledge Transfer (Team) 2020, and Dean's Research Output Prize 2021 in EdUHK.



W01: The Applications of Generative Artificial Intelligence (GAI) in Education	1
A prototype of a chatbot for evaluating and refining student startup ideas using a large language model	2
Joseph Benjamin Ilagan, Jose Ramon Ilagan	
Exploring the Integration of AI Creation and ICT-assisted Mandarin Teaching Yi-Hsuan Chen, Yun-Fang Tu, Yu-Ju Lan	8
Tools and Approaches of Generative Artificial Intelligence Used in Education Hsin-Hsuan Chung, Fu-Ling Chung, Shu-Min Lin, Yu-Ju Lan	17
Exploring ChatGPT Performance on PISA Multiple Choice Sample Questions Comparing English and Japanese Expression <i>Kyosuke Takami</i>	27
A Systematic Review of Generative Artificial Intelligence in Language Education <i>Zilin Wang, Di Zou, Lap-Kei Keith Lee, Haoran Xie, Fu Lee Wang</i>	33
Integrating ChatGPT into Synectics Model to Improve High School Student's Creative Writing Skill Worapong Khuibut, Sasivimol Premthaisong, Pawat Chaipidech	44
Generating Interactive Stories with ChatGPT to Teach Filipino Values Angelo Miguel Gregorio, Sarah Jessica Manuel, Alyssa Jaye Palmares, Sharlin Mae Tang, Ethel Ong	52
W02: The 7th Computer-Supported Personalized and Collaborative Learning	62
The Effect of Gamification with Self-Regulated Approach to Promoting Nursing Students' Leopold's Maneuvers Performance <i>Intan Setiani, Ching Yi Chang and Jie Chi Yang</i>	63
Development and Usability Evaluation of an Intelligent Personalized Erhu Pitch and Rhythm Learning System <i>Gwo-Haur Hwang, Ping-Tsung Tsai, Jenn-Kaie Lain and Shiuan-Han Huang</i>	72
Impact of Gender on Students' Classroom Engagement, Flow Experience and Learning Outcomes When Game-Based Answering Activities Are Integrated into the Curriculum <i>Gwo-Haur Hwang and Yu-Cheng Lin</i>	82
Implementation of Collaborative Project-based Learning Approach: Spherical Video-based Virtual Reality Creation <i>Husni Mubarok, Gwo-Jen Hwang, Chi-Jen Lin and Darmawansah Darmawansah</i>	92

Improving Thinking Awareness in Animation Scriptwriting Through Learning Supporting Tool <i>Panadda Jaiboonlue, Wasan Na Chai, Taneth Ruangrajitpakorn and Thepchai Supnithi</i>	102
The Trends of Computer-Supported Collaborative Learning in Two Decades <i>Hui-Chun Chu, Gwo-Haur Hwang, Han-Chieh Chao and Ching-Yi Chang</i>	112
Facilitating nursing students' critical thinking and problem-solving competence in a computer supported collaborative learning environment <i>Zi-Ying Jiang and Ching-Yi Chang</i>	118
Motivating Knowledge Seeking and Sharing Behavior in the Online Discussion Forum <i>Yue Hu, Baoping Li, Xiang Zhang and Ling Chen</i>	124
Personalized Hands-on Training Via a Hybrid Intelligent Teacher System Dongkun Han	130
Analysis of Students' Action Decisions about SDGs Issues in Open Inquiry Activity with Real Open Data Yu-Hao Lu, Chia-Chun Tseng and Ju-Ling Shih	135
Empowering Language Learners: Harnessing Computer-Based Writing for Enhanced Chinese Language Proficiency Lung-Hsiang Wong, Guat Poh Aw, Wenli Chen, Yin Ling Cheung and Seok Hwa Sim	145
Reader Characteristics, Reading Order, and Facial Emotions Expressed in Reading Science Texts <i>Yu-San Hsiang, Zheng-Hong Guan and Sunny S.J. Lin</i>	155
Educational Use of SVVR: A Scoping Review on Harnessing EduVenture-VR in Learning and Teaching <i>Youwen Shi</i>	165
Exploring Student Engagement and Teacher-Student Interaction Patterns in Collaborative STEM PBL courses through Epistemic Network Analysis <i>Ting Yen Kuo, Ssu Kuang Chen and Sunny S. J. Lin</i>	173
The Effect of Genres and Reading Orders on Interest, Reading Comprehension, and Process: Evidence from Eye Movement of Multiple-text Reading <i>Ching-Yu Luan, Zheng-Hong Guan and Sunny S. J. Lin</i>	183
Develop and Analysis of Educational Board Game <the golden="" road="" silk=""> on Cultural Cognition <i>Hsin-Ju Wang, Shaun-Wen Chen and Ju-Ling Shih</i></the>	192
Supporting Peer Help Recommendation Based on Learner-Knowledge Model Peixuan Jiang, Kensuke Takii, Changhao Liang, Rwitajit Majumdar and Hiroaki Ogata	198

Conceptual Design of WHALE: a Wise Helper Agent for the LEAF Environment Kento Koike, Rwitajit Majumdar, H. Ulrich Hoppe and Hiroaki Ogata	204
W03: The Applications of Information and Communication Technologies (ICTs) in Adult and Continuing Education	210
Investigating Students' Perceptions of Knowledge-building Environment and Learning Engagement Wei-Shou Chen, Min-Hsien Lee & Jyh-Chong Liang	211
Exploring The Framing Effect of Drawing Task Instructions on Science-Major Novice Preservice Teachers' Technology-use Knowledge Yung-Hsiang Huang, Yi-Sheng Huang, Shan-Hui Su, & Chia-Ching Lin	217
Exploring the Relationship Between Students' Preferences for Teacher Authority and Learning Approaches: An Example of Student Learning Communication Theory and Computer Technology Contexts <i>Wen-Lung Huang</i>	225
Learning Analysis infrastructure to support optimal learning based on IRT and Video Enrollment Data Masahiro Sakamoto, Yukihiro Matsubara, Kousuke Mouri	232
Measuring Understanding in Video-Based Learning Song-Yi Lin, Meilun Shih and Hsin-Mu Tsai	240
Research and Design of Digital Learning Resource Management System in Meteorological Adult Training Based on Cloud Computing <i>Jinfang Hou</i>	249
W04: The 7th International Workshop on Information and Communication Technology for Disaster and Safety Education (ICTDSE)	254
Collaborate, Design, and Generate Cybercrime Script Tabletop Exercises for Cybersecurity Education Joshua Dwight	255
Local Safety Map for a Cornerstone of Local Disaster Prevention: A Case Study of a Historical Local Town in Japan Yasuhisa Okazaki, Hiroshi Wakuya, Yukuo Hayashida, Nobuo Mishima	265
Non-Player Characters for Evacuation Training in Metaverse: Preliminary Experiment Kaito Oe, Yusaku Ichino, Hiroyuki Mitsuhara, Masami Shishibori	271

Investigating How Technology May Negatively Affect the Academic Performance and Sleep Quality of Students <i>Cedric Miguel Chan, Josh Sarte, Allen Peter Sze, Jat Cedric Talampas, Judith Azcarraga</i>	281
W05: 3rd International Workshop on Embodied Learning: Technology Design, Analytics & Practices	287
Embodied Learning of Integer Operations Using a Multitouch Design: Touchy Pinchy Integers Priyadharshni Elangaivendan, Ashwin Ramaswamy, Melwina Albuquerque, & Sanjay Chandrasekharan	288
Maker-Material Creative Embodiments in Collaborative Making Alekh V & Chandan Dasgupta	298
Virtual Reality and Embodied Learning: Unraveling the Relationship via Dynamic Learner Behavior Antony Prakash & Ramkumar Rajendran	307
Enacting Biomolecular Interactions in VR: Impact on Student Conceptual Understanding in Biochemistry <i>Prajakt Pande</i>	317
W06: The 16th Workshop on Technology Enhanced Learning by Posing/Solving Problems/Questions	326
Evaluation of an Automatic Generation System for Tracing Tasks Based on Textbook Programs Tomohiro Mogi, Yuichiro Tateiwa, Takahito Tomoto & Takako Akakura	327
Development of a Semi-Active Learning Support System with Operation Index for the Mathematics of Vectors <i>Tomohito Jumonji, Nonoka Aikawa &amp; Takahito Tomoto</i>	333
Using a Three-Problem Framework to Understand How Nursing College Students Learn to Design Healthcare Animations <i>Chun-Hao Chang</i>	343
Focal: A Proposed Method of Leveraging LLMs for Automating Assessments Peter Meyersa, Annette Hana, Razik Grewala, Mitali Potnisa & John Stampera	349
Preliminary Use of a Learning Game for Arithmetic Word Problems with Elementary School Students Yusuke Hayashi, Kohei Yamaguchi & Tsukasa Hirashima	359
The Effect of Revision-Making of Contextualized Student-Generated Questions on University Students' English Grammar Learning Performance	369

Chih-Chung Lin & Fu-Yun Yu

Review of the Literature on Goal-Oriented Self-Directed Learning Based on Question Posing Sumitra Sadhukhan, Shitanshu Mishra & Sridhar Iyer	375
W07: The 3rd Workshop on Innovative technologies for enhancing interactions and learning motivation	385
The Role of VR-Supported Co-Learning Environments in University Students' Self- Efficacy and Learning Engagement <i>Chiung-Fang Chang and Jerry Chih-Yuan Sun</i>	386
Designing and Evaluating an Attention-Engagement-Error-Reflection (AEER) Approach to Enhance Primary School Students Artificial Intelligence Literacy and Learning-to-Learn Skills: A Pilot Study <i>Siu-Cheung Kong and Yin Yang</i>	394
Exploring the Benefits of Chatbots Game-based Learning in Science Learning Outcomes and Motivation <i>Ching-Huei Chen</i>	400
Enhancing Trigonometry Learning through a Mobile App Maria Alva Aberin, Ma. Louise Antonette De Las Penas, Agnes Garciano, Juan Carlo Mallari, Jumela F. Sarmiento, Mark Anthony Tolentino, and Debbie Marie Verzosa	405
What's Going on Behind the Scenes? Peer Assessment as a Strategy for Enhancing Students' Commitment and Learning Motivation in Working Collaboratively <i>Nur Aira Abd Rahim, Norliza Ghazali, Zakiah Ponrahono, Sharifah Intan Sharina Syed Abdullah,</i> <i>Fadzilah Mohamad, and Syahidatul Khafizah Mohd Hajaraih</i>	411
The Effect of Enhancing Students' Learning Motivation on Bilingual Class with Integrating Augmented Reality in a Multimodality Learning Environment <i>Yi-Fang Lo and Jerry Chih-Yuan Sun</i>	419
The Course Learning Experience as a Forerunner in IT Internship: Assessment from the Student and the Host Training Establishment <i>Arlene Mae Celestial-Valderama</i>	428
Examining Different Affective Factors in Learning with Virtual Reality Hsing-Ying Tu, Silvia Wen-Yu Lee, and Ting-Yueh Hsu	437
W08: The 11th Workshop on Technology-Enhanced STEM Education	447
Arduino as Learning Tools for Artificial Intelligence Education: A Systematic Literature Review Pornchai Kitcharoen, Suppachai Howimanporn, and Sasithorn Chookaew	448

Proposing a Training Model on Energy Management of Compressed Air Systems with Artificial Intelligence of Things	456
Noppadon Monok, Suppachai Howimanporn, and Sasithorn Chookaew	
An Inquiry-based Learning Approach in Engineering Education regarding Simulation Practice on Automation Control <i>Santi Hutamarn</i>	462
Promoting STEM Interest through Empathy and Creative Thinking in a STEM-based Community Service Program Ma. Jenina N. Nalipay, Biyun Huang, Morris Siu-Yung Jong, Ching Sing Chai, and Eric Tsun-Hin Luk	469
A Preliminary Study on Knowledge Reconstruction Activity for Fostering Cognitive Presence in Online Discussion Lintang Matahari Hasani, Kasiyah Junus, Lia Sadita, Tsukasa Hirashima, and Yusuke Hayashi	475
Technological Tools for the Teaching and Learning of Statistics Mark Anthony C. Tolentino, Juan Carlo F. Mallari, Maria Alva Q. Aberin, Ma. Louise Antonette N. De Las Peñas, Agnes D. Garciano, Jumela F. Sarmiento, and Debbie Marie B. Verzosa	485
A Visualization App on Proving Geometric Concepts Ma. Louise Antonette N. De Las Peñas, Debbie Marie B. Verzosa, Maria Alva Q. Aberin, Agnes D. Garciano, Jumela F. Sarmiento, Mark Anthony C. Tolentino, and Juan Carlo F. Mallari	491
AquaFarm Ace: A Game on the Intensive Aquaculture Process of Oreochromis niloticus <i>Gerick Jeremiah Niño Go and Angela Nicole Yap</i>	497
Designing a Training Tool for an Industrial Robot Operating with a Programmable Logic Controller	507
Porramut Watanakul, Suppachai Howimanporn, and Sasithorn Chookaew Designing a Sorting System using Machine Vision Training Kit for Mechatronics and Robotics Engineering Students Pakorn Muangsuk, Suppachai Howimanporn, and Sasithorn Chookaew	512
Computer-Supported Collaborative Work in Academics During the COVID-19 Pandemic <i>Bjorn Svetlana Ng, John Michael Calvara, and Judith Azcarraga</i>	517
Design of a Pneumatics System Learning Material with AR Technology for Vocational Education Students	523
Using of Augmented Reality Technology in the Learning Process of Calculus2 for Higher Education Students Anek Putthidech, Amnaj Sookjam, Suwit Somsuphaprungyos and Sangtong Boonying	528

Encouraging Primary Students' Environmental Awareness by Using STEM Inquiry-based Learning Sasivimol Premthaisong and Pawat Chaipidech	533
Enhancing Understanding of Complex Systems through Analogy-Based Video Scenarios Meera Daulatrao Pawar, Sheeja Vasudevan and Sahana Murthy	543
Effectiveness of Learning Seminar Course to Promoting Research Skills : Observe-Plan- Organize-Present in Science Project <i>Chitphon Yachulawetkunakorn, Witsanu Supandee and Ratthakarn Na Phatthalung</i>	549
The Collaborative Learning of Science Project to Supporting 4C's Skills for Learning in 21st Century <i>Witsanu Supandee and Chitphon Yachulawetkunakorn</i>	555
Implementation of Professional Development Training for Industrial Employees on Artificial Intelligence of Things Sasithorn Chookaew, Suppachai Howimanporn and Warin Sootkaneung	561
STEM Education in Sustainability: A Brief Literature Review (2019 - 2023) Pawat Chaipidech, Sasivimol Premthaisong and Niwat Srisawasdi	567
A Proposal for Mobile-assisted Citizen Inquiry Learning Approach in Learning of Plastic Pollution Chawadol Sriboonpimsuay and Niwat Srisawasdi	573
W09: The First International workshop on Ethics of AIED: Challenges and Opportunities for Achieving Learner Centricity	583
Ethical Challenges of Artificial Intelligence in Education: Achieving Learner Centricity with Respect to Learner Autonomy Meera Pawar, Ram Das Rai, Ramkumar Rajendran	584
Ethical Challenges and Best Practices for Transparency in AIED: A Literature Review and Learner Centric Guidelines <i>Ram Das Rai, Meera Pawar, Ramkumar Rajendran</i>	590
Ethical Implications of Utilizing Artificial Intelligence in Education for Assessment Mihir Patki, Saira Sanadi, Shraddha Jadhav, Ashish Musale, Kapil Kadam	600
Fostering Ethics in AI: Perceptions from the Indian AI Curriculum Ashutosh Raina, Kushal Mundra, Prajish Prasad, Shitanshu Mishra C	609
ChatGPT in Education: Risks to Fairness of Access Sumitra Sadhukhan, Shitanshu Mishra, Sridhar Iyer	617

AIED in K-12 Classrooms: Challenges and Opportunities from an Ethics Lens <i>Anabil Munshi</i>	622
W10: The 12th International Workshop on ICT Trends in Emerging Economies (WICTTEE 2023)	627
Factors Affecting Sustainable Use of Minecraft-based Lessons Dominique Marie Antoinette B. Manahan & Maria Mercedes T. Rodrigo	628
A Mathematical App for the Conceptual Understanding of Area and Perimeter Jumela F. Sarmiento, Debbie Marie B. Verzosa, Maria Alva Q. Aberin, Ma. Louise Antonette N. De Las Peñas, Agnes D. Garciano, Juan Carlo F. Mallari, & Mark Anthony C. Tolentino	634
Design of a Mobile App to Promote Understanding and Fluency in Finding the Equation of a Line Agnes D. Garciano, Maria Alva Q. Aberin, Ma. Louise Antonette N. De Las Peñas, Juan Carlo F. Mallari, Jumela F. Sarmiento, Mark Anthony C. Tolentino & Debbie Marie B. Verzosa	640
Towards Identifying the Learning Affordances of Social Media as Telemedicine Platforms among Physicians in a Developing Economy <i>Michelle Bernabe, Rowena Caguiat &amp; Ryan Ebardo</i>	646
Analyzing Sentiments of ChatGPT Users: Philippine Setting Gisele Dizon & Dr. Ma. Rowena R. Caguiat	652
Implementing a Gamified Inquiry-based Learning with Mobile Learning Perspective to Promote Primary Students' Attitude toward Mathematics Learning Sakda Chaleepliam, Sasivimol Premthaisong & Pawat Chaipidech	662
Application of Blended Learning with PhET Simulation to Encourage Learning in Mathematics of Fractions <i>Atcharaporn Assawaphum, Sasivimol Premthaisong &amp; Pawat Chaipidech</i>	668
Design and Development of a Personalized Recommender System of Student Question- Generation Exercises for Programming Courses <i>Chih-Hung Lai, Pham-Duc Tho</i>	678
A Chatbot for Image Recommendation in Mobile Language Learning Mohammad Nehal Hasnine, Huyen T. T. Bui & Hiroshi Ueda	684
W12: The 1st International Workshop on Computational Thinking and Programming Education (CTPE)	689
The Impact of an Educational Board Game on Students' Learning Achievement and Flow State in Media and Information Literacy <i>Ya-Xuan Zhang, Ting-Chia Hsu</i>	690

Game-Based Learning of AI Image Recognition on Computational Thinking and Self-Efficacy of Undergraduates <i>Hui-Wen Hsueh, Ting-Chia Hsu</i>	699
Using a four-step learning activity in a programming course: classroom participation, learning performance, and attitude <i>Liang-Yi Li, Wen-Lung Huang &amp; Chien-Ping Chung</i>	708
Review of TAM used in Educational Technology Research: A Proposal Framework <i>Kai-Yu Tang, Chun-Hua Hsiao</i>	714
Exploring Motivational Differences in Competitive and Cooperative Game-Based Learning through Educational Board Games <i>Yuan-Yao Chiang, Ting-Chia Hsu</i>	719
Developing Computational Thinking: Using TurtleStitch and Physical Computing <i>Peter Seow, Chin-Lee Ker &amp; Haoran Xue</i>	724
Applying cooperative learning with role division to learn recursion Yungyu Zhuang, Hong-Wei Chen, Jen-Hang Wang & Gwo-Dong Chen	730
Design Scaffolding Inquiry-based instruction to Promote Non-engineering students in STEM Learning <i>Chia-Jung Chang</i>	736
Explainable AI in the Real World: Challenges and Opportunities Dora Horvat, Ivica Boticki, Peter Seow & Antun Drobnjak	741
Empowering Students Computational Thinking through Robotics-enabled STEM Education Daner Sun, Peiyao Tian, Kam Yuen Law, Wai Han Cheuk	750
W13: ICCE Workshop on "Towards the Practice of Seamless Interest-Driven Creators (SIDC)	759
Role of Seamless Learning in Enhancing Interest-Driven Creator Theory Su Luan Wong, Mas Nida Md Khambari and Sai Hong Tang	760
MEGA World – A Platform of Multiplayer Educational Game for All Maiga Chang	765
Web-based VR Education Contents Supporting VR-goggles and User Study Yoshihiro Okada, Kosuke Kaneko and Wei Shi	774
Development of TETPR: Technology-Enhanced Total Physical Response for Elementary Students to Learn English Vocabulary in Indonesia <i>Riska Saputra</i>	780

W14: Quantification of Learning in Virtual Learning Environments	786
Dynamic Facial Expression Recognition through Partial Label Learning and Federated Learning Mohammad Alif Daffa, Manas Gupta, Hao Chen & Cheryl Sze Yin Wong	787
An Empirical Investigation on Google Classroom: Use Behaviour of Malaysian School Teachers Priscilla Moses, Jacqueline Chung Ling Lau, Phaik Kin Cheah, Phoebe Soong Yee Yap, Mas Nida Md Khambari, Su Luan Wong	794
An Exploratory Study on Effects of WeChat-Assisted Self-Regulated English Writing Instruction on EFL Learners' Motivation <i>Ying Zhao, Di Zou, Haoran Xie, Pei Liu</i>	804
A Step toward Characterizing Student Collaboration in Online Knowledge Building Environments with Machine Learning Alwyn Vwen Yen Lee, Chew Lee Teo & Aloysius Ong	815
Student Engagement Detection: Case Study on Using Peer-to-Peer Emotion Comparison with Context Regularization Geyu Lin, Manas Gupta, Cheryl Sze Yin Wong & Huayun Zhang	825
Effectiveness of Synchronous and Asynchronous Online Learning in Bloom's Taxonomy Framework Cheryl Sze Yin Wong, Lily Hoang & Chin Tuan Tan	835
Work-In-Progress Posters (WIPP)	841
Task-based Robot-assisted learning to support L2 Speaking Practice Cheng Yueh Jao, Huichin Yeh, Shih-Hsien Yang, Ming-Chang Wu and Chen-Fu Wang	842
Designing an IoT-based Biorobotic Complex Board Game <eurasia channel=""> Song-Lin Chen, Ju-Ling Shih and Shaun-Wen Chen</eurasia>	845
Application of ChatGPT in the Role-play Game of Modeled United Nations Zhen Hung Tsai, Ju-Ling Shih and Geng-De Hong	848
IOT-integrated SDGs Scenario-based Games in the Classrooms Ching-Chieh Lin, Ju-Ling Shih and Yu-Hao Lu	851
Pupils' perceived immersion, attitudes, and learning effectiveness in virtual field trips: A comparison between immersive and projective environments <i>Kun-Hung Cheng</i>	854
Ethical and Privacy Concerns in Artificial Intelligence Dialogue Systems: Do Students in Higher Education Really Care About Them?	857

Chunpeng Zhai, Santoso Wibowo and Lily Li

Resistance Training Support System with Pose Estimation Koki Yamada, Naka Gotoda and Ryota Akagi	860
Evaluation of Simulators to Promote the Understanding of Bioaccumulation among Elementary School Students Yuka Matsuyama, Shinichi Kamiyama, Hideo Funaoi and Tomokazu Yamamoto	863
App-Infused Preschoolers' Storyline Concept-Driven Numerical Curriculum Design Ruei-Cheng Yen and Ben Chang	866
A Proposal and Evaluation of Learning Advising using a Generative AI Yasuomi Takano, Taketo Tsurube, Haruki Ueno and Hiroshi Komatsugawa	872
The Roles of Students' Help-seeking Profiles and Self-efficacy in the AI-assisted Game-based Learning <i>Ching-Huei Chen</i>	875
Question-Driven Design Process for XAI in Active Video Watching Raul Vincent Lumapas, Antonija Mitrovic, Matthias Galster, Sanna Malinen, Pasan Peiris and Jay Holland	878
Enhancing Mobile Learning App for Revamped Blended Learning Class in Beginner's Chinese Course Yuichi Ohkawa, Xiumin Zhao, Takashi Mitsuishi, Wen Gui and Xuan Wang	881
Investigating Players' Social Interactions in IOT Board Games Yueh-Chi Wang, Ju-Ling Shih and Shaun-Wen Chen	884
Method to Promote Social Facilitation of Learners by Presenting Writing Sounds Tatsuya Ueda, Tatsuya Hamada and Hironori Egi	887
Designing Learning Companions for Enhancing Students' Writing Habits Chang-Yen Liao	890
Practices of ARCS Chinese language instructional design with MR application participation Zhenni Shi, Yuto Nagata and Yusuke Morita	893
The Impact of Metaverse Worlds on International Collaborative Learning for Cross-Cultural Understanding Masako Hayashi, Takehiro Suzuki, Yuki Kawata and Keisuke Goto	896
Designing an Online Course on Learning Analytics for Educators: Preliminary Insights from a Scoping Review Nurbiha A Shukor, Norah Md Noor, Aini Marina Ma'Rof, Noor Dayana Abd Halim, Matthew Mclain, Phillip Rothwell, Graham Downes and Frances Tracy	899

Development of a Music Educational Board Game with Mobile Device: Learning Musical Theory and Emotional perception through Gameplay Song-Zhu Xiao, Chih-Chen Kuo and Huei Tse Hou	902
Designing a Virtual Reality Game for Religious Culture Guided Tour by Combining Voice Guided Scaffolding and Situated Learning Mechanism Jui-Jong Wang, Chih-Chen Kuo and Huei Tse Hou	905
Design and Preliminary Evaluation of an Educational Board Game on Urban Culture and History with Mobile Conceptual Scaffolding <i>Hau-An Yu, Chih-Chen Kuo and Huei Tse Hou</i>	908
A Preliminary Evaluation of Using Realistic Virtual Space in Designing Gamification Training Activities for Museum Interpreters <i>Shen-Yang Ni, Shu-Wei Liu and Huei Tse Hou</i>	911
Unmanned Robotic Online Laboratory with an Intelligent Cloud Teacher Dongkun Han and Martin Yun-Yee Leung	914
Learning Concentration on Virtual Reality Learning: Scale Development and a Pilot Study Yi Hsuan Wang, Chun-Ping Wu, Kuan You Pan and Yu Hui Chen	917
Methodology for the Participatory Design of a Learner-Facing Analytics Dashboard <i>Marie-Luce Bourguet</i>	920
Estimation of Features and Skills of Drawing Experts by Tracing Figures Yasuhisa Tamura, Kazunari Kaizu and Akito Hamano	923
Logical Expression Tutoring System for Controlling Smart Devices in Multi-User Environments Tomoya Takahashi, Yuko Murakami, Hidenobu Watanabe and Kouji Nishimura	926
Proposing a Processing Distribution System for Cross-Organizational Use of Educational Data Takahiro Morita, Yuko Murakami, Hidenobu Watanabe and Kouji Nishimura	929
Catalyzing Python Learning: Assessing an LLM-based Conversational Agent Daevesh Singh, Indrayani Nishane and Ramkumar Rajendran	932
How to "Unboxing" Gamer Competencies: via Systematic Literature Reviews to get Analytical Framework is workable? <i>Tieh-Huai Chang and Mingfong Jan</i>	935
A Novel Interpretation of Classical Readability Metrics: Revisiting the Language Model Underpinning the Flesch-Kincaid Index <i>Yo Ehara</i>	939
Experimental study for a computational model in ITS to predict the learners' state Yoshimasa Tawatsuji, Keiichi Muramatsu and Tatsunori Matsui	942

Gamified Learning Objects for Inclusive Programming and Science Education Saumay Garg, Seema Mittal and Mukta Goyal	945
Exploring the Possibility of Harnessing Drones in Geography Education in High Schools <i>Morris Siu-Yung Jong, Chin-Chung Tsai and De-Qi Zhang</i>	948
Impact of Augmented Reality App on EFL Young Learners' Vocabulary Learning Engagement in a Seamless Learning Environment <i>Yanjie Song, Jianfeng Zhou and Yin Yang</i>	951
Optimization of Personalized Content Providing Function for Adult Learners with Diverse Backgrounds Chikako Nagaoka, Masako Furukawa, Yuan Sun and Kazutsuna Yamaji	954
Showcase of Advancements in Technology-Enhanced Learning in Underrepresented Countries (SATELUC)	957
Promotion of ERD Design Comprehension Using Recomposing Method Putra Prima Arhandi, Bani Satria Andoko, Tsukasa Hirashima (Indonesia)	958
Designing Educational Personas using Generative AI Ivan Terzic, Antun Drobnjak & Ivica Boticki (Croatia)	961
Timorese University Students' Perception on E-learning: A Case Study Agostinho Dos Santos Gonçalves*, Sebastião Pereira, Saida Ulfa (Timor Leste)	964
Technology-Enhanced Environmental Learning: Co-design of Educational Mobile Application Case Ewa Duda, Helena Anacka, Jolanta Kowal, Iwona Nowakowska & Hanna Obracht- Prondzyńska ( <b>Poland</b> )	967
Urban Living Lab Enhanced by a Mobile Application as a New Way to Educate Towards Green and Inclusive Cities <i>Ewa Duda, Łukasz Gontar, Maksymilian Kochański &amp; Mari Hanssen Korsbrekke (Poland)</i>	970
Supporting Engineering Degree Student Wellbeing with Compulsory Lessons on Stress Management Marc Beardsley (Spain)	973
Doctoral Student Consortium (DSC)	976
Exploring the Potential of Mobile Assisted Language Learning in Rural Area of Indonesian <i>Fais Nurul Hadi</i>	977

Learning Effectiveness of Integrating Peer Assessment and Board Games in a Computational Thinking and Artificial Intelligence Unit: Taking Speech Recognition as Examples <i>Mu-Sheng Chen and Ting-Chia Hsu</i>	983
Exploring a Supportive Ecosystem for Creative Teaching in Higher Education <i>Weijing Han and Nur Aira Abdrahim</i>	988
Gamification to Train Soft Skills of Software Professionals Through Active Video Watching <i>Pasan Peiris</i>	992
An Empirical Study of the Computational Thinking Learning Game, "Online Robot City" <i>Tai-Ping Hsu and Ting-Chia Hsu</i>	996
Integrating Explainable Artificial Intelligence in Active Video Watching <i>Raul Vincent Lumapas</i>	1006
Supporting Interactive Learning in Active Video Watching Ehsan Bojnordi	1010
The Design and Use of Conversational Intelligent Tutoring Systems and Computer Simulation for The Use of Students of Technology Entrepreneurship Joseph Benjamin Ilagan	1014
Learning Habits Mining and Data-driven Support of Building Habits in Education Chia-Yu Hsu, Izumi Horikoshi, Rwitajit Majumdar and Hiroaki Ogata	1020
Applying Macro- and Micro-scripts to Facilitate Undergraduates' Interdisciplinary Competence <i>Guo Su</i>	1024
Exploring Business Students' STEM Literacy for Work Preparedness De-Qi Zhang and Morris Siu-Yung Jong	1029
Using meaningful gamification to redesign simulation-labs for engaging learners in science inquiry practices <i>Archana Rane</i>	1033
Extended Summary (ES)	1037
Learner Perceptions on Gamifying Active Video Watching Platforms Pasan Peiris, Matthias Galster, Antonija Mitrovic, Sanna Malinen and Raul Lumapas	1038
Adding Interactive Mode to Active Video Watching Ehsan Bojnordi, Antonija Mitrovic, Matthias Galster, Sanna Malinen and Jay Holland	1042

Evaluating the Assessment of Comment Quality in Learning Communication Skills using Active Video Watching	1045
Raul Vincent Lumapas, Antonija Mitrovic, Matthias Galster, Sanna Malinen, Jay Holland and Negar Mohammadhassan	
Preservice Teachers' Video-Based Reflection Supported by the Teacher Dashboard: An Epistemic Network Analysis <i>Huiying Cai, Linmeng Lu and Xiaoqing Gu</i>	1048
Extraction of Characteristic Answering Behavior Using Handwritten Log Data Junya Atake, Taito Kano, Kohei Nakamura, Chia-Yu Hsu, Izumi Horikoshi and Hiroaki Ogata	1053
Interaction Patterns between Learners and AI Tools for English Writing Yun-Shu Xie, Jin-Ho Jang, Su-Yeon Kim and Young Hoan Cho	1060
A holistic visualisation solution to understanding multimodal data in an educational metaverse platform – Learningverse <i>Yanjie Song, Jiaxin Cao, Lei Tao and Dragan Gašević</i>	1064
Embracing Synchronicity in Distance Education: Unraveling the Paradox <i>Kumiko Aoki</i>	1068
A systematic review on the competences of human-AI collaboration Youngjin Yoo, Young Hoan Cho and Jeewon Choi	1071
Influence of Robot Roles on Self-Review Shunsuke Sada and Akihiro Kashihara	1075
Metaverse and Education: A Bibliometric Analysis Based on the Past Twenty Years Chien-Liang Lin, Zeren Zhu, Yushun Su, Juan Zhou and Yu-Sheng Su	1079
An Interactive Robot Lecture System for Attention and Understanding Recovery Toshiyuki Shimazaki and Akihiro Kashihara	1084
Posthumanizing Creativity and Material Histories Alekh V and Chandan Dasgupta	1088

# W01

## The Applications of Generative Artificial Intelligence (GAI) in Education

## A prototype of a chatbot for evaluating and refining student startup ideas using a large language model

#### Joseph Benjamin ILAGAN<sup>a\*</sup> & Jose Ramon ILAGAN<sup>a</sup> <sup>a</sup>Ateneo de Manila University, Philippines \*jbilagan@ateneo.edu

**Abstract:** Assessing the soundness of business models is a critical skill for aspiring entrepreneurs and is an essential part of entrepreneurship education. However, evaluating business models can be time-consuming, costly, and subjective. This study describes the design and the prototype of a chatbot as a conversational intelligent tutoring system that assesses and gives feedback on business model soundness using natural language processing techniques and GPT-3.5, a large language model (LLM) trained by OpenAI, to help student co-founders learn and refine their startup ideas. Our method involves indexing articles and rubrics for evaluating technology startup pitches by extracting word embeddings via the OpenAI API. The chatbot accepts descriptions of startup businesses from student co-founders through a Telegram chatbot, and these are formatted as prompts and then fed into GPT-3.5. The responses are formulated by GPT-3 based on another set of prompts instructing the bot to give feedback from three virtual panelists: 1) a harsh judge, 2) a neutral expert, and 3) an optimistic investor.

**Keywords:** startups, business model, chatbot, startup panelist, simulation, conversational intelligent tutoring system, large language model, LLM, GPT

#### 1. Introduction

#### 1.1 Context: Entrepreneurship Education, Simulation, and Conversational Intelligent Tutoring Systems through Natural Language Understanding

A *startup* is a temporary organization searching for a business model (Blank, 2013). This definition of a startup is consistent with the working definition of entrepreneurship as a process in which firms search for, discover, and exploit new profit opportunities by engaging in arbitrage or innovation activities (Ross, 2011). Unlike established and traditional businesses, for startups, many unknowns about customers, markets, partners, and technology–essentially elements of the business model itself– must be discovered iteratively and incrementally. Entrepreneurship education must be flexible and adaptive, but studies show that higher education does not have the needed flexibility (Matlay & Mitra, 2002). Traditional teaching approaches to entrepreneurship have been inflexible and ineffective (Fayolle & Gailly, 2008) and not interesting (Wei, 2022). This paper is part of a larger set of studies using simulation for entrepreneurship education (Ilagan, 2022).

Aside from starting and running a business while enrolled in university, recent entrepreneurship education and entrepreneurial experiential learning attempts involve the use of computer simulations to compress learning cycles, reduce time and cost (Almeida & Simões, 2018), and model and illustrate how agent behaviors and interactions in a complex environment involving entrepreneurial ecosystems. At some point, startups will need to pitch their business ideas and models to panelists consisting of practitioners in technology, finance, investments, regulations, and various domains or industries. As it is expensive and infeasible to invite panelists frequently, simulating them to learn how to do proper pitching will be a promising option.

Generative artificial intelligence (GAI) is an unsupervised or partially supervised machine learning framework that generates content using probability and statistics based on existing digital content such as text, video, images, and audio through training examples, thus learning their patterns and distribution (Baidoo-Anu & Owusu Ansah, 2023). A large language model (LLM) is a GAI and mathematical model of the statistical distribution of tokens in the vast public corpus of human-generated text (Shanahan, 2023). From the training, LLMs are able to produce human-like language (Cooper, 2023). The tokens involved include words, parts of words, or individual characters, including punctuation marks. They are generative because we can sample and ask them guestions (Shanahan, 2023). GPT (Generative Pre-trained Transformer), an LLM-based system, is designed to generate or statistically-predict sequences of words, code, or other data, starting from a source input called the prompt (Floridi & Chiriatti, 2020). GPT is based on a deep neural network architecture called a transformer (Devlin et al., 2018; Rahaman, 2023; Vaswani et al., 2017), which trains large amounts of publicly-available data in parallel. They can exhibit creativity in writing from a paragraph to a full article convincingly on almost any topic (Baidoo-Anu & Owusu Ansah, 2023). A conversational agent or AI assistant based on LLM is ChatGPT, an interface to GPT (Introducing ChatGPT, 2023). The use of ChatGPT as chatbots in education has also been reported in blog posts and social media fields (Tlili et al., 2023).

#### 1.2 Objectives

This study describes an interface prototype to a technology startup coach simulator implemented as a chatbot. It discusses the methodology, architecture, and design based on a chatbot-based conversational intelligent tutoring system using a large language model (LLM) backend. The startup coach simulator allows students to formulate and get feedback on their startup idea pitch that discusses the business model, validation of their product-market fit, estimation of market potential, and the formulation of initial financial models. Finally, it provides a safety net for students to iteratively prepare their business model pitches, which are more high-stakes.

#### 1.3 Research Questions

What must be considered when developing and validating as virtual panelists simulator chatbot in the form of a conversational intelligent tutoring system to evaluate and provide valuable feedback on business model descriptions presented by student co-founders?

RQ1. How might we validate the efficacy of the proposed conversational virtual panelists in simulating the feedback provided by real-life panelists (harsh judge, neutral expert, and optimistic investor) based on test datasets of business model descriptions?

RQ2. How could the designed conversational intelligent tutoring system handle the complexity and variability of business model descriptions provided by potential student co-founders regarding understanding and giving relevant feedback?

#### 2. Related work

#### 2.1 Conversational Intelligent Tutoring Systems

Intelligent Tutoring Systems (ITS) (Anderson et al., 1985) simulate human tutors to help students, especially the struggling ones, to improve their learning (Cai et al., 2019) with personalized, step-by-step guidance (Atlas, 2023). An ITS waits until the student has submitted a solution. It then marks individual steps as correct or incorrect. If incorrect, it conducts a debriefing, which discusses individual steps with the student (VanLehn, 2011). Conversational Intelligent Tutoring Systems (CITS) are e-learning systems that deliver tutorial content through discussion, asking and answering questions, identifying gaps in knowledge, and providing feedback in the natural language (Holmes et al., 2017). They use

natural language to cover concepts, break down the learning material into conversations, ask and answer questions, determine knowledge gaps, and provide contextual feedback and corrective interventions (Holmes et al., 2017).

#### 2.2 Generative AI, LLMs, and Conversational Interfaces

LLMs can solve a variety of natural language processing (NLP) tasks zero-shot, without relying on any training data for a given task simply conditioning the model on appropriate prompts (Brown et al., 2020; Qin et al., 2023). Related terms to zero-shot are in-context learning (ICL) and few-shot learning (Bragg et al., 2021), which, in the context of LLMs, is the ability to learn from limited examples (Brown et al., 2020; Zhao et al., 2021). Chain-ofthought (CoT) prompting induces LLMs to generate intermediate reasoning steps before answering (Qin et al., 2023). ChatGPT (Introducing ChatGPT, 2023) is a model trained to interact with GPT conversationally. It is itself trained on GPT-3.5 (Qin et al., 2023; Tlili et al., 2023) and through reinforcement learning through human feedback (RLHF). RLHF involves three steps: training a language model with supervised learning, collecting comparison data based on human preferences and training a reward model, and optimizing the language model against the reward model using reinforcement learning (Qin et al., 2023). Prompting is much easier and cheaper than fine-tuning the whole model, especially if you only have a dozen training examples or cases (Beltagy et al., 2022; Gao, 2021; Wang et al., 2020). Numeric representations of words as vectors (Mikolov et al., 2013) have been used for semantic search in LLMs (Sai Sharath & Banafsheh, 2020). The idea is that an input string or query is compared with an expected question. Tools to make it easier to work with documents for Q&A types of systems have emerged. They use numeric representations or embeddings and chain-of-thought prompting Field (Trivedi et al., 2022) for language understanding (Ott et al., 2023; Trivedi et al., 2022).

#### 3. Approach

The prototype's architecture involves indexing articles and rubrics for evaluating technology startup pitches by extracting word embeddings via the OpenAI API. The chatbot accepts descriptions of startup businesses written in plain English through a Telegram chatbot, and these are formatted as prompt files and then fed into GPT-3.5. The responses are formulated by GPT-3 based on another set of prompts instructing the bot to give feedback from three virtual panelists: 1) a harsh judge, 2) a neutral expert, and 3) an optimistic investor. The choice of these three types of panelists stems from classifications of the types of judges students of the Bachelor of Science in Information Technology Entrepreneurship (BS ITE) from the Ateneo de Manila University in the Philippines have been encountering for their live final panel defenses.

The panelist chatbot uses a simple three-tier (i.e., database, HTTP API server, and user interface) web architecture. Telegram, the instant messaging app chosen as the user interface, sends events to FastAPI, the Python HTTP API server, through an HTTP webhook endpoint. When FastAPI receives user prompts through the webhook, it formats the prompts, sends them to gpt-3.5-turbo through the LangChain library (Chase, 2022), and sends the user the gpt-3.5-turbo response by calling the Telegram bot's API. The embeddings were stored in a vector database (originally ChromaDB, later PostgreSQL with the PGVector extension) and made available to LangChain. The prototype employs a threestep process: collecting articles, extracting text embeddings (or "indexing"), and testing prompting techniques iteratively (Wang et al., 2022) to simulate conversation with the student. LangChain's "question-answering" chains induced gpt-3.5-turbo to act as a startup panelist. Ideas from various Internet sources on tech startups (e.g., founder-market fit, known viable startup business models, how startup competition panelists evaluate pitches, common mistakes to avoid) were encoded into text files (.txt) and converted into embeddings by text-embedding-ada-002. Additional text prompts were used to describe each of the three judges to influence how they would give feedback and rate the startup

business model pitches. Figure 1 below shows the architecture and a screen grab of the session dialogue to show only the truncated pitch from a student startup on top and a portion of the harsh judge's response below (with scores out of 10 and detailed feedback). The dialogue text spans multiple pages and will not be included here.



*Figure 1.* Architecture and Screengrab of the Telegram virtual pitch session between student co-founder and virtual panelists.

The early version of the prototype was shared with a class of 22 students from the BS ITE program. Students expressed informally that the virtual panelist feedback was meaningful and that they could prepare well through multiple iterations of feedback from the virtual judges for their final defense with real human judges at the end of the semester. The prototype did have several limitations, which may be the subject of future improvements. Student testers noted that the prototype did not remember previous parts of the conversation. The developers also noted that the feedback given by the panelists was often not critical enough of what the panelist testers deemed flaws in a pitch.

#### 4. Discussion

For RQ1, the intent is to have student startups use the virtual panelist simulator with their startup pitches and answer a few questions involving the usefulness of the feedback. The typical qualitative questions would be: 1) How useful was the feedback for your startup? 2) What gaps did you realize your pitch had? 3) What changes will you incorporate in your next iteration? 4) How have you been able to improve your pitch based on earlier feedback? 5) What emotion (joy, trust, fear, surprise, sadness, anticipation, anger, or disgust) did you feel about the experience on this run? 6) Do you have suggestions on what to improve in the simulator?

For RQ2, LLMs will allow interaction with the virtual panelists through natural language. Two techniques may be used to increase the quality of the virtual panelist evaluations. First, prompt engineering can elicit more specific and more relevant paths of reasoning from the virtual panelists. Specific techniques such as the chain-of-thought prompting model can enable such specificity. Second, a body of text knowledge can be made available as context for prompts. A well-split body of knowledge can be converted into embeddings representing "text relatedness." The most relevant parts of the body of knowledge to a user query can then be passed as additional context to a prompt. Startup co-founders will be able to deliver their pitch, while the LLM back-end will help the virtual panelists pick up the important points as instructed in the hidden prompts in the backend.

#### 5. Conclusion

While testing the prototype with a few startup volunteers, the system seems ready for the next stage, involving validating and testing more formally with student startup co-founders. The next step for this study is to invite several student startups, have them enter the description of their startups as if they were pitching to the panelists, and provide feedback to the researchers as to the helpfulness of the chatbot in refining their business model.

#### Acknowledgments

We wish to thank Roberto Martin N. Galang, Ph.D., Dean of the John Gokongwei School of Management of the Ateneo de Manila University, for ideas involving validating the chatbot's responses. We also wish to thank Maria Mercedes T. Rodrigo, Ph.D. for her support throughout the process of writing this paper.

#### References

- Almeida, F., & Simões, J. (2018). Serious games in entrepreneurship education. In *Encyclopedia of* Information Science and Technology, Fourth Edition (pp. 800–808). IGI Global.
- Anderson, J. R., Boyle, C. F., & Reiser, B. J. (1985). Intelligent tutoring systems. *Science*, 228(4698), 456–462.
- Atlas, S. (2023). ChatGPT for higher education and professional development: A guide to conversational AI.
- Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Available at SSRN 4337484*.
- Beltagy, I., Cohan, A., Logan IV, R., Min, S., & Singh, S. (2022). Zero- and Few-Shot NLP with Pretrained Language Models. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts, 32–37. https://doi.org/10.18653/v1/2022.acltutorials.6
- Blank, S. (2013). Why the lean start-up changes everything. Harvard Business Review.
- Bragg, J., Cohan, A., Lo, K., & Beltagy, I. (2021). FLEX: Unifying Evaluation for Few-Shot NLP. *Advances in Neural Information Processing Systems*, *34*, 15787–15800. https://proceedings.neurips.cc/paper/2021/hash/8493eeaccb772c0878f99d60a0bd2bb3-Abstract.html
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
- Cai, Z., Hu, X., & Graesser, A. C. (2019). Authoring conversational intelligent tutoring systems. Adaptive Instructional Systems: First International Conference, AIS 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings 21, 593–603.
- Chase, H. (2022). *LangChain* [Python]. https://github.com/hwchase17/langchain (Original work published 2022)
- Cooper, G. (2023). Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence. *Journal of Science Education and Technology*, 1–9.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *ArXiv Preprint ArXiv:1810.04805*.
- Fayolle, A., & Gailly, B. (2008). From craft to science: Teaching models and learning processes in entrepreneurship education. *Journal of European Industrial Training*, *32*(7), 569–593. https://doi.org/10.1108/03090590810899838

- Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30, 681–694.
- Gao, T. (2021). Prompting: Better ways of using language models for NLP tasks. *The Gradient*. https://thegradient.pub/prompting/
- Holmes, M., Latham, A., Crockett, K., & O'Shea, J. D. (2017). Modelling e-learner comprehension within a conversational intelligent tutoring system. In W. M. & T. A. (Eds.), *IFIP Advances in Information and Communication Technology* (Vol. 515, pp. 251–260). Springer New York LLC. https://doi.org/10.1007/978-3-319-74310-3\_27
- Ilagan, J. B. (2022). The design and use of agent-based modeling computer simulation for teaching technology entrepreneurship.
- Introducing ChatGPT. (2023). OpenAI. https://openai.com/blog/chatgpt
- Matlay, H., & Mitra, J. (2002). Entrepreneurship and learning: The double act in the triple helix. *The International Journal of Entrepreneurship and Innovation*, *3*(1), 7–16.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *ArXiv Preprint ArXiv:1301.3781*.
- Ott, S., Hebenstreit, K., Liévin, V., Hother, C. E., Moradi, M., Mayrhauser, M., Praas, R., Winther, O., & Samwald, M. (2023). *ThoughtSource: A central hub for large language model reasoning data* (arXiv:2301.11596). arXiv. https://doi.org/10.48550/arXiv.2301.11596
- Qin, C., Zhang, A., Zhang, Z., Chen, J., Yasunaga, M., & Yang, D. (2023). Is ChatGPT a General-Purpose Natural Language Processing Task Solver? (arXiv:2302.06476). arXiv. https://doi.org/10.48550/arXiv.2302.06476
- Rahaman, M. (2023). Can chatgpt be your friend? Emergence of entrepreneurial research. Emergence of Entrepreneurial Research (February 18, 2023).
- Ross, R. B. (2011). Entrepreneurial behaviour in agri-food supply chains: The role of supply chain partners. *Journal on Chain and Network Science*, *11*(1), 19–30. https://doi.org/10.3920/jcns2011.x193
- Sai Sharath, J., & Banafsheh, R. (2020). Question Answering over Knowledge Base using Language Model Embeddings. 2020 International Joint Conference on Neural Networks (IJCNN), 1–8. https://doi.org/10.1109/IJCNN48605.2020.9206698
- Shanahan, M. (2023). *Talking About Large Language Models* (arXiv:2212.03551). https://doi.org/10.48550/arXiv.2212.03551
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1), 15.
- Trivedi, H., Balasubramanian, N., Khot, T., & Sabharwal, A. (2022). *Interleaving Retrieval with Chain*of-Thought Reasoning for Knowledge-Intensive Multi-Step Questions (arXiv:2212.10509). arXiv. https://doi.org/10.48550/arXiv.2212.10509
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems*, *30*.

https://proceedings.neurips.cc/paper\_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a8 45aa-Abstract.html

- Wang, B., Deng, X., & Sun, H. (2022). Iteratively Prompt Pre-trained Language Models for Chain of Thought. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, 2714–2730. https://aclanthology.org/2022.emnlp-main.174
- Wang, Y., Yao, Q., Kwok, J. T., & Ni, L. M. (2020). Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys (Csur)*, 53(3), 1–34.
- Wei, M. (2022). The Design of Intelligent Tutoring Systems Using College Students' Innovation and Entrepreneurship Education under the Background of Online Teaching. *Wireless Communications and Mobile Computing*, 2022.
- Zhai, X. (2022). ChatGPT user experience: Implications for education. Available at SSRN 4312418.
- Zhao, Z., Wallace, E., Feng, S., Klein, D., & Singh, S. (2021). Calibrate before use: Improving fewshot performance of language models. *International Conference on Machine Learning*, 12697–12706.

## Exploring the Integration of AI Creation and ICT-assisted Mandarin Teaching

Yi-Hsuan CHEN<sup>a</sup>, Yun-Fang TU<sup>b</sup>, Yu-Ju LAN<sup>a\*</sup>

<sup>a</sup> Department of Chinese as a Second Language, National Taiwan Normal University, Taiwan

<sup>b</sup> Department of Educational technology, Wenzhou University, PR China \*yujulan@gmail.com

Abstract: This study explores the integration of artificial intelligence (AI) creation techniques and information and communication technology (ICT) into Mandarin teaching, through students' perceptions of AI creation content and their feedback on engaging with technology in the process. In this study, a total of 21 students described their perceptions of Mandarin teaching as AI creation images and textual descriptions. These images were analyzed by 7 categories of encoding, including people, places, activities, technical media, objects, electronic technologies, and emotional attitudes. According to the analysis results, most students' conception of Mandarin teaching is still based on "teacher lecture" and "hands-on practice" in the "activities", followed by "books" and "classroom equipment" in "objects", and "electronic "computer" and "screens" in the "technologies" categories. There were some challenges in AI creation like an inaccurate prompt that would result in an image generation error such as certain details being missing, images appearing cluttered, blurring of visual depictions, and the barcodes and technological devices also lead to misalignment. However, certain facial expressions and the overall teaching atmosphere are also a challenge to generate during the creation. During the interview, students realized privacy and copyright issues are also a challenge. The results reveal that AI technology offers diverse, engaging teaching methods despite these challenges. By shedding light on how AI creation can enhance language learning, the study contributes to the evolving landscape of technology-assisted education. As AI technologies continue to shape modern education, these findings offer crucial guidance for educators aiming to leverage AI's benefits while addressing its limitations.

Keywords: AI creation, ICT-assisted Mandarin Teaching, Drawing analysis

#### 1. Introduction

In the rapidly evolving era of information and communication technology (ICT), both society and education are encountering transformative shifts. All educators, regardless of subject areas or grade levels, need to be competent in integrating ICT effectively into their teaching methodologies. On the other hand, students require the cultivation of ICT skills to foster cultural competency, effective communication, collaborative aptitude, problem-solving prowess, as well as self-directed learning and personal growth (Tan & Liu, 2016).

Previous research work on ICTs has focused on language teaching facilitated by electronic tools, platforms, and teaching aids, often employing interviews and surveys to solicit students' perspectives (Čok, 2016; Le Thi, 2020; Abraham et al., 2022). While, some scholars have emphasized that drawing can generate ideas and values in non-verbal expressions (Nuora et al., 2019), which not only creates an environment that makes participants feel less constrained but also enables them are able to express ideas and concepts, no limited by the questionnaire item structure (Lai, 2021). Drawing as a research method has been used as an effective tool in analyzing different educational levels, including K-12 students (Hsieh & Tsai, 2018) as well as college students (Liou, 2017; Tu et al., 2017; Lai, 2021). Within the realm of drawing research, this study selects one of the generative artificial intelligence (GAI) tools for analysis, i.e. Midjourney. This technique addresses limitations inherent in traditional hand-drawn illustrations, such as students' self-doubt, time consumption, and skill prerequisites. Utilizing GAI tools assists students in visualizing ideas,
enhancing efficiency, precision, and consistency. Furthermore, these tools adeptly handle complex concepts, facilitate easy image modifications, and maintain a lower threshold for use.

Therefore, this study seeks to explore the integration of AI creation techniques within Mandarin teaching, with a dual focus. Firstly, it aimed to delve into students' perceptions of AI authoring in terms of various dimensions, such as people, places, activities, technical media, objects, electronic technologies, and emotional attitudes, as depicted through AI images. Secondly, it intended to collect feedback from students engaging with AI technology content within their learning experiences. Through this exploration, the study aspires to shed light on the potential synergies between AI creation and ICT-assisted Mandarin language teaching.

## 2. Literature review

## 2.1 The conception of ICT-assisted Mandarin language teaching.

The concept of "ICT-assisted Mandarin teaching" entails integrating digital tools, platforms, and resources to enhance the quality of Mandarin language instruction (Čok, 2016). The appropriate use of ICT in teaching can create student-centered learning, and the role of the teacher changes to that of a navigator and co-learner, enabling diverse learners while meeting needs (Keengwe et al., 2008; Coleman et al., 2016).

ICT has now expanded to many types, such as "interactive whiteboard", "camera, laptop computer", "VLE (Virtual Learning Environment)", "Blog", "PC", "Facebook", "YouTube", "Projectors", "podcasts", "various software", etc. are all in this range (Yang, 2019). For example, Tan & Liu (2016) noted that ICT is strongly supported by the government in Singapore's Mandarin language teaching. The students can utilize Mandarin learning materials, software, and online translation tools to cultivate written language skills. They engage with platforms to create and share assignments, establish class websites and personalized e-learning profiles, and collaboratively build media content, with teachers taking on supervisory and facilitating roles. Additionally, many studies have shown that the incorporation of ICT in language teaching has many benefits, such as increasing engagement and active participation in learning, fostering learner interest and motivation, and expanding access to target languages (Golonka et al., 2014). However, there are still many challenges, such as students do not have computers at home, and are not selfdisciplined. Some schools are still in the early stages of online teaching, and teachers are not familiar with ICT operations, account management, and student privacy (Dedja, 2015; Le Thi. 2020).

The impact of ICT-assisted teaching is transformative, influencing the pedagogical insights and language competencies of language educators. ICT provides abundant learning resources, enabling more diverse learning experiences where students can adjust according to their learning needs and pace, while also fostering their digital literacy and self-management skills. Therefore, this study focus on students' perspectives on ICT-assisted Mandarin language teaching to further comprehend their conceptual differences in the actual learning contexts of ICT-assisted Mandarin language teaching, 2022).

## 2.2 Drawing as a research approach

Generative AI, a subset of artificial intelligence, utilizes algorithms like generative adversarial networks (GANs) and variational autoencoders (VAEs) to produce diverse content, such as images and text, applicable in fields like education, advertising, entertainment, and digital art (Kaplan-Rakowski, Grotewold, Hartwick, Papin; 2023). Drawing functions as a valuable research tool in educational studies, providing an alternative avenue to access participants' perspectives. It vividly illustrates their conceptualizations, bridging gaps in verbal communication (Ehrlén, 2009). In this context, drawing complements AI technology by harmoniously integrating visual and textual elements, allowing for the easy sharing of

thoughts and efficient gathering of valuable insights (Nuora et al., 2019; Lai, 2021). Scholars in educational research not only stress drawing's role in nurturing students' thinking and enabling open expression but also recognize the potential of combining it with AI technology to enhance learning (Liou, 2017). For example, Lai (2021) used drawing and co-word analysis to explore AI-assisted learning preferences. Participants described their perceptions of Al images and words. The results showed the importance of robots in Al-assisted learning and what expect of students from professional teaching in AI education. In addition, the differences in students' conceptual cognition are analyzed through the drawing analysis method (Tu et al., 2021), for example, Tu et al. (2021) studied drawings of 156 undergraduate students with textual descriptions illustrating their perceptions of a smart library. The results showed that the similarities and differences in library and information science (LIS) students prioritize technology-assisted patron services, while non-LIS students emphasize accessible resources, guiet reading, and freedom of library activity. Briefly, many studies have showcased that drawing analysis results are a valuable means to bolster learning. The research into the drawing approach aligns with the contemporary trend of technological tools, offering a means to analyze multifaceted aspects of pertinent learning concepts.

In conclusion, based on the literature review, AI creation techniques are beneficial in enhancing students' creativity in learning Mandarin and facilitating self-expression and communication through artistic creation. Additionally, the teachers can utilize AI creation techniques to generate specific pictures for language practice and assess students' proficiency (Marrone, Taddeo & Hill, 2022). To assist Mandarin teachers in developing more effective teaching strategies to enhance learners' learning.

# 3. Research Questions

The study aims to answer the following research questions: What is the concept of utilizing AI creation and ICT-assisted Mandarin teaching as a means of providing technological support to students? What feedback did the students provide regarding their learning experience with AI technology?

# 4. Study design

## 4.1 Participants and the data collection procedure

This study focuses on the concepts of AI creation and ICT-assisted Mandarin teaching for students, and students' feedback on AI technology learning. Twenty-one students participated in this study, including 4 males and 17 females. The range of age is 22 to 51 years old, with an average age of 28.

Before operating AI creation, participants participated in 3 hours of AI drawing technology courses for 2 weeks, totaling 6 hours. In the first week, the development, trends and teaching applications of AI drawing technology are introduced. In the second week, participants are asked to generate images based on prompts. Following this, they received a prompt. The prompt content asked participants to draw a "real" classroom scenario. With the support of AI creation techniques, Mandarin teachers use technology to assist teaching when preparing, teaching, guiding students, or participating in activities. Then, participants were also asked to describe the AI creation in 50 words or less, conveying what they hoped to express through their images. Finally, students are asked to give feedback on the activity.

## 4.2 Data analysis

### 4.2.1 Development of the coding scheme

To understand how students use drawing skills to represent real scenes in the classroom, this study developed a coding scheme to analyze the characteristics of students' images.

In the first step, the researchers refer to the encoding schemes proposed by scholars Hsieh and Tsai (2017), Utley & Redmond-Sanogo (2020), and Lai (2021) to design an encoding table that conforms to the drawing. In the second step, two experts in technology-enhanced language learning and an expert in AI creation research make adjustments and modifications. The third step is to encode and label the features of the generated image one by one. Through coding, images can be quantitatively analyzed to present a more complete and comprehensive analysis.

Table 1 presents the coding scheme, which consists of 7 categories and 31 subcategories: people involved, locations, activities, technical media, objects, electronic technologies, and emotions and attitudes.

Categories	Sub-categories	
1. People involved	1.1 Teacher	1.3 Peers/groups (student)
	1.2 Learner	1.4 Unspecified
2. Locations	2.1 Classroom	2.3 Unspecified
	2.2 Online	
3. Activities	3.1 Teacher lecturing	3.5 Paperwork documenting
	3.2 Student performing	3.6 Hands-on practicing
	3.3 Student observing	3.7 Resource searching
	3.4 Group working	3.8 Assessment evaluating
4. Technical media to	4.1 VR / AR	4.3 Web-based platform
support teaching/learning?	4.2 Technology devices	4.4 Unspecified
5. Objects	5.1 Technological devices	5.3 Book/handouts
	5.2 Traditional classroom	
	equipment	
6. Electronic technologies	6.1 Computer	6.4 Touch screen
	6.2 Screens	6.5 VR glasses
	6.3 Mobile devices	6.6 Unspecified
<ol><li>Emotions and attitudes</li></ol>	7.1 Positive	7.3 Unspecified
	7.2 Negative	

Table 1. Coding scheme for coding participants' AI creation images

## 4.2.2 Data analysis procedure

The study was analyzed in two parts: the AI creation images and their accompanying descriptions. A coding scheme with 7 dimensions was employed to identify specific attributes present or absent in the images. Each image could have multiple attributes, and the same index in an image was counted only once for the parts being assessed. Figure 1 provides a coded example of a participant's AI images, in which a classroom scenario with four students are situated. The context portrays an online teaching environment, with the teacher conducting a lecture virtually. The students are actively engaged in performing a task, while the scenario involves the use of a QR code on the whiteboard. They engaged in performing the task is utilizing their phone to scan the QR code, illustrating a hands-on approach to practicing the technology.



# *Figure 1.* A coding example of a participant's conception of AI creation image in Mandarin teaching.

## 5. Results

Referring to the coding scheme in Table 1, the AI creation images by 21 students were analyzed. Table 2 shows the frequencies and percentages of the coding results. Among the 21 AI creation images, a total of 234 encodings were identified, with the distribution of proportions ranked from high to low as follows: activities (23.93%), people involved (17.95%), objects (17.09%), locations (10.68%), electronic technologies (11.54%), technical media (9.83%), and emotions and attitudes (8.97%).

	Frequency	Percentage (% = N/21)
People involved		
Teacher	20	95.24%
Learner	16	76.19%
Peers/groups (student)	4	19.05%
Unspecified	2	9.52%
Locations		
Classroom	18	85.71%
Online	5	23.81%
Unspecified	2	9.52%
Activities		
Teacher lecturing	16	76.19%
Student performing	7	33.33%
Student observing	4	19.05%
Group working	5	23.81%
Paperwork documenting	9	42.86%
Hands-on practicing	10	47.62%
Resource searching	2	9.52%
Assessment evaluating	3	14.29%
Technical media to support teaching/learning?		
VR / AR	1	4.76%
Technology devices	14	66.67%
Web-based platform	1	4.76%
Unspecified	7	33.33%
Objects		
Technological devices	14	66.67%
Traditional classroom equipment	10	47.62%
Book/handouts	16	76.19%
Electronic technologies		
Computer	8	38.10%
Screens	7	33.33%
Mobile devices	3	14.29%
Touch screen	3	14.29%
VR glasses	1	4.76%
Unspecified	5	23.81%

 Table 2. Coding scheme for AI creation based on participants prompts

Emotions and attitudes			
Positive	13	61.90%	
Negative	2	9.52%	
Unspecified	6	28.57%	

In Table 2, overall, regarding the people involved, teachers account for the highest proportion, at 95.24%, and there are 2 unspecified images (9.52%) showing a cat and a robot. This demonstrates that despite the involvement of technology in the teaching process, teachers still play a crucial role in language teaching. The majority of locations represents classrooms (85.71%). In the activity part, the teacher lecturing is the predominant theme (76.19%), followed by hands-on practicing (47.62%) and paperwork documenting (42.86%). The drawing on the left of Figure 2, for the "people involved", shows that the teacher is providing a lecture, and students are engaging in completing paperwork, while on the right of Figure 2, for the "activities", depicts students doing hands-on practice and discussing assignment content.



Figure 2. Example of the AI creation in "People involved" and "Activities" categories.

The technology devices, including mobile phones and tablets, are the primary means of supporting teaching (66.17%), and there is an AI creation of a student showing the use of VR and AR teaching equipment. This reflects the student's own experience with technology devices. The picture on the left of Figure 3 shows that the students are seated in the classroom, using laptops and iPads, while the teacher uses a projector to display slides. The objects of books/handouts are most prevalent (76.19%), followed by technological devices (66.67%) and traditional classroom equipment (47.62%). Regarding the "objects", the students engage in Mandarin class with a cheerful teacher. Some students listen attentively to audio, while others are engrossed in reading digital books on iPads (see the picture on the right of Figure 3).



Figure 3. Example of the AI creation in "Technology devices" and "Objects" categories.

In the category of electronic technologies, computers (38.10%) and screens (33.33%) are the main features, with one creation depicting students wearing VR glasses during the class. In terms of emotions and attitudes, a positive attitude was prevalent (61.90%), and 6 AI images did not specify the characters' facial expressions or appearances. On the left of Figure 4, depicts four students seated in a classroom and all wearing VR glasses. The teacher is engaged in an online class. In the prompt of the AI creation generated by the students, the QR code pattern should be presented on the whiteboard, but it is not shown. It can be seen that the current AI creation technology has not yet presented the barcode. Then, it is difficult to judge the students whether emotions and attitudes are positive or negative because they are wearing the device. On the right of Figure 4, the students are actively engaged with computers, employing them for paperwork tasks, and the teacher guides by the side.



Figure 4. Example of the AI creation in "Electronic technologies" categories.

On the dimension of analyzing students' interviews, most students expressed enthusiasm for using AI tools to motivate learners, one participant particularly mentioned that their introverted students may be more motivated by such GAI technologies.. Through hands-on activities, they gained comfort and understanding of new technology, especially using AI creation images to match learning topics. While students acknowledge the benefits of AI-generated creations, they also recognize the importance of addressing copyright and privacy issues. Additionally, they have observed that certain details are missing during image generation, images appear cluttered, and there are instances of unnecessary hands and actions. In conclusion, the students found AI creation techniques in education a future trend. AI creation helped improve creativity, offering diverse teaching materials, capturing learners' attention, and fostering a positive attitude and motivation toward language learning.

# 6. Discussions and conclusions

According to the coding and analysis of the drawing of AI creation in this study, it is found that the concept and cognition of ICT-assisted Mandarin teaching are multifaceted, and through interviews, the participants expressed that it is interesting and valuable for learning about AI technology applications. Among the 7 dimensions of AI creation technology, "activities" (23.93%) generated the most dimensions, followed by "people involved" (17.95%), and finally "technical media to support teaching/learning?" (9.83%) and "emotions and attitudes" (8.97%). In other words, when AI creation presents the Mandarin teaching picture, it is more able to present the "teacher lecture" and "hands-on practice" drawing according to the prompt, and it can also generate "books" and "traditional classroom equipment". However, in the aspect of "technical media" there is a challenge in accurately representing elements like barcodes and technological devices, leading to misalignment and blurring of the visual depiction. Similarly, within the category of "emotions and attitudes," certain facial expressions and the overall teaching atmosphere might not be clearly conveyed.

With the continuous emergence of artificial intelligence technology, teaching must become more intelligent and able to meet the different needs of students. Through the

development of the coding scheme, we can understand students' current understanding of concepts and provide teachers with a reference direction for combining teaching content. It also shows what AI creation can bring to education. Despite the discovery of AI-created images during the implementation process, there are still many biases and privacy challenges. From a pedagogical perspective, however, Al-authored technologies offer opportunities for diverse, fun, and engaging teaching experiences. The efficacy of incorporating AI creation into the classroom and its potential to truly improve learning outcomes depends on how proficient teachers are in its application and design.

# References

- Abraham, M., Arficho, Z., & Habtemariam, T. (2022). Effects of ICT-assisted teaching training on teachers 'skills of integrating ICT tools in teaching English and their attitude towards using ICT tools for teaching English. The Ethiopian Journal of Social Sciences and Language Studies (EJSSLS), 9(1), 33-52. https://doi.org/10.1080/2331186X.2022.2028336
- Chiang, F. K., Li, L., Cai, R. H., & Wang, S. (2020). Investigation of elementary-school students' perception of engineering using drawing analysis. International Journal of Engineering Education, 36(1), 241-255.
- Čok, T. (2016). ICT-supported language learning tools for Chinese as a foreign language. Journal of Elementary Education, 9(3), 103-120.
- Coleman, L. O., Gibson, P., Cotten, S. R., Howell-Moroney, M., & Stringer, K. (2016). Integrating computing across the curriculum: The impact of internal barriers and training intensity on computer integration in the elementary school classroom. Journal of Educational Computing Research, 54(2), 275-294. https://doi.org/10.1177/07356331156166

Dedia, M. (2015). ICT in foreign language teaching and learning: Benefits and challenges. European Journal of Language and Literature, 1(2), 42-47. https://doi.org/10.26417/ejls.v2i1.p42-47

- Ehrlén, K. (2009). Drawings as representations of children's conceptions. International Journal of Science Education, 31(1), 41-57. https://doi.org/10.1080/09500690701630455
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. Computer assisted language learning, 27(1), 70-105. https://doi.org/10.1080/09588221.2012.700315

Hsieh, W. M., & Tsai, C. C. (2018). Learning illustrated: An exploratory cross-sectional drawing analysis of students' conceptions of learning. The Journal of Educational Research, 111(2), 139-150. https://doi.org/10.1080/00220671.2016.1220357

- Keengwe, J., Onchwari, G., & Wachira, P. (2008). The use of computer tools to support meaningful learning. AACE Journal, 16(1), 77-92.
- Lai, C.L. (2021). Exploring University Students' Preferences for AI-Assisted Learning Environment: A Drawing Analysis with Activity Theory Framework. Educational Technology & Society, 24(4), 1-15.
- Le Thi, M. A. I. (2020). Benefits and challenges to integrate ICT in EFL teaching and learning activities. Journal of Research & Method in Education (IOSR-JRME), 10(3), 46-50. https://doi.org/10.9790/7388-1003044650
- Liou, C.L. (2017). A comparative study of undergraduates' attitudes toward aging in Taiwan and the United States through student drawings. The International Journal of Aging and Human Development, 85(3), 265-288. https://doi.org/10.1177/0091415017702906
- Marrone, R., Taddeo, V., & Hill, G. (2022). Creativity and artificial intelligence-A student perspective. Journal of Intelligence, 10(3), 65. https://doi.org/10.3390/jintelligence10030065
- Nuora, P., Välisaari, J., & Kiviniemi, T. (2019). Adolescents' perception of scientific inquiry in nature: A Drawing analysis. Nordina, 15(3), 313-327. https://doi.org/10.5617/nordina.6439
- Tan, C., & Liu, M. (2016). Using ICT in teaching the Chinese language: Practices and reflections from Singapore. Teaching Chinese Language in Singapore: Retrospect and Challenges, 45-63. https://doi.org/10.1007/978-981-10-0123-9 4
- Tu, Y. F., Hwang, G. J., Chen, S. Y., Lai, C., & Chen, C. M. (2021). Differences between LIS and non-LIS undergraduates' conceptions of smart libraries: a drawing analysis approach. The Electronic Library, 39(6), 801-823. https://doi.org/10.1108/EL-07-2021-0129

- Utley, J., Reeder, S., & Redmond-Sanogo, A. (2020). Envisioning my mathematics classroom: Validating the Draw-a-Mathematics-Teacher-Test rubric. *School Science and Mathematics*, *120*(6), 345-355. https://doi.org/10.1111/ssm.12426
- Xu, Z., Ritzhaupt, A. D., Umapathy, K., Ning, Y., & Tsai, C. C. (2020). Exploring college students' conceptions of learning computer science: A Draw-a-picture technique study. Computer Science Education, 31(1), 60-82. https://doi.org/10.1080/08993408.2020.1783155
- Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., & Papin, K. (2023). Generative AI and Teachers' Perspectives on Its Implementation in Education. *Journal of Interactive Learning Research*, *34*(2), 313-338. https://www.learntechlib.org/primary/p/222363/
- Yang, X. (2019). A critical review and analysis: what role of ICT plays in improving professional teachers' English instruction? *Advances in Language and Literary Studies*, *10*(2), 129-135. https://doi.org/10.7575/aiac.alls.v.10n.2p.129

# Tools and Approaches of Generative Artificial Intelligence Used in Education

## Hsin-Hsuan CHUNG<sup>a</sup>, Fu-Ling CHUNG<sup>b\*</sup>, Shu-Min LIN<sup>c</sup> & Yu-Ju LAN<sup>d</sup>

<sup>a</sup>Department of Information Science, University of North Texas, USA <sup>b</sup>Department of Learning Technologies, University of North Texas, USA <sup>c</sup>Department of Education and Futures Design, Tamkang University, Taiwan <sup>d</sup>Department of Chinese as a Second Language, National Taiwan Normal University, Taiwan \* fu-lingchung@my.unt.edu

Abstract: This study aims to help scholars comprehend the common generative artificial intelligence (GAI) tools and approaches in the education field and to propose several potential future research directions to support scholars to make up the gap in the field. This study applied a systematic review approach to conduct content analysis and answer the research questions. We established several inclusion and exclusion criteria to identify 15 target articles from the Scopus database. The target articles should be empirical studies that discuss at least one specific GAI tool. These articles were published from 2021 to 2023 and belong to different disciplines, such as social science, computer science, and engineering. The results show that the most discussed GAI tool in education research is ChatGPT, and the most discussed GAI research method in education research is case study. Moreover, we found that ChatGPT is popular and has plenty of benefits, so educators are interested in exploring its ability, potential, and further applications. We also found that when the research objects were unusual or multiple research objects, researchers may choose to apply a case study approach. Future research can focus more on (1) conducting experiments to know what GAI tools can do for students and (2) teaching students how to use GAI to support their learning and help them recognize the concerns of using these tools.

Keywords: Generative artificial intelligence, systematic review, education, ChatGPT

## 1. Introduction

As many people know, the influence of Artificial Intelligence (AI) has been profound on organizations, communities, and individuals (Dwivedi et al., 2023). Generative artificial intelligence (GAI) is a type of AI technology, and it employs distinctive algorithms, machine learning, deep learning models, or natural language processing (NLP) to produce content resembling that created by humans, such as text, images, and music (Hsu & Ching, 2023; Lim et al., 2023). This content is generated in response to diverse and intricate prompts, which can encompass languages, instructions, and questions (Lim et al., 2023). GAI models acquire the patterns and arrangement of their input training data and subsequently produce new data with comparable attributes (Dwivedi et al., 2023). It is important to highlight that GAI possesses a distinct capability to not just offer a reply but also fabricate the content within that reply, surpassing the human-like interactions seen in Conversational AI (Lim et al., 2023).

An outstanding illustration of GAI is ChatGPT, which achieved remarkable acclaim following its public debut in November 2022, drawing in a million users within five days and amassing 100 million users in a mere two months (Hsu & Ching, 2023). The recent global embrace of ChatGPT on a large scale has showcased the extensive spectrum of applications for this technology, spanning software development, testing, creative writing like poetry and essays, composing business letters, and even drafting contracts (Dwivedi et al., 2023). Beyond text generation, the realm of GAI is witnessing a growing array of distinct products and services that utilize its power to craft diverse content forms and cater to various

objectives (Hsu & Ching, 2023). Leonardo.ai empowers users to generate distinct game assets ready for production, encompassing characters, props, textures, and environments, thereby aiding game developers in time and cost savings; Midjourney offers users the capacity to produce lifelike images from textual cues, allowing for the flexibility to experiment with artistic styles such as realism or abstraction during image creation; OpenAI's DALL-E, a text-to-image AI creator, facilitates modifications to specific portions of an image in response to supplementary natural language text inputs (Hsu & Ching, 2023). Besides, Music GAIs such as Amper Music and MuseNet also exist, employing algorithms to examine preexisting music and produce compositions across diverse genres like jazz, pop, and classical (Hsu & Ching, 2023).

People can see multiple applications of GAI in the field of education. This is because AI systems aim to foster the growth of learners' abilities in reflective and critical thinking, creativity, problem-solving, and grasping complex concepts (Vasconcelos & Santos, 2023). Viewed through a constructionist perspective, ChatGPT and Bing Chat, serving as tools for thought, present exciting opportunities to transform science, technology, engineering, and mathematics (STEM) education by promoting participation in inclusive and easily accessible learning settings (Vasconcelos & Santos, 2023). Moreover, ChatGPT is integrated into programming education, and the research results underscored the optimal utilization of AI assistance within lessons and offered diverse recommendations for researchers and educators (Yilmaz & Karaoglan Yilmaz, 2023). Besides, Vartiainen and Tedre (2023) conducted a practical workshop focused on creative production using text-to-image GAI to pre-service craft teachers and teacher educators, and the outcomes indicated that engaging in creative work with AI prompted educators to contemplate the distinct characteristics of craftsmanship and the challenges and compromises involved in incorporating GAI into craft-related activities.

However, this has also triggered various concerns related to the difficulty of discerning between human and AI authorship within academic and educational contexts, rekindling debates about the importance of traditional human contributions (Dwivedi et al., 2023). Furthermore, given its reputation for often being seen as resistant to change, it is not unexpected that many educators express reservations about GAI, particularly regarding matters of evaluation and ethical considerations like authenticity and plagiarism (Lim et al., 2023). With the rapid rise of GAI as a transformative force, education faces a golden opportunity for reimagining its future (Lim et al., 2023). In other words, educators are encouraged to understand GAI and embrace its potential rather than avoid it. Therefore, this study aims to answer the following research questions:

1. What are the characteristics of existing literature on GAI in education?

- 2. What GAI tools are used in education?
- 3. How did researchers use GAI tools in education?

This study centered on empirical studies which actually used at least a specific GAI tool in the education setting. Thus, conceptual articles, review papers, and algorithm/model articles were out of the scope of this study. Nevertheless, these types of articles are worth an independent analysis to advance the field of GAI. The following sections briefly explain the research method, display results, and discuss study implications and conclusions.

### 2. Methods

This study applied a systematic review approach to conduct content analysis and answer the research questions. Since systematic review is an approach that seeks to create a comprehensive, unbiased synthesis of multiple related studies within a single document (Aromataris & Pearson, 2014), and content analysis serves as a technique that helps classify data into defined groups, making it easier and more effective to understand and analyze the data (Harwood & Garry, 2003), so both methods support this study as a comprehensive, unbiased, and easy-to-read article.

As Figure 1 shown, we used the search keywords ("Generative Artificial Intelligence" or "Generative AI") and "education" to look for related literature in the Scopus database.

Scopus has a user-friendly interface that facilitates conducting systematic reviews, and it is one of the well-established bibliographic databases (Caputo & Kargina, 2022). Then, the searches retrieved 373 articles by August 05, 2023. The articles whose titles, abstracts, and keywords did not contain the search keywords were removed from the main interest dataset, and 77 articles remained. In addition, several criteria were utilized to evaluate the articles' eligibility to make the research quality consistent. The eligible articles should be open access, the document type should be Article, and the language should be English, and 38 articles left.



Figure 1. The Research Design.

Because this study focused on exploring the study clearly conducting research on specific GAI tools in the educational setting, we manually reviewed each article with several criteria to finalize the target articles. Table 1 demonstrates the inclusion and exclusion criteria for target articles. Target articles should be empirical studies that relate to GAI and point out at least one specific GAI tool. Finally, only 15 articles, or 39% of 38 articles, were eligible as our dataset for analysis.

Criteria target articles	for	Articles included	Articles excluded	Number articles excluded	of
Туре		Empirical study	Conceptual/ Opinion article	15	
			Review paper	4	
			Algorithm/ model article	2	
Related GAI	to	Strong	Weak	1	
Specific tool	GAI	Yes	No	1	

#### Table 1. Article Selection Criteria

# 3. Result

To answer Research Question 1, this section first presents the characteristics of the dataset, including Paper Distribution Over Time and Publication Journal of the 15 selected

articles. Then, we focus on identifying tools and approaches of GAI used by these studies to answer Research Questions 2 and 3.

## 3.1 Characteristics of the Selected Articles

### 3.1.1 Paper Distribution Over Time

The final 15 articles were published from 2021 to 2023. As Figure 2 shows, the result indicates that 93% of 15 articles were published in 2023 (n=14), and 7% were in 2021 (n=1). That is, research on GAI in the education field is still in the beginning stage, but it has attracted a lot of attention from researchers in different disciplines and increased exponentially.



Figure 2. Paper Distribution Over Time.

### 3.1.2 Publication Journal

As Table 2 shows, the 15 articles were published in 13 different journals. The Computers and Education: Artificial Intelligence and International Journal of Educational Technology in Higher Education each rank first with 2 articles, and each other journal has 1 article. Additionally, we applied Scimago Journal & Country Rank to identify the subject area of each journal. The results showed that the 13 different journals were in 7 subject areas. Social Sciences ranks first with 10 articles, followed by Computer Science (n=4) and Engineering (n=2).

Journal	Subject Area	Number of Articles
Computers and Education: Artificial Intelligence	Computer Science, Social Sciences	2
International Journal of Educational Technology in Higher Education	Computer Science, Social Sciences	2
Digital Creativity	Arts and Humanities, Computer Science	1
Educational Process: International Journal	Social Sciences	1
Eurasia Journal of Mathematics, Science and Technology Education	Mathematics, Social Sciences	1
European Journal of Engineering Education	Engineering, Social Sciences	1

Table 2. Publication Journal for 15 Articles

Higher Education Policy	Social Sciences	1
Journal of Applied Learning and	Social Sciences	1
Teaching		
Medical Education Online	Medicine,	1
	Social Sciences	
Quality Engineering	Engineering	1
Smart Learning Environments	Computer Science,	1
	Social Sciences	
STEM Education	Social Sciences	1
Sustainability (Switzerland)	Multifield	1

### 3.2 GAI Tool of the Selected Articles

As shown in Figure 3 and Table 3, in terms of the GAI tools that were studied, 86% of the 15 articles were on ChatGPT (n=13), 7% on Synthetic media (n=1), and 7% on Text-to-image generative models (n=1).



Figure 3. Frequency of Used GAI Tools.

Table 3. GAI Tool and Research Mer	thod of 15 Articles
------------------------------------	---------------------

Authors	GAI tool	Research method
Ali et al. (2021)	Synthetic media	Survey
Chan (2023)	ChatGPT	Survey
Chan and Hu (2023)	ChatGPT	Survey
Elkhodr et al. (2023)	ChatGPT	Case study
Friederichs et al. (2023)	ChatGPT	Case study
Limna et al. (2023)	ChatGPT	Interviews
Megahed et al. (2023)	ChatGPT	Case study
Nikolic et al. (2023)	ChatGPT	Case study
Perkins and Roe (2023)	ChatGPT	Case study
Tlili et al. (2023)	ChatGPT	Case study
Tülübaş et al. (2023)	ChatGPT	Comparative analysis
		of responses
Uddin et al. (2023)	ChatGPT	Experiment
Vartiainen and Tedre	Text-to-image	Case study
(2023)	generative models	
Vasconcelos and	- ChatGPT	Case study
Santos (2023)	- Bing Chat	-
Yilmaz and Karaoglan	ChatGPT	Experiment
Yilmaz (2023)		

3.3 Research Method and Approaches of How Researchers Used GAI Tools in

#### Education

As shown in Figure 4 and Table 3, in terms of research method, 54% of the 15 articles adopted Case study (n=8), 20% with Survey (n=3), 13% with Experiment (n=2), and 13% with Others (n=2). According to the results, the approaches that researchers used GAI tools in education are as follows.



Figure 4. Frequency of Used Research Method.

#### 3.3.1 Case Study (n=8)

A hands-on tutorial experiment was executed to assess the influence of employing ChatGPT as an educational aid in a classroom setting, in contrast to utilizing search engines such as Google (Elkhodr et al., 2023). In order to gauge the factual knowledge proficiency of the responses from ChatGPT, Friederichs et al. (2023) assessed how ChatGPT performed in comparison to medical students taking a progress test. Megahed et al. (2023) investigated the capability of ChatGPT in offering code, elucidating fundamental concepts, and generating knowledge relevant to the realms of Statistical Process Control (SPC) practice, education, and research. Nikolic et al. (2023) investigated the potential impact of ChatGPT on assessment within engineering education through an analysis of ChatGPT's reactions to established assessment prompts from ten courses spanning seven Australian universities. Perkins and Roe (2023) analyzed how universities worldwide address new technological threats like Automated Paraphrasing Tools and GAI tools, such as ChatGPT, in their academic integrity policies. Tlili et al. (2023) explored the role of ChatGPT in education among its initial users, utilizing a qualitative instrumental case study approach. Vartiainen and Tedre (2023) investigated the perspectives of 15 Finnish pre-service craft teachers and teacher educators regarding the possible advantages and hurdles associated with AI. specifically focusing on text-to-image GAI. The research involved a practical workshop that involved hands-on creative work using text-to-image GAI, aiming to spark discussions and gather perceptions about GAI. Vasconcelos and Santos (2023) explored the capabilities of ChatGPT and Bing Chat, sophisticated conversational Ais. These resources are designed to promote reflective and analytical thinking, as well as conceptual understanding, in the context of enhancing STEM education.

### 3.3.2 Survey (n=3)

Ali et al. (2021) familiarized middle school students with GAI methods as a creative tool while also emphasizing thoughtful deliberation about their social and ethical

consequences. Their finding indicated that students exhibited comprehension that generative media could appear credible, yet not necessarily accurate, and may contribute to the dissemination of misinformation. Chan (2023) used an online survey that encompassed the application of GAI technologies such as ChatGPT, the incorporation of AI in higher education, potential associated risks, and the influence of AI on teaching and learning to gather data from students, teachers, and staff in higher education. Chan and Hu (2023) delved into the perspectives of 399 university students regarding GAI technologies, like ChatGPT, within higher education. Their study centered on aspects such as familiarity, willingness to participate, potential advantages and obstacles, as well as optimal incorporation.

### 3.3.3 Experiment (n=2)

Uddin et al. (2023) assessed whether the integration of ChatGPT can assist in hazard recognition when incorporated into the curriculum for students preparing for a career in the construction sector. Yilmaz and Karaoglan Yilmaz (2023) explored the impact of incorporating ChatGPT into programming education on students' computational thinking skills, self-confidence in programming, and motivation for the lesson.

#### 3.3.4 Others (n=2)

Limna et al. (2023) investigated educators' and students' perspectives concerning the utilization of ChatGPT in education within the digital age, employing comprehensive interviews as a means to collect data. Tülübaş et al. (2023) examined how ChatGPT, a recent advancement in AI technology, reacts to our inquiry about emergency remote teaching (ERT). This was done to assess the precision, comprehensibility, and succinctness of the information it provides, as well as to identify any potential bias.

### 4. Discussion

This review study examines the common GAI tools and approaches in the education area. Our analysis provided answers to our research questions. Our review found the GAI research raised from 2021, and the coming out of ChatGPT in 2022 led to a trend in GAI research. Most of our 15 articles were published in the fields of social sciences with ChatGPT in 2023. The possible explanation of why educators prefer to choose ChatGPT to conduct research is that ChatGPT has multiple benefits. For example, ChatGPT can (1) help higher education institutions formulate an AI education policy (Chan, 2023), (2) enrich students' learning journeys by leveraging its capacity to generate exceptionally unique output in response to user prompts (Chan & Hu, 2023), (3) offers convenient entry to an extensive spectrum of information (Friederichs et al., 2023), (4) possesses the capability to transform the dynamics of interaction and learning between students and educators (Limna et al., 2023), (5) enhance the efficiency and productivity of educators, researchers, and practitioners (Megahed et al., 2023), and (6) be among the AI technologies that have begun to find applications in programming education (Yilmaz & Karaoglan Yilmaz, 2023). These advantages are as suggested by Hsu and Ching (2023), including (1) assisting in teaching tasks, (2) aiding in student assessment, (3) facilitating student learning, (4) providing recommendations for enhancing teaching methods, and (5) enabling communication between teachers and both parents and students. ChatGPT can also support creative thinking, personalized learning, assessment, and reading and writing comprehension (Hsu & Ching, 2023).

Based on these benefits of ChatGPT, most of the 15 articles adopted a case study approach. The possible explanation of why the authors applied the case study approach is that they conducted an in-depth analysis of unique situations. As Noor (2008) suggested case studies offer researchers a comprehensive perspective of a particular phenomenon or a sequence of events, offering a well-rounded understanding due to the utilization of various sources of evidence. Furthermore, case studies prove valuable in capturing the evolving and inherent aspects of organizational life, including the dynamic nature of organizational activities, particularly in rapidly changing environments. Additionally, case studies permit the derivation of generalizations, as findings from multiple cases can potentially lead to some form of reproducibility.

Despite the immense promise and opportunities offered by GAI, recent trends and discussions surrounding its utilization in education have brought to light certain issues and apprehensions, such as information accuracy, data privacy, bias, discrimination, and ethical issues (Hsu & Ching, 2023). According to Pavlik (2023), a primary ethical concern stemming from GAI and its potential societal implications revolves around the issue of responsibility. Given that GAI systems can generate novel content like text, audio, or visuals, there is a worry that they might be employed to produce deceptive or harmful information that is challenging for humans to differentiate from authentic content. This has the potential to propagate misinformation and other detrimental content, with significant repercussions for individuals and society. Moreover, take ChatGPT, for example. As Lin et al. (2023) indicated, ethics, plagiarism, and academic integrity are three major educational concerns when using ChatGPT in education.

In the same manner, there are weaknesses in using case study method. Critics have pointed out that case studies are often seen as lacking in scientific precision and dependability, and they may not effectively tackle concerns related to applicability across broader contexts (Noor, 2008). According to Yin (2009), case studies have often been considered a less preferred research method when compared to experiments or surveys. One significant issue that has arisen pertains to the perceived requirement for increased precision in conducting case study research. On numerous occasions, researchers employing case studies have been criticized for being careless, not adhering to systematic protocols, or letting uncertain evidence sway the outcomes and interpretations of their research.

#### 5. Conclusion

This study applied a systematic review approach to conduct content analysis and answer the research questions. Through a systematic approach, we selected 15 articles, which are empirical studies that discuss at least one specific GAI tool, from the Scopus database. Then, we focused on exploring the common GAI tools and approaches in education. The results show that the most discussed GAI tool in education research was ChatGPT, followed by synthetic media and text-to-image generative models. The most discussed GAI research method in education research was case study, followed by survey and experiment. Therefore, we took a closer examination of why researchers prefer to choose ChatGPT and case study to conduct research.

We found that ChatGPT is popular and has plenty of benefits, including formulating AI education policies, enriching students' learning journeys, offering abundant information, and enhancing the efficiency and productivity of educators, researchers, and practitioners. Thus, educators are interested in exploring its ability, potential, and further applications. Furthermore, we found that when researchers conduct research in unique situations, they may apply a case study approach. For example, the research objects were unusual or multiple research objects.

Since only 2 of the 15 selected articles conducted experiments, future research can focus more on conducting experiments to know what GAI tools can do for students, including advantages and disadvantages. Experiments may enable a comprehensive understanding of the impact of these tools on the learning process, effectiveness, and learner involvement. Also, future research can focus on teaching students how to use GAI to support their learning and help them recognize the concerns of using these tools.

### References

- Ali, S., DiPaola, D., Lee, I., Sindato, V., Kim, G., Blumofe, R., & Breazeal, C. (2021). Children as creators, thinkers and citizens in an Al-driven future. *Computers and Education: Artificial Intelligence*, 2, 100040.
- Aromataris, E., & Pearson, A. (2014). The systematic review: an overview. AJN The American Journal of Nursing, 114(3), 53-58.
- Caputo, A., & Kargina, M. (2022). A user-friendly method to merge Scopus and Web of Science data during bibliometric analysis. *Journal of Marketing Analytics*, *10*(1), 82-88.
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), 1-25.
- Chan, C. K. Y., & Hu, W. (2023). Students' Voices on Generative AI: Perceptions, Benefits, and Challenges in Higher Education. *arXiv preprint arXiv:2305.00290*.
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642.
- Elkhodr, M., Gide, E., Wu, R., & Darwish, O. (2023). ICT students' perceptions towards ChatGPT: An experimental reflective lab analysis. *STEM Education*, *3*(2), 70-88.
- Friederichs, H., Friederichs, W. J., & März, M. (2023). ChatGPT in medical school: how successful is Al in progress testing?. *Medical Education Online*, *28*(1), 2220920.
- Harwood, T. G., & Garry, T. (2003). An overview of content analysis. *The marketing review*, *3*(4), 479-498.
- Hsu, Y. C., & Ching, Y. H. (2023). Generative Artificial Intelligence in Education, Part One: the Dynamic Frontier. *TechTrends*, 1-5.
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, *21*(2), 100790.
- Limna, P., Kraiwanit, T., Jangjarat, K., Klayklung, P., & Chocksathaporn, P. (2023). The use of ChatGPT in the digital era: Perspectives on chatbot implementation. *Journal of Applied Learning and Teaching*, 6(1).
- Lin, S. M., Chung, H. H., Chung, F. L., & Lan, Y. J. (2023, August). Concerns About Using ChatGPT in Education. In *International Conference on Innovative Technologies and Learning* (pp. 37-49). Cham: Springer Nature Switzerland.
- Megahed, F. M., Chen, Y. J., Ferris, J. A., Knoth, S., & Jones-Farmer, L. A. (2023). How generative ai models such as chatgpt can be (mis) used in spc practice, education, and research? an exploratory study. *Quality Engineering*, 1-29.
- Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., ... & Sandison, C. (2023). ChatGPT versus engineering education assessment: a multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education*, 1-56.
- Noor, K. B. M. (2008). Case study: A strategic research methodology. American journal of applied sciences, 5(11), 1602-1604.
- Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, 78(1), 84-93.
- Perkins, M., & Roe, J. (2023). Decoding academic integrity policies: A corpus linguistics investigation of AI and other technological threats.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1), 15.
- Tülübaş, T., Demirkol, M., Ozdemir, T. Y., Polat, H., Karakose, T., & Yirci, R. (2023). An interview with ChatGPT on emergency remote teaching: A comparative analysis based on human–Al collaboration. *Educational Process: International Journal*.
- Uddin, S. J., Albert, A., Ovid, A., & Alsharef, A. (2023). Leveraging ChatGPT to Aid Construction Hazard Recognition and Support Safety Education and Training. *Sustainability*, *15*(9), 7121.
- Vartiainen, H., & Tedre, M. (2023). Using artificial intelligence in craft education: crafting with text-toimage generative models. *Digital Creativity*, 34(1), 1-21.
- Vasconcelos, M. A. R., & Santos, R. P. D. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *arXiv preprint arXiv:2305.02202*.

Yilmaz, R., & Karaoglan Yilmaz, F. G. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education. Artificial Intelligence, 4*, 100147.
Yin, R. K. (2009). *Case study research: Design and methods* (Vol. 5). sage.

# Exploring ChatGPT Performance on PISA Multiple Choice Sample Questions Comparing English and Japanese Expression

#### Kyosuke Takami

Education Data Science Center, National Institute for Educational Policy Research, Japan takami@nier.go.jp

**Abstract:** The utilization of generated AI such as ChatGPT in K12 is an urgent issue. However, it is not clear how much the answer ability of ChatGPT is, especially depending on the language differences such as English and Japanese. To address this issue, I examined how well the problem-solving skills of the ChatGPT perform in the context of K12 education using multiple sample test from the global PISA test comparing English and Japanese expression. I found that ChatGPT performed 100% correctly answered in English expression whereas 70% correctly answered in translated Japanese. I also conducted comparison originally Japanese quizzes and those translated English quizzes and found that the correct response rate of the ChatGPT in Japanese is 50%, while translated into English, it indicates 100% correct responses. These results suggest the performance of ChatGPT is inferior performance in languages other than English, such as Japanese, and caution should be exercised when using ChatGPT in K12 educational settings. Also, there is a need to develop and apply large language models in languages other than English.

Keywords: ChatGPT, PISSA test, English and Japanese, generative AI in education

### 1. Introduction

The Chat GPT (OpenAI, 2022), which was introduced in November 2022, has created a sensation by demonstrating unprecedentedly high interactive and generic problem-solving capabilities. Its influence has also spread to education, with some states in the U.S., first restricting its use in educational activities but now permitted (Banks, 2023) and the U.K. Department for Education announcing guidelines for the use of generative AI (DfE, 2023), including ChatGPT, which have caused controversy in the education field. Especially in Japan, elementary, junior high and high schools limited use of generative artificial intelligence (KYODO, 2023). Some of the reason for the limited use would be hallucination with human-like fluency and eloquence on things that are not based on truth (Shen et al., 2023), failing in elementary mathematical questions (Frieder et al., 2023) and a generalpurpose large language model (LLM) trained from everything on the internet web sites, thus its language coverage is dubious (Lu et al., 2022). For instance, ChatGPT-3.5, which is widely and freely used model, have been trained on a diverse set of text data, with a primary focus on the English language. Therefore, it is not well known how well they are able to answer questions on K12 education in languages other than English such as Japanese, Chinese, or Korean. In this study, for the first step, I focused on the English and Japanese expression and examined how well the problem-solving skills of the ChatGPT perform in an educational context using multiple sample test from the global PISA. The research questions is as follows.

RQ: Does the difference between English and Japanese expressions change ChatGPT's performance on K12 level multiple choice questions?

# 2. Related Works

## 2.1 PISA Test

The PISA (Programme for International Student Assessment) test is an internationally recognized assessment conducted by the Organization for Economic Co-operation and Development (OECD). It aims to evaluate and compare the knowledge and skills of 15-year-old students from various countries in reading, mathematics, and science. The test measures students' ability to apply their knowledge to real-life situations, critical thinking, problem-solving, and analytical skills. PISA is conducted every three years, and the results provide valuable insights into the education systems of different countries and help policymakers make informed decisions to improve educational outcomes worldwide. In this study, I adopted sample multiple choice test from OECD PISA Test.

## 2.2 Multiple test peformance by Generative AI

ChatGPT performance on multiple test were mainly examined in the field of medicine such as USA medical license test (Gilson et al., 2023), UK professional qualification test for anaesthetists (Birkett et al., 2023) and dermatologist (Passby et al., 2023). In educational context, multiple choice questions in higher education were used for Chat GPT performance research (Newton & Xiromeriti, 2023), however, to my best knowledge, little research was conducted for K12 education.

## 2.3 Languages used for ChatGPT training.



Figure 1. ChatGPT-3.5 trained on a diverse set of text data.

According to ChatGPT-3.5 (Actually asking with a Prompt "Please tell me the types and proportions of languages used in the language model of ChatGPT-3.5."), ChatGPT-3.5 have been trained on a diverse set of text data, with a primary focus on the English language. As of September 2021, the model includes the following languages and their respective proportions: English:80%, Catalan, Chinese and Dutch:3%, French, German and Italian:2%, Japanese, Korean, Portuguese, Russian, Spanish and Turkish:1%. Typically, performance is lower for non-English languages, especially those with fewer resources. Typically, large language performance might lower for non-English languages, especially those with fewer resources (Ahuja et al., 2023).

## 3. Methods

### 3.1 Type of ChatGPT

I utilized, in this study, the OpenAI's ChatGPT-3.5 model, which is available for free. All were generated by GPT-3.5 using chat.openai.com. The experiment was conducted using ChatGPT-3.5 for the period July 18-27, 2023.

### 3.2 Dataset

I used PISA sample test book "Take the Test: Sample Questions from OECD's PISA Assessments" (Salz & Figueroa, 2009), which is freely available in online. The book contains questions in three different categories: Reading sample tasks, Mathematics sample tasks and Science sample tasks. The Reading sample contains 17 units of questions, the Mathematics sample contains 50 units of questions, and the Science sample contains 34 units of questions, each with a set of answers and scoring criteria. The Mathematics sample and the Science sample contain 37 and 22 units, respectively, of questions that involve reading and understanding diagrams, with more than half of the questions containing diagrams. In contrast, the reading sample test contains only 8 units, and this study focused on questions in the reading sample test that did not include questions to read and understand figures and tables. I adopted Reading sample tasks, which does not require reading diagrams. This is because ChatGPT3.5, which is available for free in schools, is not capable of interpreting diagrams and tables.



For further information please contact Fiona on ext. 5577.

#### Figure 2. Example of PISA Reading Sample Test excerpted from (OECD, 2009)

As shown Figure 2, the Reading unit of the PISA sample test begins with the text for reading, followed by about 5 questions. Some of those questions are multiple choice questions, some are short answer questions, etc. In this study, only the multiple choice questions were addressed. I did not use anything other than choice questions like QUESTION2.2 that require descriptive answers.

#### 3.3 Experiment Design

Figure 3 shows an overview of the experimental design. First, the ChatGPT prompts (Fig. 3 top) included the text at the beginning of the Reading test Unit for any of the multiple-choice questions. The question text is prefixed with a #, and each choice is also prefixed with a #, e.g., #Qustion..., #A..., #D... I manually checked and scored the answers generated from the prompts that I put into ChatGPT to see if they were correct (Fig. 3 left). The same thing was done for Japanese questions as for the English questions, that is, English questions were translated into Japanese questions using the google translate function in the Google spreadsheet, then prompted and graded to see if the answers were correct (Fig. 3 right). No modifications were made to the output of Google translation. I compared the percentage of correct answers to questions originally in English to the percentage of correct answers when English is translated into Japanese. I also looked at the reverse case, i.e., the percentage of correct answers for questions that were originally in Japanese and translated into English.



Figure 3. Overview of Experiment Design.







The Figure 4 shows that accuracy rate of the multiple questions. I input twenty multiple questions (original English and translated Japanese for each) to ChatGPT and receive the answers. I checked if the answer was correct and graphed the percentage of correct answers. As you can see, in English, ChatGPT answered all the questions correctly without making a single mistake, whereas when translated into Japanese, they got 6 out of 20 questions wrong, for a 70% correct rate. This indicates that when the ChatGPT is used in Japanese, its performance is slightly lower on the multiple-choice questions.

Next, to determine whether differences in translation efficiency affected ChatGPT performance, I examined the percentage of correct answers for originally Japanese and translated English questions using four of the sample Japanese PISA questions (left side of Fig. 5) which includes questions and the correct answers for those. These quizzes are publicly available and freely accessible from National Institute for Educational Policy Research in Japan web site (<u>https://www.nier.go.jp/kokusai/pisa/pdf/2018/04\_example.pdf</u>). This material includes one Reading sample task and seven questions with four choice

questions within it. The comparison results between having ChatGPT solve these problems in Japanese and having them translated into English by Google Translate are shown on the right side of Figure 5. This figure shows that the correct response rate for the ChatGPT in Japanese is 50%, whereas when translated into English, it indicates 100% correct responses suggesting that the performance of ChatGPT for Japanese is inferior, not the performance of English to Japanese translation by Google Translate on PISA multiple choice test.

Sample questions originally in Japanese





### 5. Limitations

Not all PISA questions are open and available, and many of them are private for use in the next PISA test. For example, the questions used in the PISA 2006 survey are all private for use in the next survey in 2009. Therefore, the questions for this study were limited. It is possible that the superior ChatGPT performance in English was due to available PISA questions on the web site being used as training data for the development of the ChatGPT large language model, but we do not know what specific data was being used to train the ChatGPT because it is not publicly available. The version of ChatGPT is updated periodically, and the accuracy may change in future similar experiments. It has been reported that the accuracy differs from period to period when ChatGPT is used (Chen et al., 2023). In addition, the parameters of the temperature of the ChatGPT were not set in this study.

### 6. Implications

When using generative AI in K12, one should be careful about its performance when using local languages such as Japanese. Conversely, it may perform well when used in English classes as a foreign language. Recent studies have shown that LLMs that can handle multiple languages are more accurate when asked in English (Etxaniz et al., 2023).

For researcher and developer, when studying Generative AI in education, it is necessary to try it with various LLMs with different language resources. Also, development of language models specific to a kind of language other than English, for example, Japanese, Chinese or Korean large language models requires research and development such as Open-Calm-7B (CyberAgent, 2023), Baichuan-13B-Base (Baichuan Intelligent Technology, 2023) or HyperCLOVA (Kim et al., 2021).

### 7. Conclusion

I examined how well the problem-solving skills of the ChatGPT perform in the context of K12 education using multiple sample test from the global PISA comparing English and Japanese expressions. I found that ChatGPT performed 100% correctly answered in English expression whereas 70% correctly answered in translated Japanese. I also conducted comparison originally Japanese guizzes and those translated English guizzes and found that the correct response rate for the ChatGPT in Japanese is 50%, while translated into English, it indicates 100% correct responses. These results suggest the performance of ChatGPT is

inferior performance in languages other than English, such as Japanese, and caution should be exercised when using ChatGPT in K12 educational settings. Also, there is a need to develop and apply large language models in languages other than English.

### Acknowledgements

I thank the OECD for permitting us to make sample questions available on the web and to use excerpts from them. These excerpts are copyrighted by the OECD. This research was supported by JSPS JP23K17012 and 23H01001.

## References

- Ahuja, K., Diddee, H., Hada, R., Ochieng, M., Ramesh, K., Jain, P., Nambi, A., Ganu, T., Segal, S., Axmed, M., Bali, K., & Sitaram, S. (2023). MEGA: Multilingual Evaluation of Generative AI. In arXiv [cs.CL]. arXiv. http://arxiv.org/abs/2303.12528
- Baichuan Intelligent Technology. (2023). Baichuan-13B-Base [Software]. Hugging Face. https://huggingface.co/baichuan-inc/Baichuan-13B-Base
- Banks, D. C. (2023, May 18). ChatGPT caught NYC schools off guard. Now, we're determined to embrace its potential. Chalkbeat NEW YORK.
  - https://ny.chalkbeat.org/2023/5/18/23727942/chatgpt-nyc-schools-david-banks
- Birkett, L., Fowler, T., & Pullen, S. (2023). Performance of ChatGPT on a primary FRCA multiple choice question bank. *British Journal of Anaesthesia*, *131*(2), e34–e35.
- Chen, L., Zaharia, M., & Zou, J. (2023). How is ChatGPT's behavior changing over time? In *arXiv* [cs.CL]. arXiv. http://arxiv.org/abs/2307.09009
- CyberAgent. (2023). *Open-Calm-7B* [Software]. Hugging Face. https://huggingface.co/cyberagent/open-calm-7b
- DfE. (2023, March 29). *Generative artificial intelligence in education*. Policy Paper. https://www.gov.uk/government/publications/generative-artificial-intelligence-in-education
- Etxaniz, J., Azkune, G., Soroa, A., de Lacalle, O. L., & Artetxe, M. (2023). Do Multilingual Language Models Think Better in English? In *arXiv [cs.CL]*. arXiv. http://arxiv.org/abs/2308.01223
- Frieder, S., Pinchetti, L., Chevalier, A., Griffiths, R.-R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., & Berner, J. (2023). Mathematical Capabilities of ChatGPT. In *arXiv [cs.LG]*. arXiv. http://arxiv.org/abs/2301.13867
- Gilson, A., Safranek, C. W., Huang, T., & Socrates, V. (2023). How does ChatGPT perform on the United States medical licensing examination? The implications of large language models for medical education and .... *JMIR Medical*. https://mededu.jmir.org/2023/1/e45312/
- Kim, B., Kim, H., Lee, S.-W., Lee, G., Kwak, D., Jeon, D. H., Park, S., Kim, S., Kim, S., Seo, D., Lee, H., Jeong, M., Lee, S., Kim, M., Ko, S. H., Kim, S., Park, T., Kim, J., Kang, S., ... Sung, N. (2021). What Changes Can Large-scale Language Models Bring? Intensive Study on HyperCLOVA: Billions-scale Korean Generative Pretrained Transformers. In *arXiv [cs.CL]*. arXiv. http://arxiv.org/abs/2109.04650
- KYODO. (2023, June 22). Japanese schools to be allowed limited use of generative AI. KYODO NEWS.
- Lu, H., Huang, H., Ma, S., Zhang, D., Lam, W., & Wei, F. (2022). Advancing multilingual pre-training: TRIP triangular document-level pre-training for multilingual language models. In *arXiv [cs.CL]*. arXiv. http://arxiv.org/abs/2212.07752
- Newton, P. M., & Xiromeriti, M. (2023). ChatGPT Performance on MCQ Exams in Higher Education. A Pragmatic Scoping Review. edarxiv.org. https://edarxiv.org/sytu3
- OpenAI. (2022). Introducing ChatGPT. https://openai.com/blog/chatgpt/
- Passby, L., Jenko, N., & Wernham, A. (2023). Performance of ChatGPT on dermatology Specialty Certificate Examination multiple choice questions. *Clinical and Experimental Dermatology*. https://doi.org/10.1093/ced/llad197
- Salz, S., & Figueroa, D. T. (2009). *Take the Test: Sample Questions from OECD's PISA* Assessments. ERIC.
- Shen, Y., Heacock, L., Elias, J., Hentel, K. D., Reig, B., Shih, G., & Moy, L. (2023). ChatGPT and Other Large Language Models Are Double-edged Swords [Review of *ChatGPT and Other Large Language Models Are Double-edged Swords*]. *Radiology*, 307(2), e230163.

# A Systematic Review of Generative Artificial Intelligence in Language Education

Zilin WANG<sup>a</sup>, Di ZOU<sup>b\*</sup>, Lap-Kei Keith LEE<sup>a</sup>, Haoran Xie<sup>c</sup> & Fu Lee Wang<sup>a</sup> <sup>a</sup>School of Science and Technology, Hong Kong Metropolitan University, Hong Kong SAR <sup>b</sup>School of Graduate Studies, Lingnan University, Hong Kong SAR <sup>c</sup>Department of Computing and Decision Sciences, Lingnan University, Hong Kong SAR \*dizoudaisy@gmail.com

Abstract: This research paper presents a comprehensive exploration of the application of Generative Artificial Intelligence (GAI) within the realm of language education. Through a systematic review of 20 empirical studies conducted within a structured three-step framework, this study elucidates the multifaceted integration of GAI in language learning. The examination encompasses various dimensions, including target languages, learners' educational levels, GAI applications, language skills, and practical outcomes. Additionally, the study critically assesses the key advantages and challenges inherent in GAI's role in language education. The paper provides valuable insights into human-technology interaction by delving into language learners' attitudes toward GAI. Notably, this research identifies three pivotal roles GAI assumes within the language learning process, and they are co-author, evaluator, and learning materials provider. The study concludes by charting a path toward future research endeavors within the evolving landscape of GAI-based language learning by implicating the future research direction of integrating the GAI into language education, including collaborations between humans and GAI, clarifying the definition of GAI-powered plagiarism, GAI-based language activities design, prompting strategies, and digital literacy.

**Keywords:** Generative artificial intelligence, language education, technologyenhanced language learning, ChatGPT

## 1. Introduction

The integration of Artificial Intelligence (AI) in language teaching and learning has become increasingly essential in today's educational landscape. On the one hand, the shift towards online education, particularly in the wake of the COVID-19 pandemic, necessitates the use of AI to provide effective, accessible, and flexible language learning tools that can be utilized anytime, anywhere (Jia et al., 2022). On the other hand, AI technology has found extensive applications in modern language education. These applications encompass automated writing assessments, intelligent tutoring systems for reading, computer-mediated communication, and personalized language learning (Huang et al., 2023; Liang et al., 2021). The latest advancements in GAI technology are poised to continually influence the landscape of language learning in today's digital age.

GAI stands as a groundbreaking stride in the realm of artificial intelligence, characterized by its autonomous capability to craft coherent and contextually pertinent human-like text (Cooper, 2023). The transformative potential of GAI is progressively gaining recognition within the educational sphere, heralding the dawn of a novel epoch replete with dynamic and interactive learning encounters (Rahman & Watanobe, 2023; Wilson & Billam, 2023). In the context of language education, the integration of cutting-edge technological advancements holds a significant sway, and technology-enhanced language learning stands as a burgeoning focal point within this domain. This study embarks on a systematic review to

delve into the realm of GAI deployment in language education, seeking to unearth the present landscape of its integration and chart a course towards future research trajectories.

## 2. Literature Review

In technology-enhanced language learning (TELL), systematic reviews play a crucial role in advancing the field and informing educational practices. The systematic review is a rigorous and comprehensive approach to synthesizing and analyzing existing research on a specific topic by identifying the research trends and gaps (Zou et al., 2018), organizing diverse findings (Ghanizadeh et al., 2015; Shadiev & Yang, 2020) and providing insights for future research (Su & Zou, 2022). The current systematic review contributes to figuring out the applications of GAI in language learning based on existing empirical studies to investigate further where we start and where we could go.

Few review articles have shared viewpoints on employing ChatGPT in education contexts. Wilson and Billam (2023) reviewed ten interesting articles published between January and May 2023 about using ChatGPT in educational practice, which included radio broadcasts, research articles, and government frameworks. These published articles discussed some cases of students using ChatGPT for academic assignments and academic integrity-related issues. Ethical issues on ChatGPT used in academic language and learning were also a concern (Wilson & Billam, 2023). Similarly, Rahman and Watanabe (2023) expressed the possibility of cheating in exams by using ChatGPT, which diminished students' critical thinking capacity. Meanwhile, human teachers hardly differentiate the ChatGPT and students' work. These threats brought by ChatGPT could transform traditional teaching and learning methods. However, ChatGPT offered opportunities for learners and educators by providing personalized feedback and interactive conversations (Rahman & Watanobe, 2023). Hence, the challenges and opportunities brought by ChatGPT, and more broadly, GAI, are worth further discussion.

In specific subject education, including English language learning (Bin-Hady et al., 2023), academia and libraries (Houston & Corrado, 2023), and medical education (Armitage, 2023), ChatGPT had been first tried and discussed in the review articles. Regarding language education, ChatGPT could serve as a scaffolding to help language learners design learning plans, and ChatGPT's immediate feedback acts as a partner in language practice (Bin-Hady et al., 2023). Regarding academia and libraries, ChatGPT influenced reference practices, collection development, and metadata creation and transformation (Houston & Corrado, 2023). Regarding medical education, the impacts of ChatGPT tend to be harmful as it threatens students' ability to do clinical practice (Armitage, 2023).

The existed literature review about GAI needs to be more systematic than viewpoint reports. At the same time, ChatGPT was the only GAI tool focused on by the previous reviewers, while other diverse GAI tools, such as Bard, POE, and DALL-E2, have been popularized now. Therefore, a systematic review of GAI tools used in the educational field is needed. The current review narrows the scope in the language education fields, which already have abundant empirical studies and lacks a systematic review of current research status, integration of GAI into language education, and future research trends. The research questions for the current review are:

- What is the status of research on GAI-based language learning related to target languages, learners' educational levels, GAI applications, language skills, and practical effects of GAI?
- What are the main benefits and challenges of using GAI in language learning?
- What are the learners' attitudes toward using GAI in language learning?
- What are the roles GAI plays in the language learning process?
- What are the future directions for investigating GAI integration in language learning?

### 3. Methods

To exam the above five research questions, we conducted a systematic review based on a three-step methodology which is commonly used in technology-enhanced language learning reviews (Ghanizadeh et al., 2015; Shadiev & Yang, 2020; Su & Zou, 2022). The three steps involve article searching, screening, and coding. The first step was identifying topic-related keywords and searching them in the database. Then, the primary search results were screened by specific exclusion and inclusion criteria. Finally, the authors had an intensive reading of the last included articles, had discussions, and labeled the contents together based on the article coding scheme. The details are illustrated in this section.

## 3.1 Article Search

The databases used in the current review were Web of Science (WoS) and Scopus, which were commonly employed in systematic reviews due to their comprehensive coverage and inclusion of high-quality journals (Celaya et al., 2020). Within these two databases, we enter the keywords related to GAI ("generative artificial intelligence" OR "GAI" OR "ChatGPT" OR "POE" OR "Bing Chat" OR "GPT-4" OR "GPT" OR "Bard" OR "DALL-E2"), language acquisition (language OR second language OR foreign language OR EFL OR ESL OR TESOL OR TEFL) and education (educat\* OR acquisit\* OR learn\* OR teach\* OR student\* OR class\*). We further refined the document type as "Article" and the language as "English". The search strings in two databases and initial results retrieved in August 2023 were reported in Table 1.

Database	Search Strings	Search Results
WoS	"Generative artificial intelligence" OR	231
	"GAI" OR "ChatGPT" OR "POE"	
	OR "Bing Chat" OR "GPT-4" OR	
	"GPT" OR "Bard" OR "DALL-E2"	
	(Topic) and language OR second	
	language OR foreign language	
	OR EFL OR ESL OR TESOL OR	
	TEFL (Topic) and English	
	(Language) and educat* OR	
	acquisit* OR learn* OR teach* OR	
	student* OR class* (Topic)	
Scopus	(TITLE-ABS-KEY ("generative	197
	artificial intelligence" OR "GAI"	
	"Bard" OR "DALL-E2" ) AND	
	IIILE-ABS-KEY (language OR	
	second AND language OR	
	OR and OR togol OR tofl) AND	
	TITLE APS KEY ( aducat* OP	
	acquisit* OP loarn* OP tooch*	
	OR student* OR class*)) AND	
	$(1 \text{ MIT}_TO (DOCTVPE "ar"))$	
	$\Delta ND (IIMIT_TO(I) \Delta NGUAGE$	
	"English"))	
Scopus	OR "Bing Chat" OR "GPT-4" OR "GPT" OR "Bard" OR "DALL-E2" (Topic) and language OR second language OR foreign language OR EFL OR ESL OR TESOL OR TEFL (Topic) and English (Language) and educat* OR acquisit* OR learn* OR teach* OR student* OR class* (Topic) (TITLE-ABS-KEY ( "generative artificial intelligence" OR "GAI" OR "ChatGPT" OR "POE" OR "Bing Chat" OR "GPT-4" OR "Bard" OR "DALL-E2") AND TITLE-ABS-KEY ( language OR second AND language OR foreign AND language OR efl OR esl OR tesol OR tefl) AND TITLE-ABS-KEY ( educat* OR acquisit* OR learn* OR teach* OR student* OR class* )) AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (LANGUAGE, "English"))	197

Table 1. Primary Search in WoS and Scopus

## 3.2 Article Selection

Based on the initial results, we still needed to select well-designed empirical studies in the next step systematically. To ensure a focused selection, we established exclusion and

inclusion criteria, which were reported in Table 2. At first, we removed 98 duplications from two databases. Subsequently, 49 articles of literature review, viewpoint, or report were excluded. Then, we carefully examined the remaining items' titles, sources, and abstracts. We further identified and excluded 217 articles that did not investigate language learning, 43 that did not exploit GAI, and 1 that could not find full text. Through the selection process, shown in Figure 1, 20 articles that met the inclusion criteria were kept for further analysis.

Table 1. Exclusion and Inclusion Criteria

Exclusion Criteria	Inclusion Criteria
Review, viewpoint, or report	Empirical study
Does not investigate language learning	Investigate language learning
Irrelevant to generative artificial intelligence	Implement generative artificial intelligence
	tools
Full text unavailable	Full text available



Figure 1. The article selection process.

# 3.3 Article Coding Scheme

The data extraction and coding procedure was based on previous systematic reviews in computer and education (Crompton & Burke, 2018; Yadegaridehkordi et al., 2019). The authors reviewed 20 articles one by one, assisted by reference management software Mendeley and recorded the necessary data in MS Excel spreadsheets. Regarding five research questions, the 20 articles were coded by the following tags: target language, educational level, GAI Applications, language skills, advantages, challenges, practical effects, learners' attitudes, and the role of GAI in the learning process. Table 3 carefully presents the coding scheme.

Table 3. Article Coding Scheme

Tags	Coding Examples	Example References
Target languages	English	(Perkins, 2023)
	Greek	(Ross, 2023)
Educational levels	Higher education	(Macdonald et al., 2023)

GAI applications	Elementary education ChatGPT-2	(Lee et al., 2023) (Bulut & Yildirim-Erbasli, 2022)
Language skills	ChatGPT-3.5 Reading	(Ahmed, 2023) (Bulut & Yildirim-Erbasli, 2022)
Advantages	Writing Students were motivated in GAI-based reading activity	(Y. Su et al., 2023) (Lee et al., 2023)
	"ChatGPT could become a great help to researchers worldwide in designing their studies, conducting analyses and drafting their research articles"	(Macdonald et al., 2023)
Challenges	The frequent use of ChatGPT leads to overly similar paragraphs and structure of many papers in the same field, leading to problems with plagiarism check	(Macdonald et al., 2023)
	Lack of human connection and personalization	(Mohamed, 2023)
Practical effects	Major effect Minor effect	(Macdonald et al., 2023) (Ahmed, 2023)
Learners' attitudes	Positive Negative	(Macdonald et al., 2023) (Qasem, 2023)
The role of GAI	Work as a co-author Learning content provider	(Macdonald et al., 2023) (Lee et al., 2023)

# 4. Results

# 4.1 Target Languages of GAI-based Language Learning

Out of the total twenty reviewed articles, English emerged as the predominant target learning language, constituting 18 of the articles, whereas Greek and Arabic learning each accounted for a single research article. This distribution can be attributed to the extensive training of the large language model (LLM) on the English language corpus. Notably, one of the prominently used versions of ChatGPT, namely ChatGPT-3, was frequently employed in the research, exclusively generating outputs in English (Ross, 2023). Simultaneously, this linguistic bias resulted in varying output quality when students employed GAI to learn different target languages. Among the findings of this review, it became evident that the quality of English output surpassed that of Greek and Arabic outputs. Rose (2023) observed challenges with ChatGPT-3.5 in handling Ancient Greek, while its performance with classical Latin and Sanskrit was commendable. Conversely, Beheitt and HajHmida (2023) expressed dissatisfaction with the Arabic output.

# 4.2 Educational Level of GAI-Based Language Learning

GAI found its predominant application in higher education, as indicated by 16 articles that delved into its utilization within this context. Additionally, two articles integrated GAI into elementary education, while another piece of research explored its impact on secondary

education. Furthermore, there was an article that does not specify the educational level under consideration.

# 4.3 GAI Applications for GAI-Based Language Learning

Figure 2 presents an overview of the GAI applications utilized in the reviewed articles. The most commonly employed platform was ChatGPT, which featured in 16 empirical studies. Within this set of 16 articles, various iterations of ChatGPT, namely ChatGPT-2, ChatGPT-3, ChatGPT-3.5, and ChatGPT-4, were scrutinized for their effectiveness in the language learning process. Notably, ChatGPT-3.5 garnered the highest frequency of adoption, featuring in 10 studies. Versions 3 and 2 were each applied in 4 and 3 articles, respectively, while the most recent iteration, ChatGPT 4, found application in two articles. However, one article did not specify the particular version of ChatGPT employed.

Apart from ChatGPT, the review also examined other GAI applications, including Bard, Bing Chat, CopyAI, and Google T5, with each being the subject of investigation in a separate article.



Figure 2. GAI applications used in the reviewed articles.

# 4.4 Language Skills of GAI-Based Language Learning

In the reviewed articles, GAI proves instrumental in fostering learners' acquisition of writing, reading, communication, and vocabulary skills. Twelve studies have harnessed GAI to facilitate writing practice, encompassing diverse tasks such as digital writing (Perkins, 2023), academic writing (Fyfe, 2022; Macdonald et al., 2023; Yan, 2023), reflective writing (Li et al., 2023), poetry composition (Beheitt & HajHmida, 2023), and argumentative discourse (Hinton & Wagemans, 2023; Y. Su et al., 2023). Additionally, two studies utilized ChatGPT to generate reading materials (Bulut & Yildirim-Erbasli, 2022; Lee et al., 2023), while another pair employed the immediate feedback features of ChatGPT to cultivate learners' communication aptitude ((Mohamed, 2023; Young & Shishido, 2023)). One study recognized ChatGPT as an effective tool for language learners to enhance vocabulary acquisition (Ross, 2023). However, three studies did not specify the particular learning skills that were intended to be enhanced through GAI.

# 4.5 Advantages of Using GAI in Language Learning

The integration of Generative Artificial Intelligence (GAI) into the language learning process confers significant benefits, leveraging its multifunctional capabilities, provision of immediate and personalized feedback, and user-friendly interface.

Within the scope of the surveyed literature, GAI has showcased remarkable versatility, evidenced by its proficiency in executing diverse tasks. These tasks encompass a spectrum of activities, such as generating essays in numerous languages, crafting speeches

marked by a personalized style (Rudolph et al., 2023), succinctly summarizing complex documents (Yan, 2023), and, intriguingly, producing written content that surpasses the novelty observed in human-authored works. This achievement is attributed to GAI's extensive linguistic foundation (McCoy et al., 2023; Mohamed, 2023) An illustrative manifestation of GAI's efficacy is discerned in the realm of argumentative writing, where the arguments generated by GAI exhibit originality and potency, thereby augmenting the preparatory stages of language learners' writing endeavors (Kohnke et al., 2023; Y. Su et al., 2023). Additionally, GAI emerges as a robust tool for vocabulary and grammar verification, notably aiding in the proofreading phase (Fyfe, 2022; Su et al., 2023). Noteworthy is the discernible enhancement of students' written output through GAI adoption, evidenced by a reduction in grammatical errors and a proliferation of lexical diversity (Yan, 2023).

Concurrently, GAI substantiates its user-friendly disposition, catering to the requisites of both language educators and learners. This adaptability extends to the tailoring of outputs to align with the linguistic proficiency levels of the target audience. The intricate task of curating apt language learning materials, especially pertinent for elementary-level students, encounters a pragmatic solution in GAI's ability to swiftly generate contextually suitable texts. Educators can seamlessly elicit content suited to the cognitive capacity of young learners, enhancing the pedagogical experience (Ahmed, 2023; Bulut & Yildirim-Erbasli, 2022; Lee et al., 2023).

Moreover, empirical investigations underscore GAI's potential to engender heightened learner engagement, a facet facilitated by its interactive attributes. Immediate feedback emerges as a pivotal catalyst, invigorating students' proclivity to engage in meaningful language application, consequently fortifying their communicative self-assurance (Lee et al., 2023; Mohamed, 2023).

## 4.6 Challenges of Using GAI in Language Learning

The major defects of using GAI in language learning lie in the unstable output quality, ethical and plagiarism issues, the lack of human connection, and learners' over-dependent on it.

The poor quality of GAI output appeared in minority language conversations (Beheitt & HajHmida, 2023; Ross, 2023). On the other hand, the fake information produced by GAI had side effects on scientific research and academic writing (Bulut & Yildirim-Erbasli, 2022; Macdonald et al., 2023). Besides, for different task modalities, such as graphs and figures, GAI performed struggled with them (Li et al., 2023).

In addition, the easy accessibility of GAI to students lead to massive similar work (Macdonald et al., 2023), and human teacher find it hard to differentiate between students' work and GAI's work (Li et al., 2023). It challenged traditional academic policies about ethical and plagiarism issues (Macdonald et al., 2023).

Moreover, the fast generation lessened learners' sense of participation, and compared to the GAI partners, students preferred to interact with teachers and classmates (Ahmed, 2023). Further, learners became lazy and did not check the GAI work because they were over-dependent on GAI. This situation had a high opportunity to appear in the low language proficiency level learners (Ahmed, 2023).

## 4.7 Practical Effects of Using GAI in Language Learning

In 20 reviewed articles, 15 suggested the practical effects of GAI. 12 studies indicated the GAI had excellent performance in completing the language learning tasks or significantly improved learners' language learning performance. Other three studies indicate GAI had minor effects on promoting students' performance.

## 4.8 Learners' Attitudes of Using GAI in Language Learning

Nineteen articles investigated language learners' attitudes toward GAI's application. The participants from 13 studies held positive attitudes and highly rated GAI such as "good

learning tools" (Ross, 2023), "in favor of using GAI" (Lee et al., 2023), and "GPT architecture fits best for automatic poem generation" (Beheitt & HajHmida, 2023). Four studies kept neutral attitudes who admitted the advantages of GAI but were still concerned about its threats of plagiarism and ethical issues. Two studies expressed negative attitudes that consider the GAI cannot replace the classroom teaching modes (Ahmed, 2023; Qasem, 2023). On the one hand, students hardly check ChatGPT's output because they did not feel any sense of connect with the output they had, and instead of ChatGPT, the students preferred to cooperate with their classmates (Ahmed, 2023); On the other hand, the fake information produced by GAI made it untrustworthy in the academic writing process (Qasem, 2023).

# 4.9 The Roles of GAI in GAI-Based Language Learning

We identified GAI's three roles in the language learning process: learning materials provider, co-author, and evaluator. As the learning materials provider, GAI could be an effective tool to gather learning information in vocabulary learning and practice with the definitions, translations, and grammar concepts (Ross, 2023). Meanwhile, GAI could generate reading materials to match different language proficiency levels of learners (Bulut & Yildirim-Erbasli, 2022; Lee et al., 2023). As a co-author, GAI could provide a writing preparation plan and suggestions in completing writing tasks (Kohnke et al., 2023; McCoy et al., 2023; Su et al., 2023), and it also worked as a grammar checker in the proofreading phase (Fyfe, 2022). Being an evaluator, GAI could provide feedback based on the input comment framework. However, the feedback quality depends crucially on the evaluative rubrics (Su et al., 2023).

## 5. Discussions

Compared with the related review, the present review had consistency in that the plagiarism and ethical issues were urgent to solve (Wilson & Billam, 2023). Meanwhile, within the educational domain, GAI's multifunctions could reduce learners' active learning (Rahman & Watanobe, 2023). Within the language education field, our results also agreed with Bin-Hady et al. (2023) that GAI could be an effective tool during the phase of language learning preparation. However, we also found that GAI could provide strong support for language learning through immediate feedback and a vast knowledge base during the learning process. GAI could be a practical tool for grammar checks and language rectification during the proofreading phase. In addition, the interactive feature of GAI unsatisfied the learners compared with traditional language classrooms, which is different from Rahman and Watanobe (2023), who highly commented on this feature. Furthermore, the current study identified the roles GAI played in the language learning process as an evaluator, co-author, and learning content provider, which extends the recognition to the GAI application in language education.

We also propose future research directions in integrating GAI into language learning from five aspects (see Figure 3):

- The methods of applying GAI in language education, as reviewed in previous studies, are limited and lack systematic approaches. Based on the three roles of GAI in language learning proposed by this study, the form of collaboration between humans and GAI warrants further exploration.
- Redefine the GAI-based plagiarism and adjust academic integrity policy. This study proved that GAI has been a strong writing tool, and this factor challenged traditional academic integrity policy. The discussion of GAI-based plagiarism is necessary.
- Examine the design of GAI-based language activities for learners with varying language proficiency levels. The current study discovered that the effectiveness of GAI-based learning methods is related to the language proficiency level of the learner. Consequently, the design of GAI-based language activities should take this into account and personalize the activities accordingly.

- Identify effective prompting strategies and attempt to incorporate student feedback into these prompts. The language proficiency level of students influences the quality of the prompts and the output quality of GAI. To enhance students' questioning skills and the output quality in GAI applications, further exploration of prompting strategies is required.
- Boost the digital literacy of both language educators and learners by integrating GAI into language teacher education programs. The studies reviewed aim to incorporate GAI into educational practice, even teacher education programs. This requires both language learners and teachers to enhance their digital literacy. The methods to promote digital literacy among participants of different educational levels need further discussion.



Figure 3. Future research directions of GAI in language education.

# 6. Conclusions and Limitations

With the aim of identifying how GAI could apply to the language education field, the current systematic review analyzed 20 related empirical studies strictly under the three-step framework. We reported the current research status of GAI in language learning related to target languages, learners' educational levels, GAI applications, language skills, and practical effects. Moreover, we combed the main benefits and challenges of using GAI in language learning. In addition, the language learning process were identified as co-author, evaluator, and learning materials provider. We implicated future research directions in the GAI-based language learning field, referring to integrating the GAI into language education, including collaborations between humans and GAI, clarifying the definition of GAI-powered plagiarism, GAI-based language activities design, prompting strategies, and digital literacy.

The limitation lies in the limited empirical studies that could be analyzed because GAI was the latest concept, which is still developed in the initial phase.

# References

- Ahmed, M. A. (2023). ChatGPT and the EFL Classroom: Supplement or Substitute in Saudi Arabia's Eastern Region. https://doi.org/http://dx.doi.org/10.18576/isl/120704
- Armitage, R. C. (2023). ChatGPT: the threats to medical education. *Postgraduate Medical Journal*, qgad046.
- Beheitt, M. E. G., & HajHmida, M. Ben. (2023). EFFECTIVENESS OF ZERO-SHOT MODELS IN AUTOMATIC ARABIC POEM GENERATION. *Jordanian Journal of Computers and Information Technology*, 9(1).

Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M. (2023). Exploring the dimensions of ChatGPT in English language learning: A global perspective. *Library Hi Tech*.

- Bulut, O., & Yildirim-Erbasli, S. N. (2022). Automatic story and item generation for reading comprehension assessments with transformers. *International Journal of Assessment Tools in Education*, 9(Special Issue), 72–87. https://doi.org/https://doi.org/10.21449/ijate.1124382
- Celaya, I., Ramírez-Montoya, M. S., Naval, C., & Arbués, E. (2020). Uses of the podcast for educational purposes. Systematic mapping of the literature in WoS and Scopus (2014-2019). *Revista Latina de Comunicacion Social*, 77, 179–201.
- Cooper, G. (2023). Examining science education in chatgpt: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, *32*(3), 444–452.
- Fyfe, P. (2022). How to cheat on your final paper: Assigning AI for student writing. *AI* & SOCIETY, 1– 11.
- Ghanizadeh, A., Razavi, A., & Jahedizadeh, S. (2015). Technology-enhanced language learning (TELL): A review of resources and upshots. *International Letters of Chemistry, Physics and Astronomy*, *54*, 73–87.
- Hinton, M., & Wagemans, J. H. M. (2023). How persuasive is AI-generated argumentation? An analysis of the quality of an argumentative text produced by the GPT-3 AI text generator. *Argument & Computation, Preprint*, 1–16. https://doi.org/10.3233/AAC-210026
- Houston, A. B., & Corrado, E. M. (2023). Embracing ChatGPT: Implications of Emergent Language Models for Academia and Libraries. *Technical Services Quarterly*, *40*(2), 76–91.
- Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, *26*(1), 112–131.
- Jia, F., Sun, D., Ma, Q., & Looi, C.-K. (2022). Developing an AI-Based learning system for L2 learners' authentic and ubiguitous learning in English language. *Sustainability*, *14*(23), 15527.
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence*, 100156. https://doi.org/https://doi.org/10.1016/j.caeai.2023.100156
- Lee, J. H., Shin, D., & Noh, W. (2023). Artificial Intelligence-Based Content Generator Technology for Young English-as-a-Foreign-Language Learners' Reading Enjoyment. *RELC Journal*, 00336882231165060. https://doi.org/10.1177/00336882231165060
- Li, Y., Sha, L., Yan, L., Lin, J., Raković, M., Galbraith, K., Lyons, K., Gašević, D., & Chen, G. (2023). Can large language models write reflectively. *Computers and Education: Artificial Intelligence*, *4*, 100140. https://doi.org/10.1016/j.caeai.2023.100140
- Liang, J.-C., Hwang, G.-J., Chen, M.-R. A., & Darmawansah, D. (2021). Roles and research foci of artificial intelligence in language education: an integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments*, 1–27.
- Macdonald, C., Adeloye, D., Sheikh, A., & Rudan, I. (2023). Can ChatGPT draft a research article? An example of population-level vaccine effectiveness analysis. *Journal of Global Health*, *13*. https://doi.org/10.7189/jogh.13.01003
- McCoy, R. T., Smolensky, P., Linzen, T., Gao, J., & Celikyilmaz, A. (2023). How much do language models copy from their training data? evaluating linguistic novelty in text generation using raven. *Transactions of the Association for Computational Linguistics*, *11*, 652–670. https://doi.org/https://doi.org/10.1162/tacl a 00567
- Mohamed, A. M. (2023). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: perceptions of EFL Faculty Members. *Education and Information Technologies*, 1–23.
- Perkins, M. (2023). Academic integrity considerations of AI Large Language Models in the postpandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, 20(2), 1–19. https://doi.org/10.53761/1.20.02.07
- Qasem, F. (2023). ChatGPT in scientific and academic research: future fears and reassurances. *Library Hi Tech News*, *40*(3), 30–32.
- Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, *13*(9), 5783.
- Ross, E. A. S. (2023). A New Frontier: AI and Ancient Language Pedagogy. *Journal of Classics Teaching*, 1–19. https://doi.org/https://doi.org/10.1017/S2058631023000430
- Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning & Teaching*, 6(1). https://doi.org/10.37074/jalt.2023.6.1.23
- Shadiev, R., & Yang, M. (2020). Review of studies on technology-enhanced language learning and teaching. *Sustainability*, *12*(2), 524.
- Su, F., & Zou, D. (2022). Technology-enhanced collaborative language learning: theoretical foundations, technologies, and implications. *Computer Assisted Language Learning*, *35*(8), 1754–1788.

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, 100752. https://doi.org/https://doi.org/10.1016/j.asw.2023.100752

Wilson, O., & Billam, J. (2023). ChatGPT in June 2023: An annotated bibliography for ALL professionals. *Journal of Academic Language and Learning*, *17*(1), T9–T15.

- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 1–25.
- Young, J. C., & Shishido, M. (2023). Investigating OpenAI's ChatGPT Potentials in Generating Chatbot's Dialogue for English as a Foreign Language Learning. *International Journal of Advanced Computer Science and Applications*, *14*(6). https://doi.org/10.14569/IJACSA.2023.0140607
- Zou, D., Xie, H., & Wang, F. L. (2018). Future trends and research issues of technology-enhanced language learning: A technological perspective. *Knowledge Management & E-Learning*, *10*(4), 426–440.

# Integrating ChatGPT into Synectics Model to Improve High School Student's Creative Writing Skill

Worapong KHUIBUT<sup>a</sup>, Sasivimol PREMTHAISONG<sup>b</sup>, & Pawat CHAIPIDECH<sup>c\*</sup>

<sup>a</sup>Khon Kaen University Demonstration School MoDinDaeng Division, Khon Kaen University, Thailand

<sup>b</sup>Khon Kaen University Demonstration School International Division, Khon Kaen University, Thailand

<sup>c</sup>Faculty of Education, Khon Kaen University, Thailand \*pawach@kku.ac.th

**Abstract:** The cultivation of creative writing abilities has become more challenging, particularly with the intention of fostering creative thinking and providing learner's effective instructional support. The challenge is compounded by the present context, wherein Generative AI has seen significant advancement and plays an essential role in the field of education. ChatGPT has emerged as an important tool for offering writing recommendations. Consequently, the present study integrates the Synectic Model and ChatGPT to enhance ability in creative writing. The participants consisted of 34 high school students in the experimental group and 30 students in the control group. The experimental group reported a significant enhancement in their creative writing abilities, as well as an increased perception of the benefits of technology-supported learning.

Keywords: ChatGPT, generative AI, synectics, creative writing, language learning

### 1. Introduction

The development of reading and writing skills is essential for students and cannot be neglected. In addition, students must exercise and master these skills to the point where they can use them effectively and efficiently to contribute to positive learning outcomes (Chambers, 1997). However, attaining reading and writing fluency is difficult for students. First, fluency necessitates mastering skills, and the Thai language is notoriously difficult to master. Its intricate alphabet, vowel, and tone systems make decoding and production difficult. Particularly difficult is the skill of creative writing, necessitating the use of various learning methods and techniques to perpetually develop students. Integration of numerous learning methods, such as the Synectic teaching method, is one approach to this growth. Using the Synectic teaching method, for instance, facilitates the improvement of creative writing skills via various approaches. A case in point is the investigation of Free Writing Ability by Ina, Sa, and Yayah (2020), in which applying the Synectic learning model to the experimental group increased creativity and comprehension in poetry material development, thereby influencing students' poetry writing abilities.

Similarly, the application of the Synectic Model in developing short story writing, as examined by Nasrulloh and Waraulia (2022), revealed that the Synectic learning model was implemented successfully. Active student participation, positive teacher feedback, and positive student opinions demonstrated the program's efficacy. Therefore, it can be concluded that the learning mentioned above strategy substantially aids in improving learners' creative writing skills. In addition to the Synectic Model's function in developing writing skills, there is a notable incorporation of technology to promote skill development further. In the study by Song et al. (2021), the online collaborative writing platform CollaWrite outfitted with group awareness tools, was used. According to the findings, these tools can increase student engagement and improve the quality of online collaborative learning. This study contributes to
a greater comprehension of the function of group awareness instruments in computermediated collaborative language learning.

Evidently, the incorporation of technology into writing instruction has a possibility, particularly in the current environment where educational technology, including AI for Education, plays a significant role. Hwang et al. (2020) summarized the role of AI in education, stating that AI can serve as an intelligent tutor, instrument, or tutee and aid in educational decision-making. The paper highlights that the convergence of AI and education is a transformation of education and human knowledge, cognition, and culture. ChatGPT is a remarkably convenient tool for utilization due to its compatibility with a wide range of devices, including mobile phones, tablets, and PCs. Moreover, it possesses the ability to engage in discussions resembling those between humans, effectively giving the requested information, particularly in the areas of requesting recommendations and improving language proficiency. As witnessed in the study by Liu et al. (2021), applying AI to provide guidance and feedback on student work is particularly pertinent. This study introduced a reflective thinking promotion mechanism-based artificial intelligence-supported English writing (RTP-AIEW) approach to improve EFL learners' critical thinking and writing quality. The results demonstrated that this strategy significantly enhanced the English writing performance of the experimental group's pupils while also fostering self-efficacy and self-regulated learning and substantially reducing cognitive load. Especially the study by Su, Lin, and Lai (2023), which integrated ChatGPT with providing suggestions for argumentative writing, found that apart from aiding in grammar structuring and proofreading, ChatGPT can also assist in evaluating writing guality and offering writing suggestions.

Considering the abovementioned findings, the researchers intend to implement the Synectic learning approach with ChatGPT within guidance and instruction to fostering learners' creative writing skills in the context of language learning.

#### 2. Literature Review

#### 2.1 Generative AI

Currently, generative AI is highly significant and performs multiple roles, particularly in education. It can be used as an intelligent tutor, tutee, learning tool/partner, or policy-making advisor, according to Hwang et al. (2020), who classified these roles into four main categories: (1) Intelligent tutor, (2) Intelligent tutee, (3) Intelligent learning tool or partner, and (4) Policymaking advisor. These responsibilities are essential for the growth of education in numerous dimensions. (Li et al., 2009) Al for Education is actively constructing large-scale learning systems intending to provide quality learning experiences to millions of students through scalable technologies. For instance, chatbots such as GPT can provide immediate support and feedback to thousands of students' inquiries. To construct deep learning neural networks, GPT models were trained on a massive-tokenized language dataset, enabling them to comprehend text data and generate human-like responses by predicting subsequent words within a given string of words. The model was fine-tuned using Reinforcement Learning from Human Feedback (RLHF) to enhance ChatGPT's ability to respond to various prompts (OpenAI, 2022). This method permits the AI model to modify its behavior and enhance its performance based on a reward model developed from human feedback. ChatGPT can also maintain conversational context, enabling it to use previous interactions to inform subsequent ones and respond to users' requests for alterations or clarifications (OpenAI, 2022).

The iterative training and refining process improves ChatGPT's ability to comprehend natural language, process lengthy inputs, provide creative responses, and maintain a natural conversational flow (Adamopoulou & Moussiades, 2020; Floridi, 2023; Kasneci et al., 2023). As a result, ChatGPT is utilized for educational purposes, particularly in the writing field. For instance, Liu et al. (2021) investigated the incorporation of a mechanism that promotes reflective thinking into artificial intelligence-supported English writing environments. The results demonstrated that this strategy not only substantially improved the English writing performance of the experimental group's pupils but also increased their self-efficacy and self-

regulated learning while reducing cognitive load. In addition, the study addressed the students' learning experiences and perceptions. In addition, Su, Lin, and Lai (2023) conducted a study on integrating ChatGPT with argumentative writing in the classroom. ChatGPT was used within the context of language education. The study found that ChatGPT helps with grammar correction and sentence structure enhancement, thereby facilitating the writing process. Argumentative writing may benefit from the structural, dialogical, and linguistic aspects of interaction with ChatGPT. As a generative artificial intelligence, ChatGPT can provide personalized responses and feedback, evaluate, and recommend a content organization, conduct linguistic analysis, and proofread text. There have been studies examining the limitations of ChatGPT, particularly in the context of creative writing; Shidig (2023) provided insights into the usage of ChatGPT from the standpoint of fostering creative writing abilities and came to the following findings on the influence of the ChatGPT system, an AI-based chatbot, on students' creative writing capabilities. It emphasizes certain drawbacks, such as ChatGPT's inability to supplant human interaction, lack of creativity, failure to capture students' nuances and individual learning styles, and potential adverse effects on individuals' social and psychological aspects. The paper suggests that additional research is necessary to comprehend the potential and limitations of ChatGPT and other artificial intelligence (AI) technologies in education.

#### 2.2 Synectics Model and Creative Writing

Synectics is one of several techniques used to improve brainstorming by taking a more active role and introducing metaphor and structure into the process (Gordon, 1961). In other words, the technique of synectics is defined in a practical and applicable manner. Addressing new applicative structures of teaching techniques enables instructors to become empirically and practically familiar with various teaching methods, thereby making the classroom more dynamic and interactive through modeling and simulation. In addition, the approach is used in the learning process to cultivate creative writing abilities, as Joyce (2003) described when combining the model with creative writing. One of the synectics model's strategies can be directly applied to creative writing, not only because it encourages the use of analogies but also because it helps writers "break set," allowing them to expand the range of devices they can employ to approach expressive tasks in expository, persuasive, and narrative genres. In addition, it has been applied to develop creative writing skills. For example, Ina, Sa, and Yayah (2020) applied the model to improve students' blank verse writing. They concluded that using synectic models in the experimental group increased students' creativity and comprehension in developing poetry material, which positively impacted their writing skills. Similarly, the use of the multiliterate synectic learning model improved students' capacity to write brief stories. As research by Nasrulloh and Waraulia (2022) utilized synectics to improve students' short story writing skills. Students' writing creativity was encouraged by implementing a multiliteracy synectic learning paradigm in Ponorogo, East Java, Indonesia junior high schools. The results indicated that the implementation of the multiliteracy synectic learning model was effective in enhancing the learning outcomes of students in the field of the Indonesian language, particularly in the fundamental skill of writing short stories. In addition, a research study compared the effectiveness of using the Synectics and journal techniques to enhance creative writing skills. This study was conducted by Fatemipour and Kordnaeej (2014). The findings indicated that both techniques positively impacted the development of students' creativity, with the Synectics technique showing better results than the journal writing technique.

In conclusion, Synectics is an essential technique or process that plays a significant role in fostering and developing various forms of creative writing skills.

#### 3. An Example of Creative Writing by Synectic integrated with generative AI.

The process of fostering creative writing skills through integrating Synectic model and generative AI consists of six steps, each incorporating educational technology. In this initial phase, students are introduced to the lesson and encouraged to create basic artworks, such

as drawings or photographs, based on predetermined topics. H5P is used to facilitate classroom participation and stimulate creative output. Steps 2 through 5, Innovating Language These stages encourage students to introduce novel vocabulary and concepts using Padlet as a learning tool, thereby nurturing interaction between students. Each learner contributed actively to the growth of language proficiency. Step 6, Creative Writing - In the final stage of creative writing, students compose original verses utilizing a variety of poetic formats. The compositions must correspond to the products produced in the first step. Moreover, the assignment is subsequently evaluated and improved through consultation with Chat GPT. In summary, this approach combines the benefits of Synectic instruction with AI technology in order to improve learners' creative writing skills, as shown in figure 1.



Figure 1. An example of creative writing by synectic integrate with generative AI.

#### 4. Methodology

#### 4.1 Research design

This research involves two groups of students. The design includes pre- and post-learning evaluations for both groups. The experimental group examines the influence of creative writing skills and students' perception of collaborative learning with ChatGPT. In contrast, the purpose of the control group is to assess the impact of creative writing skills and students' perception of lecture-based education. This study employs a quantitative research methodology to determine the research objectives using the outcomes of learning experiences, Overall Impressions, and students' creative writing abilities. As shown in figure 2.



Figure 2. The research design of this Study.

#### 4.2 Participants

This study examined how to encourage creative writing through poetry-based teachings. The participants who are currently studying in 10<sup>th</sup> grade were divided into two groups for the investigation. The Control Group (CG) comprised 30 students, while the Experimental Group (EG) comprised 34 students. The competitors were selected from the 10th-grade cohort of the medium-sized school. Both groups of students have the same readiness and ability to use technology, and their initial learning abilities are similar, as indicated by their pre-test scores. As creative writing exercises, the CG received traditional lectures and blank verse poetry. While the EG received instruction in creative writing using blank verse poetry and Synectics, a method for creative problem-solving. This strategy was supplemented with ChatGPT. Each EG learning session lasted six periods and included group collaboration via ChatGPT.

#### 4.3 Research Instruments

The tools in this study included a "creative writing test" and a perception questionnaire for studying the Thai language. Creative Writing exercise Three open-ended questions made up the test: question one assessed language usage and writing style; question two set comprehension of social and cultural settings; and question three assessed word meaning development and inventiveness. A total of 25 was obtained for all three things. The creative writing test was modified from the Creative Writing Marking Key by Morris & Sharplin (2013). The questionnaire on language learning perception toward technology-supported learning consists of 21 questions developed by Peng et al. (2009). It comprises two dimensions: (1) Learning Experiences and (2) Overall Impressions. The questionnaire has been translated into Thai.

#### 4.4 Student's ChatGPT Prompting.

This research utilized various technologies tools during the implementation including interactive videos from the H5P website, Padlet, and ChatGPT. After producing inventive blank verse through a variety of processes, students then seek guidance from ChatGPT. They began by posing a query such as, "Give me suggestions on how to enhance the following creative blank verse writing". ChatGPT then provides suggestions for enhancing their creative blank verse writing. For example, ChatGPT responses in two formats. In the first format, ChatGPT edited the creative writing from the students. They passed the feedback from the first format to enhance their writing because the objective was practicing them to produce independent writing. In the second format, ChatGPT provided suggestions to enhance creative writing in various aspects, such as word choice, appropriateness of word selection, conveying emotions, using evocative language, improving text comprehension, making the text more reader-friendly, and making simile comparisons. After receiving the suggestions, students refined and improved their creative blank verse writing. As shown in figure 3.



*Figure 3.* Collaborative work activities with chatGPT for providing suggestions to enhance creative writing development.

#### 4.5 Data Collection and Analysis

Students were given 15 minutes to complete a questionnaire regarding to their Learning Experiences and Overall Impressions. They were then given another forty minutes to complete a pre-learning creative writing test. Upon completing the pre-learning data collection, the experimental group of students participated in a structured learning intervention designed to improve their ability to write creative blank verses. This intervention incorporated the Synectic learning strategy with various educational technologies, with ChatGPT as the focal point. During this intervention, a teacher instructed students on how to use ChatGPT and compose inventive blank verses. The students were then given a post-learning questionnaire and examination. As shown in figure 4.



Figure 4. The activity of creative writing by synectic integrates with generative AI.

The selected statistical techniques for evaluating creative writing skills include the Mann-Whitney U test for comparing the outcomes of creative writing skills resulting from the instructional approach of creative blank verse writing incorporating Synectic learning and generative AI. In addition, MANOVA was used to compare learning students between the Synectic model with generative AI and lecture base learning on students' perception.

#### 5. Result and Discussion

#### 5.1 Creative Writing Skill

To examine the influence of the Creative Writing learning approach by integrating Synectic with generative AI on students' creative writing skills, the Mann-Whitney U test was employed in this study, as indicated in Table 1.

Table T. Mann-Williney	U lesuit Cre	ealive writing Skill		
Groups	Ν	Mean ± SD	Z	p-value
Control group	30	18.30 ± 3.659	-2.276	.023*
Experimental group	34	20.44 ± 3.268		
* n = < 0.5				

Table 1. Mann-Whitney U result Creative Writing Skill

\* p= <.05

Table 1 demonstrated a statistically significant difference between EG and CG on posttest scores regarding creative writing skills. The result of experimental group's average posttest score was significantly higher than the control group. This implied that the instructional approach by incorporating Synectic learning with Generative AI could improve the learners' creative writing abilities. The abovementioned analysis is consistent with the study conducted by Ina, Sa, and Yayah (2020), where they applied the Synectics approach to enhance students' blank verse writing. The results of the post-test analysis indicated significant differences in poetry writing ability between the experimental and control groups. Similarly, the research by Su, Lin, and Lai (2023) investigated the integration of ChatGPT in argumentative writing within the classroom. The research revealed that the utilization of ChatGPT had a significant effect on facilitating the development of creative writing abilities among students in both experimental and control groups, as evidenced by the observed outcomes.

#### 5.2 Students' Perception toward Technology-Supported Learning

This study utilized a MANOVA test (as shown in Table 2) to investigate the effect of the Creative Writing learning approach incorporating synthetic with generative AI on students' perceptions toward technology-supported learning.

Dimension	Group	Ν	Mean	S.D.	Sig.			
Learning Experiences	CG	30	28.57	3.645	.063			
	EG	34	30.09	2.778	-			
Overall Impressions	CG	30	41.63	6.499	.001*			
	EG	34	46.88	5.881	-			

Table 2. MANOVA result Students' Perceptions toward Technology-Supported Learning

\* p= <.05

Considering Table 2 showed that there was no significant difference on learning experiences of students' perceptions toward technology-supported learning between EG and CG. Despite differences in instructional formats, the study findings indicate that both groups share similar Learning Experiences. It is perhaps due to the student in CG using their mobile device to search for more information by themself during the learning activity. This could affect their learning experience because they could find the same information that ChatGPT provided. In contrast, there was a significant difference between the groups' Overall Impressions. The statistics mentioned above support the findings of Liu et al. (2021), who discovered that implementing a reflective thinking promotion mechanism within AI-supported English writing environments increased self-efficacy and self-regulated learning.

#### 6. Conclusion

This study investigated the effects of students' perceptions toward technology-supported learning and creative writing skills on the approach which combines Synectic and generative AI. The findings revealed that this learning strategy affected the development of student's creative writing skills and their Learning Experiences and Overall Impressions. The primary findings of this study indicate that integrating the Creative Writing learning approach of Synectic with generative AI (ChatGPT) has a positive impact on the development of creative writing skills and the improvement of students' perceptions toward technology-supported learning. However, this research has limitations in terms of the time required for data collection because, during the creative writing teaching activities, it was found that students needed time to revise and refine their language expressions to achieve the desired quality. Moreover, qualitative data collection is suggested to explicit how the learners' improved their creative writing ability when participated with the generative AI.

#### Acknowledgement

This work extends its gratitude to the research team for their insightful suggestions and collaborative efforts in research development. Additionally, thank you to Khon Kaen University for their supportive contribution.

#### References

- Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*. Advance online publication. https://doi.org/10.1016/j.mlwa.2020.100006
- Chambers, F. (1997). What do we mean by fluency?. *System, 25*(4), 535-544. https://doi.org/10.1016/S0346-251X(97)00046-8
- Fatemipour, H., & Kordnaeej, M. (2014). The effect of synectics and journal creative writing techniques on EFL students' creativity. *International Journal of Applied Linguistics and English Literature*, *3*(6), 1-9.
- Gordon, W. J. J. (1961). Synectics. New York: Harper and Row Publication.
- Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100001. https://doi.org/10.1016/j.caeai.2020.100001
- Joyce, B., & Weil, M. (2003). Models of teaching. Fifth Edition. Prentice-Hall, Inc.
- Li, Q., Lau, R. W., Wah, B. W., Ashman, H., Leung, E. W., Li, F., & Lee, V. (2009). Guest editors' introduction: emerging internet technologies for e-learning. *IEEE Internet Computing*, *13*(4), 11–17.
- Liu, C., Hou, J., Tu, Y. F., Wang, Y., & Hwang, G. J. (2021). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments*, 1-19. https://doi.org/10.1080/10494820.2021.2012812
- Morris, G., & Sharplin, E. (2013). The assessment of creative writing in senior secondary english: a colloquy concerning criteria. *English in Education*, 47(1), 49–65. https://doi.org/10.1111/eie.12004
- Nasrulloh, A., & Waraulia, A. (2020). Synectics multiliteracy learning model in writing short stories. Journal of Language Teaching and Research, 11(6), 545-551.
- OpenAI. (2022, November 30). Introducing ChatGPT. OpenAI. https://openai.com/blog/chatgpt
- Peng, H., Chuang, P.-Y., Hwang, G.-J., Chu, H.-C., Wu, T.-T., & Huang, S.-X. (2009). Ubiquitous performance-support system as mindtool: a case study of instructional decision making and learning assistant. *Educational Technology & Society, 12*(1), 107–120.
- Shidiq, M. (2023). The use of artificial intelligence-based chat-GPT and its challenges for the world of education; from the viewpoint of the development of creative writing skills. *Proceeding of International Conference on Education, Society and Humanity, 1*(1), 353-357.
- Song, X., Su, Y., Ren, J., & Gong, R. (2021). Effects of group awareness tools on student engagement and enjoyment in online collaborative writing. *Journal of Educational Computing Research*, *5*9(7), 1195-1217.
- Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with chatGPT in argumentative writing classrooms. Assessing Writing, 57, 100752. https://doi.org/10.1016/j.asw.2023.100752

## Generating Interactive Stories with ChatGPT to Teach Filipino Values

#### Angelo Miguel GREGORIO, Sarah Jessica MANUEL, Alyssa Jaye PALMARES, Sharlin Mae TANG & Ethel ONG<sup>\*</sup>

De La Salle University - Manila, Philippines \*ethel.ong@dlsu.edu.ph

**Abstract:** Interactive storytelling uses digital media to engage readers to interact with the author's narrative through key decision points in the story. This has been shown to be an effective learning resource in teaching across various disciplines. Designing interactive storytelling systems typically require handcrafted storytelling knowledge and branching story graphs. Generative AI offer an opportunity for facilitating the production of interactive stories. In this research, we utilized ChatGPT to provide the storytelling knowledge for Kulturuan, a system that teaches the Filipino value of *pakikipagkapwatao* or hospitality to grade school students. Prompts are formulated to instruct ChatGPT to generate bits and pieces of the story text. Results from experiments and refinement of carefully formulated prompts showed ChatGPT's potential as a knowledge resource for interactive storytelling systems, thus eliminating the need to expend time and effort in building a separate story graph. However, further downstream tasks are needed to process the resulting story text to fit the interactive environment of Kulturuan.

Keywords: Interactive storytelling, Filipino values, Generative AI, Prompt Formulation

#### 1. Introduction

Stories predominate traditional Filipino literature as a means of imparting cultural values. To preserve and give importance to local traditions and values, the Philippine Republic Act No. 9155 or the *Governance of Basic Education Act of 2001* required the formal inclusion of the arts and culture in the basic education curriculum to instill love and pride in the country's heritage among Filipino learners. While younger generations are aware of the importance of conserving traditional Filipino culture, their awareness does not reflect their preference in which culture they choose to practice (Tabuena et al., 2022).

The introduction of digital media encouraged the notion of utilizing technology in promoting culture and enabling an interactive access to arts. This corresponds to a shift in focus from oral stories to interactive storytelling. *Adventures of Ellie* is a 2D learning environment that uses interactive stories to prepare children with autism spectrum disorder (ASD) for social situations (Ong et al., 2016). *Sarah* is a virtual peer that shares interactive healthcare stories to raise awareness on proper healthcare practices among children (Lee et al., 2019). *Amy* is a virtual talking companion that supports the social-emotional learning of children with ASD through the sharing of interactive narratives (Gagan et al., 2023).

While these studies reported the potential impact of interactive storytelling on learning, most systems are not centered on teaching culture and values. They are also mostly written in English, making such technologies inaccessible to those who are not proficient in the language. Furthermore, these systems utilize manually-built resources and predefined branching story graphs (Riedl & Young, 2006) to provide the necessary knowledge needed in the generation of stories to be imparted to the children.

Recently, large language models trained on huge amount of human-authored text have facilitated the availability of AI systems that can understand and generate human language (Radford et al., 2018). Among these is ChatGPT which produces fluent text on varying genre and disciplines, including stories (Ouyang et al., 2022; Taecharungroj, 2023). Recent studies

have also reported the opportunity in using ChatGPT as a learning resource in education (Tlili et al., 2023) and as a collaborative writing assistant (Duval et al., 2021; Pavlik, 2023).

In this paper, we describe the strategies we employed in formulating prompts to allow our interactive storytelling system, *Kulturuan*, to work with ChatGPT in the generation of culture-based story in the Filipino language. The resulting interactive stories depict the value of *pakikipagkapwa-tao* or hospitality based on themes prescribed in the Department of Education's *Good Morals and Right Conduct curriculum* (DepEd, 2016) for basic education.

#### 2. Related Work

#### 2.1 Interactive Storytelling

Interactive storytelling (IS) allows a reader to interact with an author's narrative, giving the former a key role in the virtual story world and setting that the latter created. The author defines the elements and the plot of the story in the form of branching narratives or story graphs (Riedl, 2004). A branching graph represents the multiple paths and outcomes that can occur in the story (Smed et al., 2021). Usually pre-scripted at design time, these graphs allow readers to make decisions at key points in the story by selecting from a set of candidate events, responding to some events in the story world, giving instructions and advice to other characters (Cavazza et al., 2002), and engaging other characters in a dialogue.

The story elements – setting, characters, objects, and events – collectively form the story world which dynamically changes as the story progresses. A 2D or 3D interactive virtual world interface allows the reader to participate in the story. Participation is usually in the form of a virtual player who can perform actions in the story world, talk to other characters, and manipulate story world objects. The reader can also act as a spectator who navigates the story world, or an adviser who intervenes with the characters' plans (Mateas & Stern, 2003).

An interactive story also follows the common five-act story structure that divides the narrative into five parts: exposition, rising action, climax, falling action, and resolution. These parts cover the introduction of the characters and the setting, the description of the complication or the conflict, the events leading to the turning point of the story, the unraveling of the conflict and resolution of the protagonist's struggles, and the final outcome of the story.

#### 2.2 Generative AI

ChatGPT is a language model (LM) based on the generative pre-trained transformer (GPT) model developed by OpenAI to enable computers to process human language text (Radford et al., 2018). Trained on a huge amount of human-authored text, the model has been shown to perform well in generating fluent text in various topics, genre, and languages. This led to studies that encourage the use of ChatGPT in education, notably as a writing assistant (Duval et al., 2021; Pavlik, 2023; Tlili et al., 2023), amidst concerns on plagiarism, misinformation, and academic misconduct (Cotton et al., 2023; Sullivan et al., 2023; Uzun, 2023).

Prompts are the primary mechanism for human users to communicate with LMs like ChatGPT. Prompts can be in the form of a question, a statement, or an instruction. Recent studies reported the importance of formulating effective prompts to maximize the capabilities of the generative model while avoiding its pitfalls (GPTBot, n.d.; McCoy, n.d.; WeThinkApp, 2023). A set of design principles in formulating prompts recommend the statement of a clear description of the task, explicit constraints, and context (Mayo, 2023). The quality of the LM's responses can also be improved through iterative refinement and follow-up prompts. Like human-human communication, prompts can facilitate chatbot-human communication, allowing ChatGPT to better understand the user's intent expressed through the query and correspondingly generate better quality responses (Mayo, 2023). Prompts can also unlock the LM's creativity and minimize the generation of incorrect responses (GPTBot, n.d.).

#### 3. Interactive Storytelling with Generative AI

Kulturuan aims to impart the Filipino value of *pakikipagkapwa-tao* or hospitality to Grades 1 to 3 children through interactive stories. The stories are set in three possible locations, namely home, school, and outdoors; allowing users to navigate the virtual spaces in each location and to interact with different story characters, such as family members, friends, neighbors, visitors, foreigners, and strangers. These interactions are intended to provide a platform for the children to experience real-life situations where *pakikipagkapwa-tao* can be practiced.

Kulturuan uses ChatGPT as its knowledge resource in the generation of stories. Prompts written in English are sent to the language model through the OpenAl Playground. The model then generates story text in Filipino that narrates the setting, character dialogue, decision points or choices, consequences of user choices, and story ending. To avoid overloading young children with too much textual information, the responses of ChatGPT are parsed to extract only portions of relevant story text that are then displayed in Kulturuan's web interface as illustrated in Figure 1. Because the response of ChatGPT is assumed to be correct, it is imperative that Kulturuan formulates relevant prompts to facilitate the generation of appropriate story text. In this section, we describe the types of prompts that we formulate to enable ChatGPT to generate useful story text for Kulturuan.



*Figure 1.* Kulturuan's user interface with story text extracted from ChatGPT's response.

#### 3.1 Prompt Formulation

Prompt engineering is the process of creating prompts to interact and elicit useful responses from ChatGPT (Mayo, 2023). To facilitate the generation of different story text, three types of prompts are formulated: *system message prompt, story prompt, and event prompt*.

A system message supports multi-threaded prompts while preventing the need to repeat the same instructions to be sent to ChatGPT via the OpenAl Playground. Table 1 shows the different parts of a system message prompt: (i) description of the task that the model has to perform; (ii) persona that allows the model to act accordingly with the given task description; (iii) output specifications indicating the target audience, story structure, and decision points; (iv) output constraints which include the language and the length; and (v) templates to be followed in generating the decisions points and the character dialogues.

A *story prompt*, shown in Table 2, sets the story world and its elements – the theme, characters, setting, and plot. Important specifications and constraints are reiterated in this prompt, such as the language and use of characters and locations. An *event prompt* indicates the player's choice among the candidate events generated by ChatGPT.

#### 3.2 Response Parsing

The resulting text from ChatGPT has to undergo further processing before it can be presented to the children. A downstream task, *response parsing*, is essential in removing text that is unnecessary to the overall story, i.e., text written in ChatGPT's point-of-view and labels such

as *Exposition* and *Climax*. The parser works by finding elements of the story text from ChatGPT's response and validating that the text includes mentions of characters names and location indicated in the prompt. The parsed story text is then presented to the children through an interface as seen in Figure 1. Presenting short passages of text can help understand and digest the pieces of information, as well as respond to questions posed by the story characters.

System Message	Parts
You are a <storyteller></storyteller>	Persona
who writes stories	Task
for <children 6="" 9="" aged="" to="">.</children>	Specification – Target audience
Only use the <filipino> language in your responses.</filipino>	Constraint – Language
You will follow a <5-part story structure> and each part must have at most <150 words>.	Specification – Story structure Constraint – Length
Each part must be <generated a="" at="" one="" time=""> in a separate response; <do a="" been="" choice="" generate="" has="" made="" next="" not="" part="" the="" unless="">.</do></generated>	Constraint – Response structure
Here are the five parts:	Specification – Decision points
Exposition	
Rising action. This part contains the 1 <sup>st</sup> decision point	
Climax	
Falling action. This part contains the 2 <sup>nd</sup> decision point	
Resolution	
Present the decision point in this format:	Template – Decision point
Ano ang gagawin ni Justine? (What should Justine do?)	
Positive choice: < nrovide a short description>	
Nogative choice: <provide a="" description="" short=""></provide>	
Descent a characteria dialegue in this formati	Taranlata Charactan dialamua
Present a character's dialogue in this format:	remplate – Character dialogue
Unaracter: Dialogue	

Table 1. Components of a System Message Prompt

 Table 2. Components of a Story Prompt

Story Prompt	Component
Remember that the story should be written in Filipino, and only	Language reminder
use the characters and locations you have been provided with.	
The main character of the story is a child named <justine> and</justine>	
the story takes place in the <sub-location> of the <main-< td=""><td>Main character, location</td></main-<></sub-location>	Main character, location
location>. The story's theme is practicing good hospitality to	Story theme
others. Let's start with Justine and <npc>.</npc>	



Figure 2. Parsed story text with character dialogue.

Since the prompt to ChatGPT includes the specification of the story structure, finding these elements from the model's response and displaying the corresponding UI sprites should be straightforward. To determine if a story text is to be narrated (Figure 1) or presented as a character dialogue (shown in Figure 2), the response parser checks if the LM's response stated that a character is talking. Figure 3 illustrates how a decision point is presented to the user. By parsing ChatGPT's response this way, children will be able to see how their decisions, whether positive or negative, could influence the outcome of the story.



Figure 3. Parsed story text with a decision point.

#### 4. Experiments

We conducted controlled experiments in formulating and refining our prompts to increase the quality and relevance of ChatGPT's generated story text.

#### 4.1 Basic Story Text

Instructing ChatGPT to generate an interactive story requires a prompt as simple as "*Create an interactive story for children*". The model then responds with a short story with a set of decision points, as seen in Table 3.

Table 3. Sample interactive story generated by ChatGPT from a simple prompt "Create an interactive story for children" with no constraints.

Title: The Enchanted Forest Adventure

Once upon a time, in a small village, there lived a curious and brave little boy named Leo. One sunny morning, Leo set out on an exciting adventure to the Enchanted Forest, a magical place filled with talking animals, hidden treasures, and enchanting creatures.

As Leo wandered deeper into the forest, he stumbled upon a talking squirrel named Sammy. Sammy greeted him with a cheeky smile and said, "Welcome to the Enchanted Forest, young explorer! To continue your adventure, you must solve a riddle."

Leo was intrigued and eagerly accepted the challenge. Sammy cleared his throat and asked, "What has a heart that doesn't beat?" A) A rock B) A tree C) A fish D) A bear

If you think the answer is A) A rock, turn to page 5. If you think the answer is B) A tree, turn to page 6. If you think the answer is C) A fish, turn to page 7. As seen in the example, the default language used by ChatGPT is English and the story text exceeds 150 words. The different parts of the story are also not clearly marked. To address these, the simple prompt is refined to indicate the target language; the specific names of characters, location and event that should appear in the story; the response length; and the 5-part story structure. Aside from the system message shown in Table 1 that sets the stage for story generation with ChatGPT, story prompts and event prompts are sent to the LM at each dialogue turn to request for specific story text. Shown in Table 4 is a story prompt that requests ChatGPT to generate the *exposition* text. Also included in the prompt is an explicit constraint to "*only use the characters and locations you have been provided with*" to avoid the sudden appearance of story characters without their proper introduction.

Table 4. Sample story prompt and corresponding story text generated by ChatGPT.

User	Remember that the story should be written in Filipino, and only use the characters and locations you have been provided with. The main character of the story is a child named Justine and the story takes place in the Living Room of the House. The story's theme is practicing good hospitality to others. Let's start with Justine and Uncle Renato. Justine offers a drink to Renato.
ChatGPT	Exposition: Isang maganda at mainit na araw. Nasa living room si Justine kama and kanyang Uncle Renato. Si Justine ay isang masayahing bata na laging handang tumulong sa ibang tao. Si Uncle Renato naman ay isang masipag na tao na madalas bumibisita sa kanilang bahay upang magkuwentuhan.

#### 4.2 Decision Points and Choice Generation

Interactive stories include key decision points where the children are presented with choices that influence how the story will move forward. In Kulturuan, the decision points are presented in the *rising action* and *falling action*. We conducted experiments to evaluate how ChatGPT generates choices, paying particular attention to the format and types of choices in the response. The following constraints in the system message prompt restricts the LM to conform to the specified template in generating a positive and a negative choice to enable the response parser to extract the choices and present these in the user interface.

Present the decision point in this format: Ano ang gagawin ni Justine? (What should Justine do?)

Positive choice: <provide a short description> Negative choice: <provide a short description>

Table 5 shows an example. Also illustrated in the table is the structure of a dialogue in ChatGPT's response, which is used by the response parser to determine if a given story text is narrated or delivered as a dialogue.

Table 5. Sample story text containing decision points generated by ChatGPT.

Rising Action:

Justine: Uncle Renato, gusto niyo po bang inumin ang aking gawaing juice? Ginawa ko po ito para sa inyo. Uncle Renato: Salamat, Justine. Mabuti ka talagang bata. Oo, gusto ko nga ng isang mainit na inumin ngayong maininit na araw. Maraming salamat sa'yo.

Ano ang gagawin ni Justine?

Positive Choice: Ihandog kay Uncle Renato and gawing juice na may ngiti sa kanyang labi. Negative Choice: Sabihin kay Uncle Renato na wala nang juice tumawid na lang sila sa kabilang bahay. The choice of the child is formulated as a prompt and sent to ChatGPT to request for the outcome of such a decision. An example is shown in Table 6.

Table 6. Sample story prompt with the child's choice and the subsequent story text.

User	I choose the positive choice, continue the story.					
ChatGPT	Climax: Justine: (may ngiti sa kanyang labi) Ito po, Uncle Renato. Sana'y magustuhan niyo ang aking gawaing juice.					
	Uncle Renato: (kinuha ang baso ng juice at ngumiti) Maraming salamat, Justine. Ang sarap ng juice mo! Talagang napasaya mo ako ngayong araw.					
	Nagpatuloy ang kanilang kwentuhan habang nag-eenjoy sila sa juice na handog ni Justine. Habang nagmamadali ang oras. Biglang may narinig silang kalabog mula sa kusina.					
	Falling Action: Justine:(nag-aalala) Uncle Renato, may narinig akong kalabos sa kusina. Baka may nagyaring hingi maganda. Dapat po ba tayong pumunta at tingnan kung ano iyon?					
	Ano ang gagawin ni Justine?					
	Positive Choice: Tumakbo papunta sa kusina upang masiguradong walang problema. Negative Choice: Tumuloy lang sa kwentuhan at hayaan na lang ang kalabog sa kusina					
User	I choose the negative choice, continue the story.					
ChatGPT	Climax: Uncle Renato: Huwag kang mag-alala, Justine. Baka may tao lang na nagkamali ng lagay ng mga kawali sa kusina. Hayaan na nating magpatuloy ang ating kwentuhan.					
	Nagpatuloy sila sa kanilang kwentuhan at hindi na pinansin ang kalabog sa kusina. Ngunit habang nagpatuloy ang kanilang pag-uusap, biglang dumating si Auntie Marissa, ang ina ni Justine.					

It was noticed that ChatGPT may at times deviate from following the prescribed 5-part structure for indicating the parts of the story and presenting the choices. An instance of this is depicted in Table 7 where the LM presented the full story as one response, given the same story prompt shown in Table 4. Addressing this deviation will be part of our future work.

#### 4.3 Vocabulary

Stories can serve as an avenue for children to acquire language and increase their vocabulary. Thus, it is essential for ChatGPT to produce story text that uses Filipino words suitable for the intended audience. Tests were conducted to determine if ChatGPT uses Filipino vocabulary that are familiar for young children. Prompts containing the story characters and location were created. Table 7 illustrates the responses from the LM with the target audience ("for children") explicitly indicated in *Test Case 2*. As seen in column 2, *Test Case 2* uses simpler words to describe the stranger, i.e., "isang taong di niya kilala" (a person he does not know) as opposed to the unfamiliar word "estranghero" (stranger). an example response from the LM using an uncommon word in Filipino when it translated the word "stranger" into "estranghero," which can be complicated for children to understand.

Table 7. Sample response of ChatGPT containing the full story text and disregarding the given structure and template.

Noong isang malamig na gabi, nasa sala ng bahay sina Justine at Uncle Renato. Si Justine ay isang batang mabait at maalalahanin na laging handang tumulong sa ibang tao. Siya ay may malambot na puso at gusto niyang iparamdam sa lahat ng kanyang nakakasalamuha ang mainit na pag-aalaga at kabaitan.

"Uncle Renato, mainit po ba kayo?" tanong ni Justine na may ngiti sa kanyang labi.

"Talaga bang mayroon kang mainit na inumin, Justine?" sabi ni Uncle Renato habang inaayos ang kanyang mga damit. "Salamat naman sa inyong pag-aalala."

"Opo, Uncle Renato! Mayroon po akong tsaa na nilagang saging. Gusto niyo po ba?" sagot ni Justine na puno ng excitment.

Napangiti si Uncle Renato at nagpatuloy sa pag-aayos ng kanyang mga sapatos. "Salamat, Justin. Tunay nga ngang napakabuti mong bata."

•••

Inilapag ni Justine ang tasa sa mesa at umupo sa tabi ni Uncle Renato. Habang umiinom sila, nagkuwentuhan sila tungkol sa kanilang mga karanasan at mga pangarap. Napuno ang sala ng tawa at ligaya, at sa bawat sandali, lumalalim ang pagkakaibigan nina Justin at Uncle Renato.

Sa simpleng pag-aalaga ni Justine at sa pagiging magiliw ni Uncle Renato, naipakita nila ang kahalagahan ng mabuting pananawagan at kabaitan sa isa't isa. Ipinakita nila na sa pamamagitan ng simpleng pagbibigay ng inumin, nagawa nilang palakasin ang kanilang samahan at magbigay ng kasiyahan sa isa't isa.

....

Prompt	ChatGPT
<i>Test Case 1.</i> Create a sentence in Filipino about interaction between Justine and a stranger.	Nakita ni Justine ang <u>isang estranghero</u> na naglalakad sa kalsada at nagpasya siyang lumapit sa kanya.
<i>Test Caes 2.</i> Create a sentence <b>for children</b> in Filipino about interaction between Justine and a stranger.	Nakita ni Justine ang <u>isang taong di niya</u> <u>kilala</u> na naglalakad sa kalsada at nagtatakang kung ano ang sasabihin nito.

Table 8. Comparison of output using prompts with and without target audience indication.

#### 5. Discussion

Our study makes an innovative attempt at integrating Generative AI, particularly ChatGPT, with the domain of interactive storytelling for educational purposes. Much of the potential that generative AI can contribute to this field lies in the model's ability to produce relevant story text with the target theme to impart cultural values. For this to happen, it is imperative that prompts are properly formulated to direct ChatGPT in performing the expected task and response. Following the design principles outlined in Mayo (2023) and (Zhao, 2023), our experiments confirmed the essential guidelines below in the formulation of prompts.

**Clear task descriptions**. Prompt should give a clear description of the task to enable ChatGPT to carry out the instruction and generate the desired response. In Kulturuan, these descriptions include the role of the LM (storyteller), input data (story character, location), expected output (story), and response restrictions.

**State explicit constraints**. Constraints include (i) the format of the expected response, i.e., 5-part story structure; (ii) the intended audience, i.e., Filipino children who are between 6 - 9 years old; (iii) the language; and (iv) the response length.

**Provide more context**. Contextual information allows the LM to understand the scope and purpose of the request. In Kulturuan, these are the specification of the 5 parts of the story and the decision points, i.e., two choices for the activities that a child may do in the living room.

**Decompose the task**. It is also important to decompose instructions into smaller and more detailed subtasks to avoid overloading the LM with information.

**Issue follow-up prompts**. When a prompt is decomposed into smaller subtasks, followup prompts can be issued that make additional requests to form a cohesive conversation thread. This is particularly useful in our study where we ask ChatGPT to formulate different parts of an interactive story, such as the setting, the problem or conflict, and the resolution.

**Refine the prompt**. To improve the quality and relevance of the generated story text, prompts must be iteratively refined through experimentation. While this does not guarantee that ChatGPT will consistently generate good quality and relevant story text, experiments can shed insights on the extent of the LM's capabilities and limitations, allowing downstream tasks to address the gap in ChatGPT's response generation and storytelling abilities.

Through prompt engineering, generative AI can bridge narrative science with natural language processing strategies to deliver an interactive learning environment for children to learn about culture and values. While our study barely scratched the surface on prompt engineering, conducting user experience studies among the target audience can provide empirical evidence on the impact of Kulturuan on the students' learning performance, engagement, and appreciation of the hospitality value. Interviews with educators and writers of children's stories may also provide valuable feedback on the relevance of an AI-based technology in cultural narrative generation.

#### 6. Conclusion

Kulturuan is an interactive storytelling environment designed to deliver the Filipino value of *pakikipagkapwa-tao* or hospitality to children who are 6 to 9 years old. Generative AI, specifically ChatGPT, is utilized as the knowledge resource in the generation of the interactive stories. Controlled experiments highlighted the importance of properly formulated prompts in enabling the model to produce story text that are relevant to the story theme, leading to a set of guidelines in prompt engineering. However, downstream tasks such as a response parser are needed to process the resulting text of ChatGPT to fully align with the requirements of the Kulturuan environment. Further work includes enhancing the performance of the *response parser* in verifying the consistency and controlling the randomness of ChatGPT's generated text. Future empirical evaluation with child educators and literary experts can lead to valuable insights on the quality of the resulting stories in imparting the target values and the overall relevance and effectiveness of the system as an environment for learning. Children can share their perception on the use of Kulturuan as an educational resource to understand and appreciate Filipino culture and values.

#### References

- Bravo, C. D., Dimalanta, F. D., Jusay, K. A., Vitug, M. Y., & Tabuena, A. (2022). Inclination state on the Philippine culture and arts using the Appraisal Theory: Factors of progress and deterioration. *Participatory Educational Research*, 9 (1), 388-403. doi: 10.17275/per.22.21.9.1.
- Cavazza, M. O., Charles, F, & Mead, S.J. (2002). Character-based Interactive Storytelling. *IEEE Intelligent Systems*, 17(4), 17-24.
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*. doi:10.1080/14703297.2023.2190148.

DepEd (2016). *K to 12 Curriculum Guide: Edukasyon sa Pagpapakatao*. Department of Education Learning Portal. Retrieved July 10, 2023 from https://lrmds.deped.gov.ph/detail/5451

Duval, A., Lamson, T., & de Kerouara, G. d. L. (2021). Breaking writer's block: Low-cost fine-tuning of natural language generation models. *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, pp. 278–287.

Gagan, I. T., Matias, M. A. M., Tan, I., Vinco, C. M., & Ong, E. (2023). Preparing children with Level 1 ASD for social interactions through storytelling with Amy: An exploratory study. *Proceedings of CHI Conference on Human Factors in Computing Systems Extended Abstracts*. ACM, New York. doi:10.1145/3544549.3585916.

GPTBot (n.d.). Mastering ChatGPT: How to craft effective prompts (full guide). https://gptbot.io/master-chatgpt-prompting-techniques-guide/

- Lee, R. G., Regala, J. B., Tan, A., Tan, J. M., & Ong, E. (2019). Learning about good healthcare practices through interactive storytelling with virtual peers. *Philippine Computing Journal*, 14(1), 15-24. Computing Society of the Philippines.
- Mateas, M., & Stern, A. (2003). Façade: An Experiment in Building a Fully-realized Interactive Drama. In *Proceedings of the Game Developers Conference*, San Jose, CA.
- Mayo, M. (2023). Prompt Engineering for Effective Interaction with ChatGPT. https://machinelearningmastery.com/prompt-engineering-for-effective-interaction-with-chatgpt/
- McCoy, J. (n.d.) The inside scoop: The secret to ChatGPT prompts for amazing results. https://contentatscale.ai/secret-to-chatgpt-prompts/.
- Ong, E., Consignado, D. G., Ong, S. J., & Soriano, Z. (2016). Building a semantic ontology for virtual peers in narrative-based environments. In M. Numao et al. (Eds.), Trends in Artificial Intelligence PRICAI 2016 Workshops, LNCS 10004, pp. 65-76. Springer International Publishing AG.
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. arxiv.org. doi:10.48550/arXiv.2203.02155.
- Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. Journalism & Mass Communication Educator. doi:10.1177/10776958221149577.
- Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. OpenAI. https://cdn.openai.com/research-covers/language-unsupervised/language\_understanding\_paper.pdf.
- Riedl, M. (2004). Narrative Generation: Balancing Plot and Character. *PhD dissertation*, North Carolina State University.
- Riedl M. O., & Young, R. M. (2006). From linear story generation to branching story graphs. *IEEE Computer Graphics and Applications*, 26(3), 23-31. doi:10.1109/MCG.2006.56.
- Smed, J., Suovuo, T., Skult, N., & Skult, P. (2021). *Handbook on Interactive Storytelling*. Wiley. https://doi.org/10.1002/9781119688167.
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023) ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6(1). doi:10.37074/jalt.2023.6.1.17.
- Taecharungroj, V. (2023) "What can ChatGPT do?" Analyzing early reactions to the innovative Al chatbot on Twitter. *Big Data and Cognitive Computing*, 7(1). doi:10.3390/bdcc7010035.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10, 15. doi: 10.1186/s40561-023-00237-x.
- Uzun, L. (2023) ChatGPT and academic integrity concerns: Detecting artificial intelligence generated content. *Language Education and Technology*, 3(1), 45-54.
- WeThinkApp (2023). Guide for crafting clear and effective prompts for ChatGPT prompting. https://wethinkapp.medium.com/the-art-of-chatgpt-prompting-a-guide-to-crafting-clear-andeffective-prompts-1dfc2589295b.
- Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., & Wen, J.-R. (2023). A survey of large language models. arXiv.org. https://arxiv.org/abs/2303.18223,

# W02

## The 7th Computer-Supported Personalized and Collaborative Learning

## The Effect of Gamification with Self-Regulated Approach to Promoting Nursing Students' Leopold's Maneuvers Performance

#### Intan SETIANI<sup>a\*</sup>, Ching-Yi CHANG<sup>b</sup> & Jie-Chi YANG<sup>c</sup>

<sup>a</sup>Graduate Institute of Network Learning Technology, National Central University, Taiwan <sup>b</sup>School of Nursing, College of Nursing, Taipei Medical University, Taiwan <sup>c</sup>Graduate Institute of Network Learning Technology, National Central University, Taiwan \*<u>intansetiani@g.ncu.edu.tw</u>

Abstract: Fetal appearance and proper position evaluation are crucial for determining the best labor treatments. Healthcare or midwifery nurses, require knowledge about the fetus's position and health to make decisions on whether to induce labor. In addition, attending to the patient's needs throughout labor is essential. The fetal positioning and monitoring process may be uncovered with the use of Leopold's techniques. Students find it challenging to apply what they learn in the classroom to real-world challenges since standard training lacks chances for clinical experience and application. This study aims to investigate the impact of integrating online game-based environments with gamification and a self-regulated learning (SRL) strategy on nursing students' Leopold's maneuvers skills. The use of games as a learning tool in education has gained popularity due to its ability to provide a playful and motivating experience for students, as well as fostering cooperation among them. However, the application of game-based learning and gamification in nursing skill instruction remains understudied. The researchers propose integrating an online game-based environment called Gather with gamification features and a self-regulated learning strategy for nursing students. The self-regulated learning approach in this study includes goal setting, performing tasks with gamification elements such as points, badge, leaderboards, and self-evaluation. An experiment will be conducted at a nursing school in a Taiwan university using a quasi-experimental design to investigate the learning performance, motivation, self-efficacy and perceptions.

**Keywords:** Gamification, online game-based environment, self-regulated learning, Leopold's maneuvers, nursing education.

#### 1. Introduction

An accurate assessment of the fetus's presentation and position is crucial for determining the exact time of birth (McFarlin et al., 1985). The four Leopold's maneuvers identify fetal orientation (Mirghani et al., 2007). This technique gives expecting mothers a sense of their child's position in the womb, promoting awareness of the growing fetus (Nishikawa & Sakakibara, 2013). Nursing education emphasizes the importance of training in fundamental competencies such as knowledge, skills, and attitudes (Hwang & Chang, 2020). Nurses employ various interventions, such as Leopold's maneuvers, to strengthen maternal bonding and increase fetal awareness (Brandon et al., 2009). However, recognizing an occiput posterior fetus challenges traditional teaching and learning methods (Biancuzzo, 1993). Traditional training lacks clinical practice and application opportunities, making it difficult for students to apply classroom knowledge to real-world problems (Chang et al., 2020).

Furthermore, research has indicated that the implementation of successful pedagogical approaches is imperative for improving academic performance among students

(Chang et al., 2020). The researcher establishes a correlation between variances in students' learning abilities and insufficiency in self-regulatory skills. This viewpoint prioritized the identification of students' self-awareness to manage their boundaries and pursue their academic pursuits effectively. While educators need to know their students' learning abilities, their primary objective should be facilitating students' self-awareness of these differences. According to Zimmerman (2002), if students encounter difficulty comprehending a particular aspect of a course during class, it is incumbent upon them to exhibit self-awareness and strategic understanding to implement corrective measures. Hence, it is imperative to facilitate nursing students' learning through a self-regulated strategy.

The use of game-based learning and various forms of gamification is increasingly gaining traction within the realm of education. One of the rationales behind utilizing games as a learning tool is that it provides a playful experience that provides a source of motivation for students. Another factor pertains to the games' social dimension, which fosters student cooperation (Vidergor, 2021). However, this strategy has been investigated infrequently for nursing skill instruction (Chang et al., 2022). Previous research has demonstrated the effectiveness of online game-based learning (Yang et al., 2021). Using games as an educational tool has had positive results, as evidenced by documented improvements in academic performance, student motivation, and diligence across various fields of study (Barata et al., 2017). Gamification is a new method that tries to motivate and engage people (Bozkurt & Durak, 2018). To further understand the impact of certain gamification features or designs, academic research is increasingly looking to motivation, behaviour, and learning theories (Nacke & Deterding, 2017).

Therefore, its important to guide nursing students in a self-regulated learning strategy with gamification inside. To this end, this study proposed an approach for integrating online game-based environments using Gather with gamification inside with a self-regulated learning strategy to investigate its impact on nursing students' Leopold's maneuvers skills. Gamification with self-regulated learning is integrated in this study. The self-regulated learning in this study included: goal setting, doing the task and having gamification inside; the students can get the points, rewards, and leaderboard after finishing the task, and the last step is self-evaluation. To investigate the potency of the proposed approach, an experiment will be conducted at a school of nursing in a Taiwan university. We will use a quasi-experimental design to answer the following research questions.

- 1.) How does gamification with a self-regulated approach improve the students' learning performance compared to gamification without an SRL approach to learning Leopold's maneuvers skill?
- 2.) How does gamification with a self-regulated approach improve the students' learning motivation compared to gamification without an SRL approach to learning Leopold's maneuvers skill?
- 3.) How does gamification with a self-regulated approach improve the students' self-efficacy compared to gamification without an SRL approach to learning Leopold's maneuvers skill?
- 4.) What is students' perception toward gamification with self-regulated learning?

#### 2. Literature Review

#### 2.1 *Gamification in Nursing Education*

The utilization of game-based learning and diverse forms of gamification is progressively garnering momentum in the domain of education (Vidergor, 2021). Gamification can enhance learning experiences by positively impacting various factors such as extrinsic motivation, enjoyment, feedback, collaborative work, and active learning (Ferriz-Valero et al., 2020). Several studies have explored the effectiveness of gamification in various health care domains, such as clinical skills training and teaching learning in medical education (Gentry, et al., 2019; Kim & Kim, 2022; Sardi et al., 2017). These studies have demonstrated positive effect, including increased knowledge, skills, professional attitudes, outcomes, and satisfaction attitude; was effective in helping students emphatize

with their patients and improve their problem-solving skills; gamification of e-Health frequently involves the utilization of rewards, feedback, and socialization elements. However, the use of gamification in the context of promoting nursing students' performance in Leopold's maneuvers remains limited.

For instance, a study by Koivisto, et al., (2016) implemented experiences of learning clinical reasoning processes by playing a 3D simulation game. The utilization of gamification has the potential to facilitate the development of clinical reasoning, problem identification, and the application of theoretical knowledge, as evidenced by the findings of this investigation. The utilization of 3D patient scenarios in clinical reasoning, which incorporate interactive patient and equipment features, can provide prompt feedback to healthcare professionals.

Another study by Brull et al., (2017) implemented a gamified learning platform to investigate the efficacy of three pedagogical approaches, namely didactic, online modules and gamification, in the context of wound management, pain management, and fall prevention during orientation. The platform integrates avatar-based gaming with educational content accompanied by knowledge-based challenges and the opportunity to earn badges. The results indicate that gamification is a viable instructional approach in contrast to conventional pedagogical techniques. The incorporation of gaming elements in the learning process was found to be positively received by staff and resulted in higher knowledge retention. The findings indicate that gamification effectively increases the users attention, elicits a sense of pleasure and foster the utilization of cognitive abilities. Therefore, to increases nursing students performance, motivation, self-efficacy and perceptions, researcher will conduct gamification elements during learning activity.

#### 2.2 Self-Regulated Learning in Nursing Education

Self-regulated learning (SRL) is a construct that emphasizes learners active engagement in planning, monitoring and evaluating their learning process (Zimmerman, 2002). The ability to engage in self-regulated learning (SRL) is an essential skill that physicians must possess in order to pursue lifelong learning and provide superior healthcare services (Bransen et al., 2022). According to (Yang et al., 2018), students may lose sight of their learning objectives and performance if they don't use proper learning tactics when playing games. Several studies have explored the effectiveness of self-regulated learning in various health care domains, including SRL in a competency-based, e-learning on patient safety (Zheng et al., 2020; Gaupp et al., 2018). These studies examined the employment of self-regulated learning (SRL) by students across their self-reported levels of academic achievement. The findings revealed that students at all levels utilized planning and reflection strategies.

Furthermore, the study by (Yoo & Jung, 2022) pointed out that self-regulated learning was observed to manifest when individuals were provided with chances to engage in self-regulated study and leadership activities, as well as when they established interpersonal connections with their professors and peers. These methods have been acknowledged to enhance the level of satisfaction in the process of learning. Based on the findings of this investigation, it is imperative to devise pedagogical approaches that augment the degree of contentment experienced by learners enrolled in remote nursing education curricula. However the use of self-regulated learning in the context of Leopold's maneuvers skills or other clinical skills remains limited. Therefore, to enhance students self-regulated learning by their own learning, this study will conducted its impact on nursing students clinical skills, especially for Leopold's maneuvers skills.

#### 3. Methodology

#### 3.1 Study design and participants

We will adopt a quasi-experimental design and use online game-based learning environments using Gather and gamification inside with a self-regulated learning strategy to know its impact on nursing students Leopold's maneuvers skills. First-year nursing students will be recruited as experiment participants, and the same teacher will teach all participants. We will randomly assign 25 participants to the experimental group to conduct gamification with a self-regulated learning strategy and 25 participants to the control group using traditional course learning. The experimental group use Gather as an online game-based learning environment.

## 3.2 Development gamification with self-regulated learning strategy through an online game-based environment using Gather

The adoption of digital technology in teaching and learning has garnered significant attention from educators, particularly those in higher education. The process of digitalized learning entails incorporating technological tools to transform educational materials into digital formats, which may include video recordings, audio files, and video conferencing (Yow, 2022). We used Gather to develop an online game-based environment as an interactive role-playing game to teach Leopold's maneuvers skills. The tool facilitates the creation of virtual spaces that accurately represent physical locations, with the added capability of integrating game settings such as tables, offices, and classrooms into virtual rooms. According to (Najjar et al., 2022), the tool facilitates the creation of virtual parts where users can position their avatars and communicate with other participants who are also present. They can also use video, audio, and chat to communicate.

By integrating gamification inside the Gather, we used a self-regulated learning strategy to regulate their limit throughout their learning interests. Figure 1. illustrates self-regulated learning with a gamification process through an online game-based environment using Gather. There are three phases the students must follow after they enter the learning game environment: forethought phase; goal setting, performance phase; doing the task, and self-reflection phase; self-evaluation. During the "forethought phase" the system prompts students to set learning objectives, including anticipated classification, anticipated correct-answer rate, and anticipated learning duration. During the "performance phase" students are instructed to observe and track their learning progress and rate of goal attainment while participating in the game. Also, upon completion of each phase of the game, the system provides feedback to students based on their game results. In the "self-reflection phase", students are guided to assess and reflect on their learning outcomes. They are then encouraged to modify their goal-setting approach for the next stage of the game. Before starting the game or learning activities, the teacher will introduce the system and learning material. To start the game, the students need to go in the direction in the first phase; they set their learning goals by embedding Google Forms into Gather. Then, before the students do the task, they can learn by watching a video about Leopold's maneuvers in Gather. After that, in the second phase, students do the task by following the directions in the game; while students do the task, they can see their points, badge, and leaderboard and other students' leaderboards. To end this game, in the third phase, students need to follow the direction to fill self-evaluation form in Gather to evaluate their learning according to their doing the task and goal-setting status.

The educational scenarios that were formulated within the game were derived from a nursing textbook and a formal nursing protocol of fundamental nursing proficiencies (Superville et al., 2023), especially in Leopold's maneuvers skills (such as indications, contraindications, preparations, technique, complications and clinical significance) which were reviewed for content validity by two experts in the field of nursing education. The storyline in this game is based on the development of Leopold's maneuvers skills. Students need to learn the process of four Leopold's maneuvers by completing the five-step learning task (i.e., indications, contraindications, preparation, procedure and complication). In the first step, students need to understand what the indications are and the reasons or criteria that make Leopold's maneuvers of action appropriate or necessary for a patient. In the second step, students need to understand what are the specific circumstances, conditions, or factors that make a particular medical treatment, procedure, or intervention inadvisable or potentially harmful for a patient, especially for Leopold's maneuvers intervention. In the third step, nursing students need to be familiar with the specific equipment, instruments, and supplies

required for a particular procedure. They must know how to operate the equipment, ensure its proper functioning, and be aware of any specific considerations or precautions. In the fourth step, that important step, students need to understand well about Leopold's maneuvers skills procedure and how to implement the intervention. In the last step, students need to understand who can take risks, complications unexpected or undesirable outcomes that may occur during or after a medical procedure. Complications can range from mild and temporary side effects to severe and potentially life-threatening events.



Figure 1. Self-regulated learning with gamification process

#### 3.3 Experimental procedure

Figure 2. illustrates the experimental procedure of this study. Before the learning activities, the teacher introduces Leopold's maneuvers skills syllabus and learning goals. All students will take the learning performance pre-test and complete the learning motivation and self-efficacy pre-questionnaire. The students will be divided into the experimental group and control group to learn about Leopold's maneuvers skills. More specifically, the experimental group will complete gamification with a self-regulated learning strategy, while the control group will complete this course using traditional classroom. The experimental group utilized gamification with self-regulated learning strategy, which is gamification involves incorporating gamification elements such as points, badge and leaderboard into the learning process to enhance learning outcomes and motivation. And self-regulated learning involves the students taking an active role in their learning process. They set their learning goals, watch instructional video, complete task and then evaluate their learning performance. While the control group follows a traditional approach to learning. Students in this group watch instructional video, practice and complete learning sheets. Furthermore, the learning activity will carry out once a week for 100 mins each time, for a total of 200 mins over two weeks. The purpose of this study is to investigate whether the gamification and self-regulated learning approach in the experimental group leads to better learning performance, increase motivation and improved self-regulated learning compare to traditional classroom approach in the control group. After the learning activities, all students will complete the learning performance post-test and the post-questionnaire of learning motivation and self-efficacy.



Figure 2. Experimental procedure

#### 3.4 Instruments

The Leopold's maneuvers test was created by two experienced nursing educators to assess students' pre- and post-knowledge of Leopold's maneuvers. Each test had 20 multiple-choice questions about Leopold's maneuvers, with a total score of 100.

The learning motivation scale was adopted by Wang and Chen (2010). It consisted of three items for intrinsic motivation such as "In this course, I prefer course material that really challenges me so I can learn new things.", "In this course, I prefer course material that arouses my curiosity, even if it is difficult to learn." and "When I have the opportunity, I choose to do the quizzes in this course that I can learn from, even if they don't guarantee a good grade." and three items for extrinsic motivation such as "Getting a good grade and mastering Leopold's maneuvers skills in this course are the most satisfying things for me.", "If I can, I want to get better grades in this course than most of the other students." and "I want to do well in this course because it is important to show my skills to my family, friends, employer, and others.". A 5-point Likert scale will be adopted in this measure, ranging from 1 (strongly disagree) to 5 (strongly agree).

The self-efficacy scale was adopted by Pintrich et al. (1991). It consisted of eight items, such as "I believe I will receive an excellent grade in this class.", "I'm certain I can understand the most difficult learning material presented in this game.", "I'm confident I can understand the basic concepts taught in this course.", "I'm confident I can understand the most complex material presented in this game.", "I'm confident I can understand the tasks in this game.", "I expect to do well in this course.", "I'm certain I can master Leopold's maneuvers skills being taught in this course." and "Considering the difficulty of this course, the teacher, and the Leopold's maneuvers skills, I think I will do well in this course.". A 5-point Likert scale will be adopted in this measure, ranging from 1 (strongly disagree) to 5 (strongly agree).

The online game-based environment using Gather.town was developed in this study. This is capable of running on individual personal computers. The students can choose their avatar and sign in to the game map and enable to open the video, audio and chat with other students. By embedding learning strategies into Gather and gamification inside the task. Furthermore, a set of questions was adopted by Chang et al. (2020) to interview to know the students' perceptions after using this system: 1.) What was the difference between this kind of instruction and the way you expected? 2.) As a whole, what were the advantages of this learning approach from your perspectives? 3.) In which part did you gain the most by using this approach? In which part did you learn the most? 4.) What should be improved with regard to the function or interface design of the system? 5.) Do you look forward to learning with this learning approach again? What kind of subject? Why? Why are these subjects suitable for this type of learning? 6.) Would you recommend that your classmates learn by using this system or this kind of method? Do you think they need to learn using this method? 7.) Would you recommend that your teachers teach using this system or this kind of method? Do you think they would like to teach using this method? The you think they need to teach in this way? Or, do you think they would like to teach using this method?

#### 4. Expected Results

The study aims to examine the impact of a gamified learning intervention with a self-regulated approach on nursing students' performance in conducting Leopold's maneuvers. It is anticipated that gamification and self-regulated learning strategies will lead to several positive outcomes. To begin, nursing students engaging in the intervention are required to exhibit enhanced skill in completing Leopold's maneuvers. Gamified learning and self-regulation approaches are expected to improve students' performance, motivation, and self-efficacy due to their interactive and engaging character. Second, the intervention is intended to improve information retention among nursing students since gamification and self-regulated learning promote deeper comprehension and memory consolidation. Third, implementing self-regulatory skills inside the gamified learning is expected to boost students' development of self-regulatory abilities such as goal-setting and self-evaluation. This may result in enhanced self-awareness, self-efficacy, and the capacity to recognize and address areas for improvement in performing Leopold's maneuvers. Finally, the intervention is projected to elicit positive student feedback and satisfaction since the engaging and immersive character of gamification, combined with the autonomy and control offered by self-regulated learning, is likely to promote a good learning experience.

#### 5. Conclusions

While gamification and self-regulated learning have each shown favorable impacts in nursing education, their combined influence on enhancing nursing students' performance in Leopold's maneuvers is unknown. As a result, future research should focus on designing gamified learning procedures that combine self-regulated learning principles in order to increase nursing students' knowledge of Leopold's maneuvers.

Nursing educators may improve the efficacy of teaching and learning approaches, thereby boosting the quality of care offered by future nurses, by exploiting the potential of gamification and self-regulated learning. Interactive simulations, virtual reality settings, or instructional games that involve students in practicing and implementing Leopold's maneuvers in a controlled and immersive environment are examples of gamified learning interventions.

Integrating self-regulated learning into these gamified interventions might help nursing students take responsibility of their learning experience even more. Students may improve their metacognitive abilities and gain a better comprehension of Leopold's skills by encouraging goal-setting and self-evaluation. The purpose of this study is to improve nursing students' motivation, self-efficacy, and perceptions of competence.

Furthermore, the use of gamified learning interventions can help to bridge the gap between theoretical knowledge and practical application. Students can obtain vital experience and confidence in completing Leopold's maneuvers by imitating real-world conditions and offering instant feedback. This method not only increases students' learning performance but also prepares them for the difficulties they may face in clinical settings. Finally, the combination of gamification and self-regulated learning has significant promise for increasing nursing education and nursing students' competency in completing Leopold's maneuvers. More study in this area might give useful insights into creating and implementing successful gamified learning interventions that increase skill development, motivation, and self-efficacy among nursing students. Finally, these improvements have the potential to improve the overall quality of care offered by future nurses.

#### References

- Anguas-Gracia, A., Subiron-Valera, A. B., Antón-Solanas, I., Rodriguez-Roca, B., Satustegui-Dorda, P. J., & Urcola-Pardo, F. (2021). An evaluation of undergraduate student nurses' gameful experience while playing an escape room game as part of a community health nursing course. *Nurse Education Today*, *103*, 104948.
- Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2017). Studying student differentiation in gamified education: A long-term study. *Computers in Human Behavior*, *71*, 550-585. https://doi.org/10.1016/j.chb.2016.08.049
- Biancuzzo, M. (1993). How to recognize and rotate an occiput posterior fetus. *The American Journal* of Nursing, 93(3), 38-41.
- Bozkurt, A., & Durak, G. (2018). A systematic review of gamification research: In pursuit of homo ludens. *International Journal of Game-Based Learning (IJGBL)*, 8(3), 15-33. https://doi.org/10.4018/IJGBL.2018070102
- Brandon, A. R., Pitts, S., Denton, W. H., & Stringer, E. H. M. (2009). A history of the theory of prenatal attachment. *Journal of Prenatal and Perinatal Psychology & Health*, 23(4), 201–222.
- Bransen, D., Driessen, E. W., Sluijsmans, D. M., & Govaerts, M. J. (2022). How medical students co-regulate their learning in clinical clerkships: a social network study. *BMC Medical Education*, *22*(1), 193.
- Brull, S., Finlayson, S., Kostelec, T., MacDonald, R., & Krenzischeck, D. (2017). Using gamification to improve productivity and increase knowledge retention during orientation. *The Journal of Nursing Administration*, 47(9), 448–453.
- Chang, C. Y., Chung, M. H., & Yang, J. C. (2022). Facilitating nursing students' skill training in distance education via online game-based learning with the watch-summarize-question approach during the COVID-19 pandemic: A quasi-experimental study. *Nurse Education Today*, 109, 105256. <u>https://doi.org/10.1016/j.nedt.2021.105256</u>
- Chang, C.-Y., Kao, C.-H., Hwang, G.-J., & Lin, F.-H. (2020). From experiencing to critical thinking: A contextual game-based learning approach to improving nursing students' performance in electrocardiogram training. *Educational Technology Research and Development*, 68(3), 1225-1245.
- Chang, C.-Y., Panjaburee, P., Lin, H.-C., Lai, C.-L., & Hwang, G.-H. (2022). Effects of online strategies on students' learning performance, self-efficacy, self-regulation and critical thinking in university online courses. *Educational Technology Research and Development*, *70*(1), 185-204
- Dastjerdi, N. B. (2016). Factors Affecting ICT Adoption among Distance Education Students Based on the Technology Acceptance Model - A Case Study at a Distance Education University in Iran. *International Education Studies*, 9(2), 73-80. <u>https://doi.org/10.5539/ies.v9n2p73</u>
- Gaupp, R., Fabry, G., & Körner, M. (2018). Self-regulated learning and critical reflection in an e-learning on patient safety for third-year medical students. *International Journal of Medical Education*, *9*, 189.
- Gentry, S. V., Gauthier, A., L'Estrade Ehrstrom, B., Wortley, D., Lilienthal, A., Tudor Car, L., ... & Car, J. (2019). Serious gaming and gamification education in health professions: systematic review. *Journal of medical Internet research*, *21*(3), e12994. <u>https://doi.org/10.2196/12994.</u>
- Hwang, G.-J., & Chang, C.-Y. (2020). Facilitating decision-making performances in nursing treatments: a contextual digital game-based flipped learning approach. *Interactive Learning Environments*, 1-16. <u>https://doi.org/10.1080/10494820.2020.1765391</u>
- Kim, H., & Kim, B. (2022). Effects of Situation-Based Flipped Learning and Gamification as Combined Methodologies in Psychiatric Nursing Education: A Quasi-Experimental Study. In *Healthcare* (Vol. 10, No. 4, p. 644). MDPI. <u>https://doi.org/10.3390/healthcare10040644</u>

- Koivisto, J. M., Multisilta, J., Niemi, H., Katajisto, J., & Eriksson, E. (2016). Learning by playing: A cross-sectional descriptive study of nursing students' experiences of learning clinical reasoning. *Nurse Education Today*, 45, 22–28. <u>https://doi.org/10.1016/i.nedt.2016.06.009</u> PMID:27429399
- McFarlin, B. L., Engstrom, J. L., Sampson, M. B., & Cattledge, F. (1985). Concurrent validity of Leopold's maneuvers in determining fetal presentation and position. *Journal of Nurse-Midwifery*, *30*(5), 280-284.
- Mirghani, H. M., Khair, H., & Ahmed, B. (2007). LEOPOLD'S MANEUVERS: HOW COMMON IS IT IN OUR DAILY PRACTICE? *Gynaecologia et Perinatologia: Journal for Gynaecology, Perinatology, Reproductive Medicine and Ultrasonic Diagnostics*, *16*(2), 79-82.
- Nacke, L. E., & Deterding, S. (2017). The maturing of gamification research. *Computers in Human Behavior*, 71, 450-454. <u>https://doi.org/10.1016/j.chb.2016.11.062</u>
- Najjar, N., Stubler, A., Ramaprasad, H., Lipford, H., & Wilson, D. (2022). Evaluating Students' Perceptions of Online Learning with 2-D Virtual Spaces. Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 1,
- Nishikawa, M., & Sakakibara, H. (2013). Effect of nursing intervention program using abdominal palpation of Leopold's maneuvers on maternal-fetal attachment. *Reproductive Health*, *10*(1), 1-7.
- Sardi, L., Idri, A., & Fernández-Alemán, J. L. (2017). A systematic review of gamification in e-Health. *Journal of biomedical informatics*, *71*, 31-48. <u>https://doi.org/10.1016/j.jbi.2017.05.011</u>
- Superville SS, Siccardi MA. Leopold Maneuvers. [Updated 2023 Feb 19]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK560814/
- Vidergor, H. E. (2021). Effects of digital escape room on gameful experience, collaboration, and motivation of elementary school students. *Computers & Education*, *166*, 104156. <u>https://doi.org/10.1016/j.compedu.2021.104156</u>
- Yang, J. C., Chung, C. J., & Chen, M. S. (2021). Effects of performance goal orientations on learning performance and in-game performance in digital game based-learning. *Journal of Computer Assisted Learning*, 38(2), 422-439. <u>https://doi.org/10.1111/jcal.12622</u>
- Yang, J. C., Lin, M. Y. D., & Chen, S. Y. (2018). Effects of anxiety levels on learning performance and gaming performance in digital game-based learning. *Journal of Computer Assisted Learning*, 34(3), 324–334. <u>https://doi.org/10.1111/jcal.12245</u>
- Yoo, L., & Jung, D. (2022). Teaching presence, self-regulated learning and learning satisfaction on distance learning for students in a nursing education program. *International Journal of Environmental Research and Public Health*, 19(7), 4160.
- Yow, H.-Y. (2022). A Case Study of Virtual Anatomy Museum: Facilitating Student Engagement and Self-paced Learning through an Interactive Platform. *International Journal of Information and Education Technology*, *12*(12).
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64-70.

## Development and Usability Evaluation of an Intelligent Personalized Erhu Pitch and Rhythm Learning System

**Gwo-Haur HWANG**<sup>a\*</sup>, **Ping-Tsung TSAI**<sup>b</sup>, **Jenn-Kaie LAIN**<sup>b</sup> & **Shiuan-Han HUANG**<sup>c</sup> <sup>a</sup>Graduate School of Intelligent Data Science, National Yunlin University of Science and

Technology, Taiwan <sup>b</sup>Department of Electronic Engineering, National Yunlin University of Science and Technology, Taiwan <sup>c</sup>Chinese Music Department, Tainan National University of the Arts, Taiwan \*ghhwang0424@gmail.com

Abstract: Pitch accuracy and rhythm errors are the most common problems encountered by erhu beginners. The traditional way of learning the erhu involves using a tuner and a metronome to improve these two issues. However, for erhu beginners, without the assistance of a teacher, it is often difficult to address these problems effectively. Therefore, this study has developed an "Intelligent Personalized Erhu Pitch and Rhythm Learning System" to help erhu learners learn to pitch and rhythm anytime and anywhere without the presence of a teacher while providing professional feedback. This system included pitch and rhythm training features. In addition to automatically generating sheet music in the MML (Music Macro Language) format, it allows the erhu learners to play along with the sheet music. After each performance, the system provides learners with expert-like comments on pitch and rhythm, respectively. The system also keeps a record of learners' practice sessions and provides appropriate encouragement and guidance based on their learning progress by creating an intelligent and personalized learning scaffold. After the development of the system, this study conducted a usability evaluation and investigated the relationship between usability and erhu playing experience. The results showed that the average for all usability dimensions was above 4.6, indicating that the system is acceptable for erhu learners. Moreover, learners with more erhu playing experience showed higher acceptance of the system compared to those with less experience in erhu.

**Keywords:** Intelligent, personalized, erhu, pitch learning, rhythm learning, usability evaluation

#### 1. Introduction

#### 1.1 Background

Erhu is popular among the Chinese community because of its cultural identity and expressiveness. Erhu lessons can be conducted in both one-on-one and group classes. Although the quality and progress of group classes may not be as good as one-on-one teaching lessons, most amateur learners prefer group classes because of the low price, the ease of making friends, and the relaxed learning environment.

#### 1.2 Motivation and Purpose

Erhu is not easy to learn because it needs proficient skills to control the tone color, pitch, and rhythm. Beginners will need to rely on a teacher for a long time to master those skills. However, it takes more time for a teacher to demonstrate and explain in group classes, which makes it impossible to correct every problem with rhythm or pitch (Hwang, Chen, Sung, & Kuo, 2015).

In traditional erhu group teaching, students practice pitch and rhythm with the teacher and using a tuner and metronome to correct pitch and rhythm. Learners are often unable to remember the mistakes they have made. The metronome stabilizes the rhythm, but it is less likely to be used as a tool for learning rhythm. Therefore, this study has developed an "Intelligent Personalized Erhu Pitch and Rhythm Learning System" to address the above problems of erhu learning to improve erhu learners' pitch and rhythm.

#### 2. Literature Review

#### 2.1 Technology-Assisted Pitch Learning

The sound of a bowed string instrument is controlled by pressing the strings, which is difficult for beginners because they need to control fingering, bowing posture, and develop pitch listening skills. Therefore, Wang, Wang, Chen, Chang, and Chen (2012, July) developed a violin pitch detection system that provides scales and arpeggio scores as teaching materials. Violin learners can choose different rhythms to practice according to their level. Unlike the violin, the erhu does not have a fingerboard to press, and all pitches need to be pressed with the left hand floating in the air.

Previous research has developed two erhu learning systems to improve finger position, pitch (Kikukawa, Ishihara, & Soga, 2013), and bowing (Kikukawa & Soga, 2014) using magnetic position sensors and display interface. Although these systems for erhu beginners are accompanied by a visualized interface, the device is still too complicated for those who need to practice constantly. Each time they wear the device, they may become less willing to learn. Therefore, there is a need for a more convenient erhu learning system that allows learners to learn anytime and anywhere without wearing any equipment, as long as they have a tablet.

#### 2.2 Technology-Assisted Rhythm Learning

Another important element of playing an instrument is rhythm. It is very important to have a correct and stable rhythm when playing music. However, beginners are often influenced by their own senses, resulting in a fast or slow rhythm that affects their overall performance. Therefore, learners usually use a metronome to stabilize their instruments.

Many studies have used body postures and beats to improve learners' rhythms, either indirectly or directly. For example, Fonteles and Rodrigues (2021) proposed a Kinect-based conducting system that allows students to imitate a conductor and thus train rhythm stability. At the same time, the system proved to be effective, simple and flexible. As another example, Dos Santos, Loke, Yacef and Martinez-Maldonado (2022) proposed a method for modeling rhythmic movements through smartphones. The smartphone's motion sensor enhanced the dance teachers' assessment of rhythm-related dance skills. These two studies were conducted to generate and evaluate music-related motion features and to enhance rhythm evaluation through dance teacher examination.

Although these systems are helpful for learners' rhythm, they are not suitable for erhu learners because the erhu requires the coordination of left and right hands. Therefore, we developed the "Intelligent Personalized Erhu Pitch and Rhythm Learning System" to help learners learn how to play the erhu.

#### 2.3 Intelligent Personalization Systems

In order to achieve personalized teaching and learning, an assisted learning system should not only provide learning resources and a learning environment for learners, but also actively interact with them to understand their learning situation and imitate the teaching style of real teachers to provide different teaching materials for different users. For example, Desmarais and Baker (2012) suggested that intelligent learning systems should embed learner models that effectively deal with uncertainty and partial evidence. This helps students know more about their learning progress. In addition, Zhe (2021) designed an Intelligent Guided Learning System (IGLS) model for music sight-reading, which not only combines the characteristics of the music sight-singing subject in teaching contents and teaching methods but also proposes a difficulty feature-based score recommendation algorithm and the sight-singing scoring algorithm.

Therefore, the system developed in this study incorporated the concept of intelligent personalized learning, allowing learners to learn on their own at any time and from anywhere, as well as incorporating comments from real erhu teachers and recording their learning history to provide a personalized learning guide.

#### 3. System Design

#### 3.1 System Architecture

The system architecture diagram is shown in the left of Figure 1. After entering the system, the learners can first tune and then choose to practice the pitch or rhythm. The system will generate the corresponding sheet music for the learners. It will analyze and provide a score and suggestions for each pitch after the learner completes their performance according to the instructions. The results will be recorded for learners to check at any time.

#### 3.2 Tuning Mode

The system has a built-in tuner for learners to tune before practicing. The screenshot of the built-in tuner is shown in the right of Figure 1. Before using the tuner, learners can choose to adjust the A4 frequency. Although the international standard for A4 is 440 Hz, some performers or orchestras will adjust the A4 frequency from 436 Hz to 444 Hz in order to achieve the best performance. When the learner adjusts the A4 frequency of the tuner of this system, the system will update the A4 frequency of the practice mode at the same time.



Figure 1. System Architecture Diagram and Screenshot of the Built-in Tuner.

#### 3.3 Pitch Training Diagnostic Mode

Since a semitone is equal to 100 cents, a reasonable range for a pitch is plus or minus 50 cents, i.e., a frequency exceeding plus or minus 50 cents is considered out of tune. The human ear can distinguish at least 25 cents (Peretz & Hyde, 2003), so when comparing, the exact pitch can be set at about 20–25 cents, and when the detected frequency falls within plus or minus 20 cents after the cents calculation, it can be judged as the correct pitch. The detailed phonetic score judging criteria are shown in Table 1.

Judgment	Judgment Criteria
Correct	cent <20
Too high	20< cent <50 & cent>0
Too low	20< cent <50 & cent<0
Out of tune	cent ≥50
Not played	f = 0

 Table 1. Tone Division Judgment Standard

The pitch training mode, as shown on the left of Figure 2, allows learners to choose the level they want to practice at. There are descriptions of the topic below each level's button so that learners can choose the appropriate level according to their ability.

The sheet music produced by the pitch training mode includes scale exercises based on the commonly used keys of the erhu (D, G, C, F, A, <sup>b</sup>B) and three positions (two octaves). The note will be marked during the performance to guide the learner. The correct frequency waveform (black) and detected frequency waveform (red) will be displayed on the top of the interface so that the learners can instantly check their performance status, as shown in the right of Figure 2.



Figure 2. Screenshot of Level Selection and During Performance.

After playing, the system will analyze the learner's performance and display the detection results for each note below the corresponding note, as shown in Figure 3. There are four judgment levels, each represented by a color, and each judgment level is scored according to the judgment level. In the performance results, learners can click on a note on the score to enter the individual diagnosis interface. The individual diagnosis interface of the pitch mode (Figure 4) shows information about the performance judgment (correct, high, low, severely high, and severely low), the judgment problem, and suggestions for correction.



Figure 3. Screenshot of the End of Performance.



Figure 4. Individual Diagnostic Screenshot of the Pitch Training Diagnostic Mode.

#### 3.4 Rhythm Training Diagnostic Mode

The criteria for detecting rhythm is to compare the detected start time, end time and duration with the correct start time, end time and duration, respectively. The threshold value for both start and end times is TH, and the threshold value for the duration is 20% of the correct time. There are four judgments for start time, end time, and duration. The four judgments for start time are "not played", "beat rush", "beat drop", and "correct"; the four judgments for end time are "not played", "beat rush", "beat drag", and "correct"; and the four judgments for duration are "not played", "too short", "too long", and "correct". The criteria for calculating the rhythm score are shown in Table 2.

Status	Judgment criteria Score		Weight	
Ditab	Exact string	1	1	
FIIGH	Wrong string	0	I	
Start time	Time Difference < TH	1	1	
Start time	Time Difference > TH	0	I	
End time	Time Difference< TH	1	1	
End time	Time Difference > TH	Time Difference > TH 0		
Duration	Time Difference < 20% 1		1	
Duration	Time Difference > 20%	0	I	
	Coverage rate >80%	2		
Rhythm coverage rate	50%< Coverage rate <80% 1		4	
	Coverage rate <50%			
The number of points obtained = Score x Weight				
Rhythm score = (Total number of points obtained / 12) x 100				

Table 2. Rhythm Score Calculation Criteria

The rhythm training diagnostic mode generates sheet music in different rhythms, including the common erhu rhythms (4 beats, 3 beats, 2 beats, 1 beat, half beat, etc.), and the rhythm practice is in D major. The notes are composed of the inner string and the outer string played together. The individual diagnostic interface of the Rhythm mode (Figure 5) shows information about the correctness of the empty strings (inner and outer strings), the error time of the entry point, the error time of the breakpoint, the error of the playing length, the rhythm coverage, and the rhythm score, with the criteria shown in Table 3.



Figure 5. Rhythm Mode Individual Diagnosis Interface.

Table 3. Judgment Levels and	Scoring	Criteria
------------------------------	---------	----------

Color	Sound level mode	Rhythm mode	Score	Weight
Blue	Cent  ≤ 20	Rhythm Score ≥ 90	1	3
Green	20 <  Cent  ≤ 50	60 ≤ Rhythm Score < 90	0.5	2
Orange	Cent >50	0 < Rhythm Score < 60	0	1
red	Not played	Rhythm Score = 0	0	0

#### 3.5 Intelligent Personalized History Analysis and Guidance Model

After each performance, the system will record the results of the performance. In the history analysis interface, each record will show the serial number, user name, selected practice mode, level, score, and performance date. The history analysis interface is shown in Figure 6.

After playing, the learner can click on the button in the upper right corner for history analysis to view past performances and compare the accuracy of each month's performance.



Figure 6. History Analysis.

The system will analyze the learner's learning history by month, mainly divided into three analysis modes: "Analysis of this month", "Comparison of specified months" and "Trend of performance accuracy". Learners can click on the upper-right corner button to check their past record in the History Analysis and use the drop-down menus to select the items they want to view:

Analysis of this month: The analysis for this month is shown on the left of Figure 7. The system first collects all the records of the learner's performance at the specified level in this month and then calculates the accuracy (percentage) of each item in the level (pitch training mode is the pitch that appears in the track, rhythm training mode is the beat that appears in the track) according to the weights in Table 3. The system calculates the sum of the weights of each item in the month and the total number of items and then divides the sum and the total number by 3 (the weight is 3) to calculate the accuracy of the item in the month. At last, the accuracy of every item will be presented in a bar chart.

Specified month comparison: The schematic diagram of the specified month comparison is shown on the right of Figure 7. The calculation method is the same as the "analysis of this month". Learners can compare whether they have progressed or regressed in playing the item this month with previous months.



Figure 7. Analysis of the Month (Rhythm Mode) and Comparison of Specified Months.

Performance Accuracy Trend: The performance accuracy trend graph is shown in Figure 8. This mode analyzes the average accuracy of the learner for each month at the specified level. The system lists the months in which the learner played the specified level from the oldest to the newest, calculates the average accuracy for each month, and then presents the average accuracy trend for the specified level as a line graph.



Figure 8. Performance Accuracy Trend.

#### 4. Experimental Results and Analysis

#### 4.1 System Availability Evaluation

After the system was developed, in order to investigate whether the interface and operation of the system are suitable for erhu learners, erhu students and teachers were invited to try out and operate the "Intelligent Personalized Erhu Pitch and Rhythm Learning System" developed by this study. They need to fill out the usability evaluation questionnaire, and an interview will be conducted to find out their thoughts after operating the system.

The usability evaluation questionnaire used in this study was planned and designed according to the 10 user interface usability evaluation criteria of the system proposed by Nielsen (1994) and Nielsen (1994, April) and was designed using Likert (1932) five-point scales. Each question has 5 options: "Strongly agree (5 points)", "Agree (4 points)", " Neither agree nor disagree (3 points)", "Disagree (2 points)" and "Strongly disagree (1 point) ". Each dimension has three questions according to the system design, and the entire questionnaire has a total of 30 questions.

#### 4.2 Experimental Subjects

The subjects of this study were 34 erhu learners from a music classroom in central Taiwan and erhu majors from the Department of Chinese Music at a university of the art in southern Taiwan. The age distribution ranges from 9 to 74 years old, from 1 to 10 hours of erhu practice per week, and from less than 1 year to more than 10 years of erhu playing experience.

#### 4.3 Experimental Process

The experimental procedure for system usability evaluation in this study is shown in Figure 9. Before the experiment starts, the experimental subjects will be informed of the experimental precautions and asked to fill out the informed consent form for 10 minutes. After that, the experimental subjects were introduced to the functions of the system and explained the system's operation procedure for about 10 minutes. Each experimental subjects will be allowed to use the tuner of the system to tune the erhu for about 5 minutes after the explanation of the system operation. Then, the experimental subjects will be allowed to practice pitch and rhythm, respectively. They can check the sheet music for each mode and choose two of them to practice for about 15 minutes. After all the tracks are practiced, the experimental subjects will be allowed to view the performance records and analysis of the performance for about 5 minutes. Lastly, a 15-minute usability evaluation questionnaire and a post-use interview were conducted to find out what they thought of the system. The experiments are shown in Figure 10.





Figure 10. Usability Evaluation Experiments.

#### 4.4 Data Collection and Analysis

After all the subjects had completed the experiment and filled out the questionnaires, all the data were analyzed using SPSS statistical software. First, a reliability analysis was carried out. The results showed that the Cronbach's Alpha values of each facet were greater than 0.7 (between 0.755 and 0.906), indicating that all questionnaires reached an acceptable level.

After analyzing the reliability of the questionnaire, we then analyze the average of each dimension. The usability evaluation results can be seen in Table 4. The average of all dimensions is greater than 4.6. The results show that the system developed by this research is suitable for erhu learners. Among them, the top three with higher averages are H4, H7 and H5.

Usability evaluation dimensions	Count	Min	Max	Ave	SD	Ranking
H1 : Visibility of system status	34	3.0	5.0	4.690	0.478	5
H2 : Match between system and the real world	34	3.0	5.0	4.618	0.520	10
H3 : User control and freedom	34	3.0	5.0	4.710	0.448	4
H4 : Consistency and standards	34	4.0	5.0	4.770	0.374	1
H5 : Error prevention	34	3.0	5.0	4.740	0.433	3
H6 : Recognition rather than recall	34	3.0	5.0	4.670	0.519	8
H7 : Flexibility and efficiency of use	34	3.0	5.0	4.750	0.421	2
H8:Aesthetic and minimalist design	34	3.0	5.0	4.680	0.522	6
H9 : Help users recognize, diagnose, and recover from errors	34	4.0	5.0	4.670	0.512	9
H10 : Help and documentation	34	3.0	5.0	4.680	0.522	7
Average of the dimensions	34	3.4	5.0	4.696	0.437	

Table 4. Result of Usability Evaluation

In Table 5, erhu learners with more erhu playing experience had significantly higher averages in H6, H7, H8, and H9 than erhu learners with less experience, which means they have higher acceptance of the system.

Usability evaluation dimensions	Erhu playing experience	Count	Ave	SD	t
H1 : Visibility of system status	>=8 years	16	4.792	0.382	- 1.221
	<=7 years	18	4.593	0.543	
H2 : Match between system and the real world	>=8 years	16	4.688	0.412	- 0.734
	<=7 years	18	4.556	0.605	
H3 : User control and freedom	>=8 years	16	4.854	0.344	- 1.891
	<=7 years	18	4.574	0.496	
H4 : Consistency and standards	>=8 years	16	4.854	0.344	- 1.179
	<=7 years	18	4.704	0.394	
H5 : Error prevention	>=8 years	16	4.875	0.295	- 1.894
	<=7 years	18	4.611	0.502	
H6 : Recognition rather than recall	>=8 years	16	4.875	0.342	- 2.425*
	<=7 years	18	4.481	0.585	
H7 : Flexibility and efficiency of use	>=8 years	16	4.917	0.258	- 2.315*
	<=7 years	18	4.611	0.488	
H8 : Aesthetic and minimalist design	>=8 years	16	4.896	0.264	- 2.590*
	<=7 years	18	4.481	0.618	
H9 : Help users recognize, diagnose, and	>=8 years	16	4.875	0.363	— 2.450*
recover from errors	<=7 years	18	4.481	0.563	
H10 : Help and documentation	>=8 years	16	4.854	0.344	- 2.010
	<=7 years	18	4.519	0.607	
Average of the dimensions	>=8 years	16	4.848	0.307	- 2.048*
	<=7 years	18	4.561	0.497	

 Table 5. Impact of Erhu Playing Experience on Usability Evaluation

\*p<.05

#### 5. Conclusion and Future Prospects

#### 5.1 Conclusion

In this study, an "Intelligent Personalized Erhu Pitch and Rhythm Learning System" was developed using Unity. We evaluated the usability of the system and investigated the impact
of erhu playing experience on the usability evaluation. According to the experimental result, the average of each usability dimension is higher than 4.6, which means the system is acceptable for erhu learners. Learners who have more erhu playing experience have higher acceptance of the system. In other words, the more experienced erhu learners will be more receptive to this system than the less experienced erhu learners.

#### 5.2 Future Works

After the usability evaluation experiment, post-experimental interviews were conducted with all subjects in order to find out what they really thought about the system developed in this study. In terms of advantages, we hope to add some functions such as tempo customization, voice control, etc. In terms of system planning, we hope to diversify the topics or allow users to decide their own sheet music so that the practice will not be limited by the system. In the future, we will continue to modify the system according to the suggestions above and improve the stability of the system. We expect to conduct more complete teaching experiments on the learning effectiveness of erhu learners to investigate whether the system can improve their accuracy and stability of pitch and rhythm. In addition, in the future, the influence of learners of different ages on the acceptance of this system and the learning effect can also be explored.

## Acknowledgements

This research was funded by the National Science and Technology Council, Taiwan, R.O.C. with grant number NSTC 111-2410-H-224-004-MY2.

#### References

- Desmarais, M. C., & Baker, R. S. D. (2012). A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction*, 22, 9-38.
- Dos Santos, A. D. P., Loke, L., Yacef, K., & Martinez-Maldonado, R. (2022). Enriching teachers' assessments of rhythmic Forró dance skills by modelling motion sensor data. *International Journal of Human-Computer Studies*, *161*, 102776.
- Fonteles, J. H., & Rodrigues, M. A. F. (2021). User experience in a kinect-based conducting system for visualization of musical structure. *Entertainment Computing*, *37*, 100399.
- Hwang, G. H., Chen, B., Sung, C. W., & Kuo, T. H. (2015, November). To explore the effect of age on cognitive load when using social websites to conduct flipped classrooms on musical instrument performance teaching-Taking amateur erhu learners as an example. In *Proceedings of the 23rd International Conference on Computers in Education (ICCE 2015)*, Hangzhou, China (pp. 1-8).
- Kikukawa, F., Ishihara, S., Soga, M., & Taki, H. (2013, May). Development of a learning environment for playing erhu by diagnosis and advice regarding finger position on strings. In *NIME* (Vol. 2013, pp. 271-276).
- Kikukawa, F., Soga, M., & Taki, H. (2014). Development of a gesture learning environment for novices' erhu bow strokes. *Procedia Computer Science*, *35*, 1323-1332.
- Likert, R. (1932). A technique for the measurement of attitudes. Archives of Psychology, 22(140), 1-55.
- Nielsen, J. (1994). Heuristic Evaluation. Usability Inspection Methods. New York: John.
- Nielsen, J. (1994, April). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 152-158).
- Peretz, I., & Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Sciences*, 7(8), 362-367.
- Wang, J. H., Wang, S. A., Chen, W. C., Chang, K. N., & Chen, H. Y. (2012, July). Real-time pitch training system for violin learners. In 2012 IEEE International Conference on Multimedia and Expo Workshops (pp. 163-168). IEEE.
- Zhe, T. (2021). Research on the model of music sight-singing guidance system based on artificial intelligence. *Complexity*, 2021, 1-11. https://doi.org/10.1155/2021/3332216.

# Impact of Gender on Students' Classroom Engagement, Flow Experience and Learning Outcomes When Game-Based Answering Activities Are Integrated into the Curriculum

#### Gwo-Haur HWANG<sup>a</sup> & Yu-Cheng LIN<sup>b\*</sup>

 <sup>a</sup>Graduate School of Intelligent Data Science, National Yunlin University of Science and Technology, Taiwan
 <sup>b</sup>Center of General Education, National Taichung University of Science and Technology,

Taiwan

\*Linyuch8@gmail.com

Abstract: Many studies pointed out that when game-based answering activities are integrated into lecture teaching courses, students' classroom participation, flow experience, and learning effectiveness can be improved. However, some studies have also pointed out that gender differences could affect students' classroom participation. flow experience, and learning effectiveness when learning in physical classrooms or on the Internet. This study explores the impact of gender differences on students' classroom participation, flow experience, and learning effectiveness when the Kahoot game-based answering platform is integrated into lecture teaching courses. The subjects of the experiment were 31 students in the 12th grade of a senior high school in central Taiwan who were taking the "Arts and Life-applied Music" course. They are between 17 and 18 years old. Among them, there are 15 boys and 16 girls. The results showed that girls' classroom participation and flow experience were significantly better than boys'. However, in terms of learning outcomes (gain scores), girls are slightly better than boys, but not at a significant level. This study speculates that it may be because the theme of "Arts and Life-applied Music" is biased towards the field of music art, and girls are more interested in knowledge in this field. In addition, in the Kahoot game, because girls have better language skills, even if they use mobile phones to answer in class, the girls had a lot of speeches and discussions. During these processes, it is easier to strengthen the learning impression invisibly. On the other hand, boys tended to think in silence and reduced their interaction in language communication, and rushed to answer. This contrasts with the rational classroom atmosphere of boys and girls and boys tend to be more rational.

**Keywords:** Gender, game-based answering activities, classroom participation, flow experience, learning outcomes

## 1. Introduction

Although traditional lecture teaching is very efficient, it is often criticized for reducing students' classroom engagement and flow experience, which in turn affects learning outcomes. Many scholars have explored the factors that affect the classroom engagement of university students. The results showed that the factors that affect students' classroom engagement include teachers' motivation, encouragement or teasing among students, and interactions among students (Aziz & Kazi, 2019; Ghalley & Rai, 2019). Teachers asking questions to students in the classroom can increase interaction and thus increase engagement in the classroom. However, Mundelsee and Jurkowski's (2021) study indicated that students are often less willing to respond to teacher questions due to a lack of confidence, which in turn

affects classroom engagement. Therefore, it is important for students to be willing to participate actively when teachers ask questions in the classroom.

In addition to increasing student engagement in the classroom, immersing students in the instructional environment can have a significant impact on student learning outcomes. Clarke and Haworth's (1994) findings suggested that optimal mindfulness experiences are characterized by a high degree of cognitive engagement. Kiili, De Freitas, Arnab and Lainema (2012) found that well-designed game-based learning is effective in engaging students. The principles of good educational game design should include pedagogical elements that emphasize feedback and process principles.

Kahoot is a game-based scramble activity system suitable for integration into classroom lectures (Wang, Øfsdahl, & Mørch-Storstein, 2007, November). It is designed to provide highly interactive lecture robocall games using the classroom's existing presentation equipment and infrastructure (e.g., teacher's laptop, large screen, monopod projector, Internet connection, and students' cell phones). Dellos (2015) created quizzes on Kahoot with pictures and music to promote thinking for elementary school students. It enables students to use Kahoot to answer questions according totheir understanding without talking out loud in the classroom. This is beneficial to both teachers and students. This improves the disincentive for students to answer teacher questions individually as mentioned previously. Zhang and Yu (2021) analyzed 26 Kahoot-related journal articles to examine the impact of Kahoot on student learning outcomes, classroom interactions, and collaboration. The results showed that Kahoot can improve students' learning outcomes, interactions between students and teachers, and extracurricular collaborations among students in both traditional and flipped classroom learning environments if it is accompanied by good curriculum planning. From the above literature, Kahoot is a practical and effective tool for classroom interaction.

However, boys and girls in middle school begin to show significant differences in brain personality (Voyer & Voyer, 2014). Boys and girls have different brain and physiological development, and many educators believe that today's school systems: rigid classroom schedules, predominantly verbal classroom practices, less space and free time, and the need to work on multiple tasks at the same time prevent boys from being effective learners and potential troublemakers (Gurian, 2001). For example, the researcher observed that boys would often raise their hands during lecture time to leave the classroom and go to the restroom, which may be due to boys' lack of focus or comprehension of narrative knowledge. There have been many studies discussing the importance of gender differences in learning (Hyde, 2014). Failure to understand the importance of gender differences in academic development can be extremely risky (Green, 1993). If teachers are not attentive to the differences in learning between boys and girls, adjusting teaching methods and strategies for gender differences can result in compromising the effectiveness of well-designed instruction (Lesperance, Hofer, Retelsdorf, & Holzberger, 2022). Or, strategies and activities that are appropriate for girls may not also be appropriate to apply to teaching boys' classes (Cengiz Yakut, 2017).

To summarize, Kahoot is a game-based answering activity system that can increase students' classroom engagement, flow experience, and learning outcomes in traditional lectures. However, whether it has the same effect on male or female students has not been explored in the literature. Therefore, this study is a worthwhile research topic. The research questions of this study are specifically described as follows:

- (1) Does gender have a significant impact on students' classroom engagement when gamebased answering activities are integrated into the curriculum?
- (2) Does gender have a significant impact on students' flow experience when game-based answering activities are integrated into the curriculum?
- (3) Does gender have a significant impact on students' learning outcomes when game-based answering activities are integrated into the curriculum?

## 2. Literature Review

2.1 Impact of Game-Based Answering on Learning

#### 2.1.1 Advantages of Game-based Answering

Aljezawi and Albashtawy (2015) compared the effects of question-and-answer game-based instruction and lecture-based instruction in a physical course on student performance, satisfaction, and knowledge retention. The study utilized a true experimental research method and involved a total of 66 students. The results showed that there was no significant difference in the performance scores of the instant test between the two teaching methods. However, in the immediate achievement post-test and the retention test, students in the question-and-answer format scored significantly higher than students in the lecture format. Satisfaction surveys showed that the answering format was preferred and recognized by students as a more satisfactory teaching method.

With the popularity of unlimited internet in classrooms, Kahoot is increasingly being used in physical courses for game-based answering activities. Licorish, Owen, Daniel and George (2018) explored the impact of Kahoot on students' classroom dynamics as well as their learning experience. The study conducted semi-structured interviews with 14 students. The results of the study showed that Kahoot enriched the quality of students' learning in the classroom and had the greatest impact on improvements in classroom engagement, motivation, and learning experience. Their findings also suggest that the use of educational games in the classroom may minimize disruption and thus improve the guality of teaching beyond the traditional classroom. Uzunboylu, Galimova, Kurbanov, Belyalova, Deberdeeva and Timofeeva (2020) also explored teachers' use of "Kahoot" as a teaching aid. The subjects of the study were 38 teachers, including 17 males and 21 females. In the study, semistructured interviews were used as the data collection tool. The results of the study showed that the teachers perceived the advantages of Kahoot as helping to eliminate deficiencies in learning and increasing students' engagement, motivation, and activity in the classroom. Rajabpour (2021) explored the advantages and disadvantages of using Kahoot in the university classroom. Focus group interviews were used. Interviews were conducted with 13 faculty members working at the Center for Matriculation Studies at a university in Oman. Seven of them were male and six were female. The strengths of Kahoot as perceived by these teachers were: increased student participation, motivation, and activity in the classroom.

In summary, game-based answering activities, especially Kahoot, can indeed increase overall student engagement, motivation, activity, learning motivation, and learning experience in the classroom.

#### 2.1.2 Limitations of Game-based Answering

Although game-based answering activities, especially Kahoot, have many advantages, there are some problems and limitations that exist within it in today's teaching and learning environments. Wang and Tahir (2020) conducted a literature review on the effects of using Kahoot, which included a total of 93 relevant studies. The impact of Kahoot on learning performance and classroom dynamics was examined from the perspectives of students and teachers. The main difficulties cited by students were unreliable Internet connection and time pressure to answer questions, while teachers cited difficulties in mastering the difficulty of questions and unreliable Internet connection. Rajabpour (2021) also suggested that using Kahoot in the classroom may have problems such as network congestion, the negative influence of students on ranking, and difficulty in maintaining freshness. In addition, Uzunboylu, Galimova, Kurbanov, Belyalova, Deberdeeva and Timofeeva (2020) also mentioned that Kahoot is an Internet-based game, so using the program may be problematic for students who have little knowledge of the Internet domain. Also, gender differences may have an impact on Internet learning activities. For example, Chen and Tsai (2007) explored gender differences in attitudes toward Internet learning among university students in Taiwan. The study collected responses from 1,866 Taiwanese university students (940 males and 926 females) between the ages of 18 and 23 through a web-based survey. While females may perceive the Internet as a technology that is more favorable for male use, the analysis showed that females had higher performance outcomes in e-learning than males. In addition, females

held more positive attitudes toward the usefulness and diversity of e-learning content than males.

#### 2.2 Impact of Gender Differences on Learning

In the past, many scholars have mentioned that gender differences have an impact on students' learning.

Lesperance, Hofer, Retelsdorf and Holzberger (2022) explored gender differences in learning motivation. The study synthesized 71 stressors from 20 preliminary studies. The results of the analysis indicated that male and female students would differ in motivation. interest, self-confidence, stereotypes, and affective factors and that making adjustments in teaching styles based on these factors would have a positive impact on both male and female students. In addition, Hyde (2014) also explored gender similarities and differences. The study summarized past research on gender similarity and difference theories. The results of the analysis showed that females and males performed similarly in mathematics; males outperformed females in timing tests and spatial performance; females were more neurotic and gentler than males in temperament; males were more irritable in emotional experience. while females were more prone to sadness, fear, and happiness; females were more prone to guilt and shame when expectations were placed on them, whereas males were more confident. From the above two studies, we can see that there are differences in the characteristics of different genders, and the teaching methods used by teachers will also produce differences in the effectiveness of teaching, so the teaching methods can be adjusted by the above analysis. In addition, Voyer and Voyer (2014) explored gender differences in academic achievement. The study included an assessment of gender differences in student achievement at the elementary, middle, high school, or college level (undergraduate and graduate). A final analysis was conducted on 502 measures drawn from a sample of 369 respondents. The results of the study indicate that there is a stable female advantage in school achievement. Notably, females had the greatest advantage in language courses and the least advantage in math courses. However, while girls outperform boys, there are hidden problems. For example, high-achieving girls who clearly outperform boys often have the same level of confidence as boys with lower scores, i.e., they appear to be less confident despite their good grades (Kent, Sharma, Malliaris, Jukic, & Varma, 2023).

Gender differences may be manifested differently in different domains, and Green (1993) conducted a study on music, gender, and education. The study was conducted on 78 music teachers in the United Kingdom. The results of the study showed that girls were perceived to be more emotionally close, open, communicative, hardworking, driven, reliable, and persistent than boys, while boys had more talent for imagination, ability to manipulate technique, and innovative or challenging forms of music. Yakut (2017) also explored gender stereotypes related to athletic training and practice factors. The study was conducted on 36 young adults, 20 males and 16 females with an average age of 23 years. The results of the study showed that males had a greater advantage in performing visual exercises, whereas females showed a more deliberate exercise strategy and seemed to emphasize accuracy more than males.

However, unlike gender differences in academic performance, there is no significant difference in the number of students' classroom discourse Raviv and Aflalo (2023) investigated gender differences in students' classroom discourse in physics. The study was conducted on two unisex high school classes, boys' and girls' classes, aged 16 to 17 years. The results of the study showed that the average number of all classroom discourse parameters was similar in both classes. From the teachers' perspective, if there were differences between the two classes, they were related to students' personality and classroom environment rather than students' gender.

Sigit, Suryanda, Suprianti and Ichsan (2019) to study the effect of gender on learning outcomes of biodiversity among high school students. The study was conducted on a simple random sample of 114 students (57 males and 57 females). The results of the study showed that there was a significant difference between the average learning outcomes of male and female students. Female students had higher academic achievement than male students.

Another example is Noroozi, Banihashem, Taghizadeh Kerman, Parvaneh Akhteh Khaneh, Babayi, Ashrafi and Biemans (2022) explored gender differences in argumentative essay writing and peer review performance and absorptive capacity in a higher education context. The study utilized a pre-test and post-test design for the experiment. A total of 101 students majoring in the field of environmental science were enrolled in the experiment. These students were asked to individually write a dissertation, participate in a peer review exercise, and revise their original dissertation based on the peer reviews received. The results of the study showed that female students performed better than male students in terms of topic stance in writing the paper. There were also gender differences in peer review performance, with females providing better reasoning, more constructive reviews, and higher quality peer reviews for issues identified during peer review compared to males. However, in another study, the results were opposite to the previous two results. Hashemi, Si Na, Noori and Orfan (2022) explored the gender acceptance of ICT use in English language learning. The study used a guantitative research method. The population of the study was 152 students (94 males and 58 females) from Takhar University in Afghanistan. The results of the study showed that most of the male respondents believed the use of ICT tools and applications could help them to improve their English language skills. On the other hand, female respondents agreed that the use of ICT could help improve reading comprehension, but their acceptance level was generally lower compared to that of male respondents.

Summarizing the above, regarding the impact of gender on learning, there is no certain result that male and female students have different strengths and weaknesses in different subjects or situations. Therefore, the effect of gender on Kahoot's learning pattern is a research topic worth exploring and is the focus of this study.

## 3. Method

#### 3.1 Research Architecture

The research architecture is shown in Figure 1. The independent variable is gender. The dependent variables are classroom engagement, flow experience, and learning effectiveness. The control variables are teacher, teaching content, grade level of participants, and prior education.



Figure 1. Research Architecture

#### 3.2 Subjects

The subjects of this study were 31 students in the 12th grade of a high school in central Taiwan who were enrolled in the "Arts and Life-applied Music" course. Their ages ranged from 17 to 18 years old. There were 15 male students and 16 female students.

## 3.3 Research Tools

The research tools used in this study included the Kahoot game-based answering platform, the pre- and post-tests of the "Arts and Life-applied Music", the classroom participation questionnaire, and the flow experience questionnaire. The pre- and post-tests were provided by the instructor and consisted of 20 multiple-choice questions each with a similar level of difficulty. The questions were all about the basic core knowledge of the content of "Arts and Life-applied Music". The classroom engagement questionnaire is based on a questionnaire designed by Elmaadaway (2017). The questionnaire was modified from the one proposed by Jamaludin and Osman (2014) and was divided into three components: behavioral engagement (10 questions), cognitive engagement (7 questions), and emotional engagement (8 questions), with a total of 25 questions. The Heart Flow Experience Questionnaire was modified by Pearce, Ainley and Howard (2005) and consisted of 8 questions. The questionnaires were reviewed and modified by two experts in the field of digital learning with more than 10 years of experience in the field for expert validity.

#### 3.4 Procedure

The experimental procedure of this study is shown in Figure 2. The teacher first gave the students a 15-minute pre-test to find out the students' prior knowledge of the "Arts and Life-applied Music" program. Then, the teacher gave a lecture, which was accompanied by an instructional video and a textbook to explain the subject matter of Arts and Life-applied Music. Each 15-minute lecture was interspersed with a 5-minute Kahoot game, with a total of three rounds of Kahoot games, as shown in Figure 3. Afterward, a 10-minute post-questionnaire on classroom participation was filled out. Finally, a 15-minute post-test will be administered.



Figure 2. Procedure Chart



Figure 3. Kahoot Game in Classroom

#### 3.5 Data Analysis

In this study, the reliability of the questionnaires was analyzed, and then a paired-sample ttest was used to determine whether there was a significant improvement in the learning outcomes of the overall student population. Next, an independent sample t-test was conducted to determine whether there was a significant difference between male and female students in terms of classroom engagement (including behavioral participation, cognitive participation, and affective participation), flow experience, and gain scores.

## 4. Results

## 4.1 Reliability Analysis

The reliability of the questionnaires in this study ranged from 0.921 to 0.955, which were all greater than 0.9, indicating that the results of the questionnaires in this study were very reliable.

#### 4.2 Paired Sample t-test

The results of t-test of the paired samples of the pre- and post-tests are shown in Table 1. The mean of the post-tests is higher than that of the pre-tests, which means that using Kahoot for teaching activities has a significant effect on students' performance (t=-20.476, p<0.001). From the results of the analysis, using Kahoot for teaching can increase male and female students' classroom interaction, learning motivation and confidence, and also make students more immersed in classroom teaching, which can effectively improve students' academic performance.

Group	Sample size	Average value	Standard deviation	<i>t</i> -value
Pre-test	31	42.90	12.300	20 476***
Post-test	31	88.06	14.299	-20.4/0

#### Table 1. The Results of Paired Sample t-test

\*\*\*p<0.001

#### 4.3 Independent Sample t-test

The t-test results of the independent samples of male and female students are shown in Table 2. The mean values of all constructs of the test were higher for female students than for male students, with significant differences in behavioral involvement (t=-2.666, p<0.05), cognitive involvement (t=-3.263, p<0.01), affective involvement (t=-2.979, p<0.01), and flow experience (t=-2.621, p<0.05). In other words, when Kahoot was applied to classroom teaching, female students' classroom engagement and flow experience were significantly higher than male students.

Facet	Group	N	Average	SD	<i>t</i> -value
Behavioral	Male	15	3.720	0.679	2666*
Engagement	Female	16	4.363	0.662	-2.000*
Cognitive	Male	15	3.505	0.691	2 762**
Engagement	Female	16	4.304	0.672	-3.203
Emotional	Male	15	3.758	0.709	2 070**
Engagement	Female	16	4.492	0.662	-2.9/9

Table 2. The Results of Independent Samples t-test

Flow	Male	15	3.950	0.789	2 621*
Experience	Female	16	4.641	0.677	-2.021
Cain Saara	Male	15	42.00	12.780	1 410
Gain Score	Female	16	48.13	11.383	-1.410

\*p<0.05 \*\*p<0.01

## 5. Discussion

#### 5.1 Classroom Engagement

The results of this study showed that girls were significantly more engaged than boys when using Kahoot in the classroom. However, this result is contrary to the findings of Aquillon, Siegmund, Petipas, Drake, Cotner and Ballen (2020). The study examined the impact of gender differences on students' classroom engagement. The study was conducted with 244 and 265 undergraduates enrolled in a two-semester introductory biology course at a large research university, with a gender ratio of 55.9% and 57.1%, respectively. It was found that males in most classes' engagement more than expected, especially after the group discussions in both semesters, with a significantly higher proportion of males than females responding voluntarily. This difference may be due to different subjects and different teaching activities. For example, in math and science related classes, males will have higher participation while in language related classes, females will have higher participation (Hyde, 2014). Males are likely to be more engaged than females when using physical teaching methods, while the opposite may be true when using online teaching methods (Chen & Tsai, 2007). Various factors interact with each other to influence student engagement and understanding the impact of gender differences in learning can help to make appropriate adjustments in teaching.

## 5.2 Flow Experience

The results of this study showed that female students had significantly higher flow experiences than male students when using Kahoot in the classroom. This result is like the findings of Hsieh, Lin, and Hou (2016). The study investigated the relationship between mental stream experiences and academic performance of primary school students. Gender and grade differences were also investigated. A total of 34 elementary school students participated in the study. The results of the study showed that students with higher levels of mindfulness tended to have higher levels of learning performance. Gender differences showed that female students in this study had higher performance scores in the mini-educational game and had a better mind-flow experience.

## 5.3 Learning Outcomes

In this study, because of the small sample size and the unequal distribution of pre-test scores, the learning outcome of this study is the students' gain score, which is the post-test minus the pre-test scores. The results showed that when Kahoot was used in the classroom, the gain scores (learning outcomes) of female students were slightly higher than male students, but not to a significant level. This result is close to that of Johnson (2011) and Nguyen, Hou, Richey and McLaren (2022). In Johnson's (2011) study, the performance of gender differences in digital learning was examined. The subjects of the study were 303 males and 252 females. The results of the study showed that females had slightly better learning outcomes than males when compared to males. In Nguyen, Hou, Richey and McLaren's (2022) study, differences in learning outcomes between male and female students were examined in a Math Learning Game. The study was conducted with 624 students in grades 5 and 6 at the elementary school level. The results of the study showed that the improvement scores of female students were significantly higher than those of male students. In other words, Math Learning Game is more valuable to female students.

#### 6. Conclusions and Recommendations

#### 6.1 Conclusions

In this study, the male students were not as engaged as the female students in terms of participation, flow experience, and gain scores. This study hypothesizes that it is possible that the subject matter of the Arts and Life-applied Music textbook is oriented towards the music art, and that the female students are more interested in this area of knowledge. In addition, for the answering game, because girls have better verbal ability, even if they use cell phones to press the button to answer the question, the girls' group will have a lot of speeches and discussions, and it is easier to strengthen the impression of learning in these processes. On the other hand, male students tend to be silent and thinking more than talking, reducing the language exchange and interaction during the answering, which makes the classroom atmosphere of male students much different from that of female students in terms of rationality.

#### 6.2 Recommendations

In order to adjust how to improve boys' participation and test scores, we can consider adding graphical symbols or mind maps in classroom lectures or answering games to enhance the logical view instead of purely textual narration. In addition, male students may be more sensitive to body movement than female students, and a long classroom lecture is not likely to attract the attention of high school-aged male students for a long period of time. Therefore, in addition to the Kahoot answering, the addition of hands-on activities or the need to move seats to change the grouping of seats may increase the sense of freshness in the classroom and increase male students' participation.

#### Acknowledgements

This research was funded by the National Science and Technology Council, Taiwan, R.O.C. with grant number NSTC 111-2410-H-224-004-MY2.

## References

- Aguillon, S. M., Siegmund, G. F., Petipas, R. H., Drake, A. G., Cotner, S., & Ballen, C. J. (2020). Gender differences in student participation in an active-learning classroom. *CBE—Life Sciences Education*, *19*(2), ar12.
- Aljezawi, M. E., & Albashtawy, M. (2015). Quiz game teaching format versus didactic lectures. *British Journal of Nursing*, 24(2), 86-92.
- Aziz, F., & Kazi, A. S. (2019). Role of teachers in students' classroom participation in universities. International Journal of Educational Enquiry and Reflection, 4(1), 46-57.
- Chen, R. S., & Tsai, C. C. (2007). Gender differences in Taiwan university students' attitudes toward web-based learning. *Cyberpsychology & Behavior*, *10*(5), 645-654.
- Clarke, S. G., & Haworth, J. T. (1994). 'Flow'experience in the daily lives of sixth-form college students. *British Journal of Psychology*, *85*(4), 511-523.
- Dellos, R. (2015). Kahoot A digital game resource for learning. *International Journal of Instructional Technology and Distance Learning*, *12*(4), 49-52.
- Elmaadaway, M. A. N. (2017). The effects of a flipped classroom approach on class engagement and skill performance in a blackboard course. *British Journal of Educational Technology*, *49*(3), 479-491.
- Ghalley, L. R., & Rai, B. M. (2019). Factors influencing classroom participation: A case study of Bhutanese higher secondary student. *Asian Journal of Education and Social Studies*, *4*(3), 1-14.
- Green, L. (1993). Music, gender and education a report on some exploratory research. *British Journal of Music Education*, *10*(3), 219-253.

- Hashemi, A., Si Na, K., Noori, A. Q., & Orfan, S. N. (2022). Gender differences on the acceptance and barriers of ICT use in English language learning: Students' perspectives. *Cogent Arts & Humanities*, 9(1), 2085381.
- Hsieh, Y. H., Lin, Y. C., & Hou, H. T. (2016). Exploring the role of flow experience, learning performance and potential behavior clusters in elementary students' game-based learning. *Interactive Learning Environments*, 24(1), 178-193.
- Hyde, J. S. (2014). Gender similarities and differences. Annual Review of Psychology, 65, 373-398.
- Jamaludin, R., & Osman, S. Z. (2014). The use of a flipped classroom to enhance engagement and promote active learning. *Journal of Education and Practice*, *5*, 124–131.
- Johnson, R. D. (2011). Gender differences in e-learning: Communication, social presence, and learning outcomes. *Journal of Organizational and End User Computing (JOEUC)*, 23(1), 79-94.
- Kent, P., Sharma, A., Malliaris, M., Jukic, N., & Varma, A. (2023). Perceived differences in confidence and ability of females: the role of human resources. *International Studies of Management & Organization*, 53(2), 104-123.
- Kiili, K., De Freitas, S., Arnab, S., & Lainema, T. (2012). The design principles for flow experience in educational games. *Procedia Computer Science*, *15*, 78-91.
- Lesperance, K., Hofer, S., Retelsdorf, J., & Holzberger, D. (2022). Reducing gender differences in student motivational-affective factors: A meta-analysis of school-based interventions. *British Journal of Educational Psychology*, *92*(4), 1502-1536.
- Licorish, S. A., Owen, H. E., Daniel, B., & George, J. L. (2018). Students' perception of Kahoot's influence on teaching and learning. *Research and Practice in Technology Enhanced Learning*, *13*(1), 1-23.
- Mundelsee, L., & Jurkowski, S. (2021). Think and pair before share: Effects of collaboration on students' in-class participation. *Learning and Individual Differences*, *88*, 102015.
- Nguyen, H. A., Hou, X., Richey, J. E., & McLaren, B. M. (2022). The impact of gender in learning with games: A consistent effect in a math learning game. *International Journal of Game-Based Learning (IJGBL)*, *12*(1), 1-29.
- Noroozi, O., Banihashem, S. K., Taghizadeh Kerman, N., Parvaneh Akhteh Khaneh, M., Babayi, M., Ashrafi, H., & Biemans, H. J. (2022). Gender differences in students' argumentative essay writing, peer review performance and uptake in online learning environments. *Interactive Learning Environments*, 1-15.
- Pearce, J. M., Ainley, M., & Howard, S. (2005). The ebb and flow of online learning. *Computers in Human Behavior*, *21*(5), 745-771.
- Rajabpour, A. (2021). Teachers' perception of advantages and disadvantages of Kahoot. *English Linguistics Research*, *10*(4), 49-61.
- Raviv, A., & Aflalo, E. (2023). Classroom Discourse in Single-Sex Physics Classes: A Case Study. *European Journal of Science and Mathematics Education*, *11*(1), 182-196.
- Sigit, D. V., Suryanda, A., Suprianti, E., & Ichsan, I. Z. (2019). The effect of adversity quotient and gender to learning outcome of high school students. *International Journal of Innovative Technology and Exploring Engineering*, 8(6), C2.
- Uzunboylu, H., Galimova, E. G., Kurbanov, R. A., Belyalova, A. M., Deberdeeva, N. A., & Timofeeva, M. (2020). The Views of the Teacher Candidates on the Use of Kahoot as a Gaming Tool. *International Journal of Emerging Technologies in Learning*, *15*(23).
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174-1204.
- Wang, A. I., & Tahir, R. (2020). The effect of using Kahoot for learning–A literature review. *Computers & Education*, *149*, 103818.
- Wang, A. I., Øfsdahl, T., & Mørch-Storstein, O. K. (2007, November). Lecture quiz-a mobile game concept for lectures. In Proceedings of the 11th IASTED International Conference on Software Engineering and Application (SEA'07) (pp. 305-310).
- Yakut, C. (2017). Gender stereotyping, training and practice factors related to learning a complex visual motor task. *Cogent Psychology*, *4*(1), 1326576.
- Zhang, Q., & Yu, Z. (2021). A literature review on the influence of Kahoot on learning outcomes, interaction, and collaboration. *Education and Information Technologies*, 26(4), 4507-4535.

# Implementation of collaborative projectbased learning approach: Spherical videobased virtual reality creation

# Husni MUBAROK<sup>a</sup>, Gwo-Jen HWANG<sup>a,b\*</sup>, Chi-Jen LIN<sup>c</sup> & Darmawansah DARMAWANSAH<sup>a</sup>

 <sup>a</sup> Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taiwan
 <sup>b</sup> Graduate Institute of Educational Information and Measurement, National Taichung University of Education, Taiwan
 <sup>c</sup>Higher Education SPROUT Project Office, Chung Shan Medical University, Taiwan
 \*gjhwang.academic@gmail.com

**Abstract:** There have been obstacles for individual students to participate in projectbased learning, such as the lack of learning engagement and the frustration of encountering difficulties. Therefore, this study proposed a collaborative project-based learning approach with SVVR technology to enhance students' project-based learning outcomes. A quasi-experimental design was conducted to show the effectiveness of the proposed approach. Two classes students in a university participated in the experiment. On class with 25 students were in the experimental group, and 20 students were in the control group. The experimental group experienced collaborative projectbased learning with SVVR, while the control group underwent individual project-based learning with SVVR. The results of this study revealed that the experimental group outperformed the control group in terms of students' computational thinking and the SVVR project quality. This study highlights the effectiveness of collaborative projectbased learning in enhancing students' computational thinking and creativity.

Keywords: PBL, SVVR, project-based, creativity, collaborative

## 1. Introduction

Project-based learning (PBL) has become increasingly popular as a way to adapt to a changing world, and it allows students to learn by doing through a project (Oguz-Unver & Arabacioglu, 2014; Thomas, 2000). This approach has been shown to have a positive impact on students' deep learning and to develop their higher-order skills (Baran et al., 2018; Kızkapan & Bektaş 2017). Despite these benefits, there have been obstacles and skepticism in implementing project-based learning successfully. For example, if the design of a project-based learning the project (Huang et al., 2023). Additionally, evaluating the results of project-based learning can be challenging, as most assessments are group-based (Loertscher, 2008). To address these challenges, virtual reality (VR) technology is proposed as an alternative solution for project-based learning. This technology can serve as an effective learning medium for both individual and group projects, and it offers a more engaging and immersive experience for students.

Over the past ten years, virtual reality has become a commonly used tool in both research and teaching, particularly in the PBL approach (Won et al., 2022). For instance, a qualitative case analysis conducted by Morales et al. (2013) showed that project-based learning in virtual reality classrooms can be effective, even with minimal teacher guidance, and has great potential for independent and peer-mentored learning. Hou et al. (2023) conducted a project-based learning approach for virtual reality-aided green building education, which demonstrated the effectiveness of this novel approach in delivering green building

education. Additionally, Zang et al. (2022) linked project-based learning and semi-immersive virtual reality to improve learning performance, providing valuable insights for educational institutions looking to enhance student performance through project-based learning opportunities.

On the other hand, scholars have pointed out that creating realistic 3D interactive VR content can be expensive and time-consuming, making it difficult for many schools to afford (McFaul & FitzGerald, 2019). Additionally, designing VR content can be challenging for most teachers (Merchant et al., 2014). Furthermore, teachers require user-friendly tools and materials to use as learning resources for PBL (Gerhátová et al., 2021). To address these challenges, spherical video-based virtual reality (SVVR) is an excellent option as it reduces the cost and time required to create VR content. Importantly, creating SVVR content does not require advanced technical skills (Chien et al., 2020), making it a suitable option for students' PBL activities.

To improve students' computational thinking skills and promote innovative thinking in a project, it is essential to encourage teamwork and the exchange of ideas with peers, as suggested by Baser et al. (2017) and Kerimbayev et al. (2023). Vasileva et al. (2011) also support this idea by using project-based collaborative learning in a web-based virtual environment to facilitate synchronous and asynchronous communication among students. Computer-supported collaborative learning (CSCL) environments, as noted by Chan (2012) and Sidek et al. (2020), can assist students in self-regulated learning, sharing ideas, and enhancing their creativity. Therefore, this study proposed a collaborative project-based learning approach with SVVR technology to improve students' computational thinking and creativity. The research questions were proposed as follows:

- 1. Is there any difference between the SVVR-collaborative project-based learning approach and the SVVR-individual project-based learning approach in terms of the students' computational thinking?
- 2. Is there any difference between the SVVR-collaborative project-based learning approach and the SVVR-individual project-based learning approach in terms of the students' SVVR project?

#### 2. Literature review

In the rapidly evolving digital era, characterized by rapid technological advancements and an increasing reliance on information technology, computational thinking has emerged as a crucial skill to navigate the complexities of modern life effectively. This cognitive skill encompasses a multifaceted approach to problem-solving, enabling individuals to analyze, and synthesize complex problems into manageable components, ultimately arriving at innovative and efficient solutions (Wing, 2006). Collaborative project-based learning, integrated with SVVR technology, has been identified as a promising approach to foster and apply computational thinking skills among students (Saad & Zainudin, 2022). Engaging in collaborative tasks within virtual environments allows students to practice problem-solving, enhancing their ability to tackle complex challenges in a systematic manner (Esteves et al., 2006). In addition, collaborative project-based learning with SVVR has garnered considerable attention as an effective means to promote deeper learning and teamwork among students (Baran et al., 2018; Kızkapan & Bektaş, 2017).

The integration of SVVR technology into collaborative project-based learning has demonstrated significant benefits, leading to improved computational thinking and creativity among students. The experiences within virtual environments equip students with practical skills and the confidence to apply their knowledge effectively (Hou et al., 2023). Collaborating in SVVR environments fosters creativity by encouraging students to share ideas, explore diverse perspectives, and experiment with innovative solutions (Bilyatdinova et al., 2016; Guan et al., 2023). The nature of SVVR technology provides a fertile ground for students to explore their creative potential and express their ideas through unique content creation (Parmaxi, 2023).

## 3. Methodology

#### 3.1 Participant

This study adopted a quasi-experimental design, with a total of participants were 45 freshmen in a university. These students were from two classes, which were randomly treated in the experimental group (n = 25) and the control group (n = 20). The experimental group received instruction through a collaborative project-based learning approach with the use of spherical video-based virtual reality, while the control group received instruction through an individual project-based learning approach with the use of spherical video-based virtual reality.

#### 3.2 Experimental procedure

The experimental procedure of this study is shown in Figure 1. Both the experimental and control groups participated in classes and activities that lasted for four weeks, meeting once a week for 100 minutes each time. In the first week, both groups received an introduction to project-based learning activities and SVVR technology to learn how to use the technology, and pre-questionnaire of students' computational thinking. The same instructor taught both groups during the learning activities, and they used the same SVVR app, specifically the iStaging app. The interface of the iStaging app is shown in Figure 2. This application allowed students to practice creating content, designing the interface, organizing information, and sharing their creations. Both groups used the same learning material for two weeks. The learning activities process is shown in Figure 3. The major difference between the two groups was the treatment of student activities (collaborative and individual). For the experimental group, students used a collaborative project-based learning approach with spherical videobased virtual reality. While in the control group, students used an individual project-based learning approach with spherical video-based virtual reality. After that, the teacher gave students a project assignment and a post-questionnaire of students' computational thinking to evaluate the impact of the treatment individually.



Figure 1. The experimental procedure

**Photo embedded:** Description of the picture and location

Various component: 360 images, immersed pictures, notes, voices, and videos.



Video embedded: Share link of youtube video. And watch the video in iStaging App

Page: Istaging App provide several pages for input some information with different location



Figure 2. The interface of SVVR in the iStaging app



*Figure 3.* PBL activities process

## 3.3 Instrument

A questionnaire was utilized to evaluate the students' computational thinking. The item in the questionnaire was presented with a 5-point Likert scale, ranging from 1 as strongly disagree to 5 as strongly agree. The students' computational thinking questionnaire was adopted from Hwang et al. (2020), which consisted of six items with a Cronbach's alpha of .77.

To examine students' SVVR projects, two instructors was developed a rubric by referring to Chang and Hwang (2018), Fegely and Cherner (2021), and Hwang et al. (2018). The rubric consists of four dimensions: content knowledge, complexity, appearance, and creativity, as presented in Table 1. Each criterion is rated on a five-point scale ranging from 0 (lowest) to 5 (highest). The first dimension, content knowledge, evaluates students' ability to provide accurate, detailed, and appropriate information. The second dimension, complexity, measures students' proficiency in using the iStaging app to include various components, such as 360 images, immersed pictures, notes, voices, and videos. The third dimension, appearance, assesses students' ability to organize the content effectively and create a visually appealing layout for easy viewing. Lastly, the dimension of creativity measures students' level of creativity in developing their SVVR projects.

Dimension		Scale								
DIMENSION	1	2	3	4	5					
Content knowledge	No information	Simple information, but not appropriate	Lots of information, but not appropriate	Simple and appropriate information	Detailed and appropriate information (such as, video, voice, link, etc.)					
Complexity	Only 360 image	Two components (360 image	Three components (360 image,	Four components (360 image,	Five or more components (360 image,					

 Table 1. The rubric of SVVR project

Dimension			Scale		
Dimension	1	2	3	4	5
		and immersed picture)	immersed picture, and note)	immersed picture, note, and voice)	immersed picture, note, voice, video, etc.)
Appearance	Uncomforta ble to watch and poor layout	Uncomfortabl e to watch and normal layout	Comfortable to watch, but poor layout	Comfortable to watch and normal layout	Comfortable to watch and good layout with a neat structure
Creativity	The work creativity is extremely poor	The work creativity is poor	The work creativity is mediocre	The work creativity is good	The work creativity is excellent

#### 4. Result

To analyze the prior competence of students' computational thinking in the two groups was used the independent sample t-test, and the result showed that there was no significant level was found (t = 0.64, p > .05), indicating that the prior students' computational thinking of the two groups was equivalent before the learning activity. Subsequently, a within-group homogeneity regression analysis indicated no significant difference between the two groups (F = 0.001, p > .05), confirming the homogeneity test passed. Then, Analysis of Covariance (ANCOVA) was employed to analyze the post-questionnaire scores of the two groups by excluding the effect of the pre-questionnaire. Table 2 shows the ANCOVA result. The adjusted scores of the experimental and control groups are 4.25 and 3.71, and the F score is 9.57 (p < .01,  $\eta^2 = 0.19$ ), showing a high effect size (Cohen, 1988). Consequently, it was concluded that the students who learned with a collaborative project-based learning approach had significantly better computational thinking than those who learned with an individual project-based learning approach.

Variable	Group	Ν	Mean	SD	Adjusted mean	SE	F	η²
Computational	Experimental	25	4.26	0.51	4.25	0.12	9.57**	0.19
thinking	Control	20	3.69	0.68	3.71	0.13		

Table 2. The ANCOVA result of computational thinking students

\*\*p < .01.

The students' SVVR project score in this study was analyzed by independent sample t-test. For the first dimension, content knowledge, was evaluated using Levene's test of error variance, which showed that the two groups were homogeneous (F = 0.12, p = 0.73 > 0.05). Table 3 shows the mean and standard deviation for the experimental group were 3.88 and 1.13, respectively, while the mean and standard deviation for the control group were 3.00 and 1.17. A t-test revealed a significant difference in content knowledge dimension between the two groups (t = 2.56, p < 0.05), indicating that the use of the SVVR-collaborative project-based learning approach improved students' competence in this dimension.

Similarly, for the complexity dimension, Levene's test for error variance showed that the two groups were homogeneous (F = 1.45, p = 0.24 > 0.05). The mean and standard deviation for the experimental group were 3.76 and 1.01, respectively, while the mean and standard deviation for the control group were 2.90 and 1.45. A t-test showed a significant difference in complexity between the two groups (t = 2.34, p < 0.05), it means that the use of

SVVR-collaborative project-based learning enhanced students' capability in the complexity dimension.

However, for the appearance dimension, Levene's test for error variance indicated that the two groups were also homogeneous (F = 0.73, p = 0.40 > 0.05). The mean and standard deviation for the experimental group were 3.80 and 0.87, respectively, while the mean and standard deviation for the control group were 3.40 and 0.94. A t-test showed no significant difference in appearance dimension between the two groups (t = 1.48, p > 0.05), suggesting that the use of SVVR-collaborative project-based learning did not significantly enhance students' capability in this dimension.

For the creativity dimension, Levene's test for error variance also showed that the two groups were homogeneous (F = 0.08, p = 0.78 > 0.05). The mean and standard deviation for the experimental group were 4.04 and 0.89, respectively, while the mean and standard deviation for the control group were 3.10 and 0.97. A t-test revealed a significant difference in creativity between the two groups (t = 3.39, p < 0.01), indicating that the use of SVVR-collaborative project-based learning enhanced students' creativity.

When considering the average for all four dimensions, Levene's test for error variance indicated that the two groups were homogeneous (F = 0.09, p = 0.77 > 0.05). The mean and standard deviation for the experimental group were 3.87 and 0.87, respectively, while the mean and standard deviation for the control group were 3.10 and 0.98. A t-test showed a significant difference in the average dimension between the two groups (t = 2.79, p < 0.01), it shows that the use of SVVR-collaborative project-based learning enhanced students' capability in each dimension.

Test	Group	Ν	Mean	SD	t	d
Content	Experimental	25	3.88	1.13	2.56*	0.76
knowledge	Control	20	3.00	1.17		
Complexity	Experimental	25	3.76	1.01	2.34*	0.69
	Control	20	2.90	1.45		
Appearance	Experimental	25	3.80	0.87	1.48	
	Control	20	3.40	0.94		
Creativity	Experimental	25	4.04	0.89	3.39**	1.01
	Control	20	3.10	0.97		
Average	Experimental	25	3.87	0.87	2.79**	0.83
-	Control	20	3.10	0.98		

 Table 3. Independent sample t test of students' SVVR project

\**p* < .05, \*\**p* < .01.

#### 5. Discussion and conclusion

This study proposed a collaborative project-based learning approach with SVVR technology to improve students' computational thinking and creativity in SVVR content creation. According to the results of the study, the students' computational thinking and the average dimension of the SVVR project revealed that students in the experimental group have significant differences compared to students in the control group. This shows that the collaborative project-based learning approach can be effective in enhancing productivity and positive feelings in facing challenges (Song, 2018; Spoelstra, 2014) and improve students' computational thinking (Li et al., 2023; Zhou & Tsai, 2022).

The study found a significant difference between the two groups in terms of content knowledge, complexity, and creativity dimensions. This finding supports the previous research indicating that collaborative project-based learning improves students' creativity and communication skills to discuss project ideas (Sidek et al., 2020; Vasileva et al., 2011). To produce exceptional and engaging content, one idea is often not enough, and teamwork is crucial for sharing ideas and complementing each other's strengths. The collaborative project-

based learning approach is a way to train students in teamwork and communication, particularly in project creation (Lee et al., 2017).

However, no significant difference was found in the appearance dimension between the two groups. Although the experimental group had a higher score than the control group, researchers consider that the collaborative project-based learning approach did not really affect organizing the layout of the SVVR project. Because the SVVR technology is userfriendly and does not require high-tech capabilities (Chien et al., 2020).

Additionally, to enrich the learning experience further, future studies could incorporate multiple learning strategies in conjunction with collaborative project-based learning. Techniques such as mind mapping, question prompts, and metacognitive strategies have shown promise in augmenting students' critical thinking and problem-solving abilities. By integrating these strategies into the SVVR project creation process, teachers may unlock additional potential for students to tap into their creative thinking and collaborative skills, propelling the learning experience to new heights.

In conclusion, the investigation of the collaborative project-based learning approach with SVVR technology has shed light on its multifaceted benefits. The findings support the notion that this pedagogical approach enhances students' computational thinking and creativity skills while fostering a positive and collaborative learning environment. While the appearance dimension did not exhibit significant differences, the overall impact on students' learning experiences and content creation was evident. The potential of SVVR technology coupled with collaborative project-based learning to revolutionize education is promising, and with further research and refinement, this approach could open up exciting possibilities for transforming how students engage with learning content and collaborate with their peers.

#### Acknowledgments

This study is supported in part by the Ministry of Science and Technology of Taiwan under contract numbers MOST 111-2410-H-011 -007 -MY3 and MOST-109-2511-H-011-002-MY3.

#### References

- Baran, M., Maskan, A., & Yasar, S. (2018). Learning Physics through Project-Based Learning Game Techniques. *International Journal of Instruction*, *11*(2), 221-234.
- Baser, D., Ozden, M. Y., & Karaarslan, H. (2017). Collaborative project-based learning: An integrative science and technological education project. *Research in Science & Technological Education*, *35*(2), 131-148.
- Bilyatdinova, A., Karsakov, A., Bezgodov, A., & Dukhanov, A. (2016). Virtual environment for creative and collaborative learning. In *Knowledge, Information and Creativity Support Systems: Selected Papers from KICSS'2014-9th International Conference, held in Limassol, Cyprus, on November 6-8, 2014* (pp. 371-381). Springer International Publishing.
- Chan, C. K. (2012). Co-regulation of learning in computer-supported collaborative learning environments: A discussion. *Metacognition and learning*, 7, 63-73.
- Chang, S. C., & Hwang, G. J. (2018). Impacts of an augmented reality-based flipped learning guiding approach on students' scientific project performance and perceptions. *Computers & Education*, *125*, 226-239.
- Chien, S. Y., Hwang, G. J., & Jong, M. S. Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Computers & Education*, *146*, 103751.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Esteves, M., Morgado, L., Martins, P., & Fonseca, B. (2006). The use of Collaborative Virtual Environments to provide student's contextualisation in programming. *Current developments in technology-assisted education*, *2*, 1496-1500.
- Fegely, A., & Cherner, T. (2021). A Comprehensive Rubric for Evaluating EduVR. *Journal of Information Technology Education*, 20.

Gerhátová, Ž., Perichta, P., Drienovský, M., & Palcut, M. (2021). Temperature measurement— Inquiry-Based learning activities for third graders. *Education Sciences*, *11*(9), 506.

- Guan, J. Q., Wang, L. H., Chen, Q., Jin, K., & Hwang, G. J. (2023). Effects of a virtual realitybased pottery making approach on junior high school students' creativity and learning engagement. *Interactive Learning Environments*, *31*(4), 2016-2032.
- Hou, H., Lai, J. H., & Wu, H. (2023). Project-based learning and pedagogies for virtual realityaided green building education: case study on a university course. *International Journal of Sustainability in Higher Education*.
- Huang, W., Li, X., & Shang, J. (2023). Gamified Project-Based Learning: A Systematic Review of the Research Landscape. *Sustainability*, *15*(2), 940.
- Hwang, G-J., Li, K-C. & Lai, C-L. (2020). Trends and strategies for conducting effective STEM research and applications: A mobile and ubiquitous learning perspective. *Int. J. Mobile Learning and Organisation*, *14*(2), 161–183.
- Hwang, G. J., Tu, N. T., & Wang, X. M. (2018). Creating interactive e-books through learning by design: The impacts of guided peer-feedback on students' learning achievements and project outcomes in science courses. *Journal of Educational Technology & Society*, 21(1), 25-36.
- Kızkapan, O., & Bektaş, O. (2017). The effect of project-based learning on seventh grade students' academic achievement.
- Lee, H. J., Kim, H., & Byun, H. (2017). Are high achievers successful in collaborative learning? An explorative study of college students' learning approaches in team project-based learning. *Innovations in Education and Teaching International*, *54*(5), 418-427.
- Li, W., Liu, C. Y., & Tseng, J. C. (2023). Effects of the interaction between metacognition teaching and students' learning achievement on students' computational thinking, critical thinking, and metacognition in collaborative programming learning. *Education and Information Technologies*, 1-25.
- Loertscher, D. (2008). Powerful Learning: What We Know About Teaching for Understanding. *Teacher Librarian, 36*(1), 57-58.
- McFaul, H., & FitzGerald, E. (2020). A realist evaluation of student use of a virtual reality smartphone application in undergraduate legal education. *British Journal of Educational Technology*, *51*(2), 572-589.
- Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education*, *70*, 29-40.
- Morales, T. M., Bang, E., & Andre, T. (2013). A one-year case study: Understanding the rich potential of project-based learning in a virtual reality class for high school students. *Journal of Science Education and Technology*, *22*, 791-806.
- Parmaxi, A. (2023). Virtual reality in language learning: A systematic review and implications for research and practice. *Interactive learning environments*, *31*(1), 172-184.
- Saad, A., & Zainudin, S. (2022). A review of Project-Based Learning (PBL) and Computational Thinking (CT) in teaching and learning. *Learning and Motivation*, 78, 101802.
- Sidek, R., Halim, L., Buang, N. A., & Arsad, N. M. (2020). Fostering scientific creativity in teaching and learning science in schools: A systematic review. *Jurnal Penelitian dan Pembelajaran IPA*, 6(1), 13-35.
- Song, Y. (2018). Improving primary students' collaborative problem-solving competency in project-based science learning with productive failure instructional design in a seamless learning environment. *Educational Technology Research and Development*, *66*, 979-1008.
- Spoelstra, H., Van Rosmalen, P., & Sloep, P. (2014). Toward project-based learning and team formation in open learning environments. *Journal of Universal Computer Science*, *20*(1), 57-76.
- Oguz-Unver, A., & Arabacioglu, S. (2014). A comparison of inquiry-based learning (IBL), problem-based learning (PBL) and project-based learning (PJBL) in science education. *Academia Journal of Educational Research*, *2*(7), 120-128.
- Thomas, J. W. (2000). A review of research on project-based learning. The Autodesk Foundation. *California: San Rafael*.

Vasileva, T., Tchoumatchenko, V., Lakkala, M., & Kosoncn, K. (2011). Infrastructure supporting collaborative project-based learning in engineering education. *International Journal of Engineering Education*, 27(3), 656.

Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33-35.

- Won, M., Ungu, D. A. K., Matovu, H., Treagust, D. F., Tsai, C. C., Park, J., ... & Tasker, R. (2022). Diverse approaches to learning with immersion virtual reality identified from a systematic review. *Computers & Education*, 104701.
- Zang, J., Kim, Y., & Dong, J. (2022). New evidence on technological acceptance model in preschool education: Linking project-based learning (PBL), mental health, and semiimmersive virtual reality with learning performance. *Frontiers in Public Health*, 10.
- Zhou, X., & Tsai, C. W. (2022). The Effects of Socially Shared Regulation of Learning on the Computational Thinking, Motivation, and Engagement in Collaborative Learning by Teaching. *Education and Information Technologies*, 1-18.

# Improving Thinking Awareness in Animation Scriptwriting Through Learning Supporting Tool

#### Panadda Jaiboonlue<sup>a</sup>, Wasan Na Chai<sup>b</sup>, Taneth Ruangrajitpakorn<sup>b</sup>, and Thepchai Supnithi<sup>b</sup>

<sup>a</sup>Department of Animation and Digital Media, Faculty of Science and Technology Bansomdejchaopraya Rajabhat University, Bangkok, Thailand panadda.ja@bsru.ac.th <sup>b</sup>Language and Semantic Technology Laboratory, National Electronics and Computer Technology Center, Pathumthanee, Thailand {wasan.na\_chai,taneth.rua,thepchai.sup}@nectec.or.th

Abstract: In the media industry, there is a high demand for animation films, leading to the establishment of animation film courses in various universities to train future animators. However, writing animation scripts poses challenges for students, as it requires critical thinking skills to craft captivating and coherent story ideas, akin to other creative works that adhere to general principles for a complete and logical narrative. To solve this issue, we developed a tool as part of our project to help students systematically organize the essential components of a story based on the fundamental principles of animation, consisting of three acts. This tool creates a conducive learning environment by breaking down crucial elements into clear sections, prompting students to reflect on their own ideas about animation stories. Additionally, the tool encourages collaboration between students and instructors, enabling constructive feedback and reflections through error corrections. The results of our experiments showed significant improvements in students' development after using the tool. Notably, the recurring errors in omitting vital parts of the script did not reoccur after the initial use of the tool. Its implementation heightened students' awareness of the importance of each component. Furthermore, the assessment scores of all students demonstrated a significant improvement, with 34% of students displaying increased awareness in scriptwriting. The elements that students commonly missed were 'conflict' and 'progression of complications,' respectively.

Keywords: scriptwriting theory, learning supporting tool, awareness of thought

## 1. Introduction

Animation is one of the creative media that can engage audiences with interesting stories and expressions of imagination (Redvall, 2009). Because animated films are more visually appealing and reach more imaginative possibilities than movies at a lower cost. In addition to the above, animated films are important in other areas such as education or marketing and advertising. By education, animation is used extensively in educational materials and e-learning platforms. It helps simplify complex concepts, making learning engaging and accessible for students of all ages or animation is widely used in marketing and advertising campaigns. For marketing and advertising, animated advertisements, explainer videos, and animated graphics help capture the attention of consumers. Additionally, animated media effectively communicates messages to consumers.

Currently, teaching and learning in the field of animation and digital media are popular among university students in Thailand. Learning about animation development in classrooms involves teaching from creating digital media to scriptwriting for animation films. Upon evaluating the teaching and learning in the aforementioned field between teachers and students. We found that in the subjects of theory and scriptwriting, they are among the most challenging topics for students to comprehend. Because it requires creative thinking coupled with a theoretical understanding of the essential elements of writing, including narrative techniques to captivate and engage the audience. Including creating an interesting story with clear elements is also important for screenplays and animated screenplays. It's important to keep the plot running smoothly and logically.

In this work, we develop a tool to help students to conceptualize components in their animation scripts into explicit objects. With the use of an ontology, conceptualized components are connected following script theory (Demorest & Alexander, 1992). This will help both students to recheck their idea and instructors to point out the mistakes. Moreover, the students will become more aware of necessary script elements and train their story composing for improvement.

## 2. Background

#### 2.1 Basic Theory in Animation Script Writing

Fundamentally, a script for an animation is similar to a screenplay of other medias including movie and theater play as they share common elements (Scott, 2003). An animation script is a form of literary as it is intended to be interpreted by other such as producer, graphic designer, illustrator, and voice actor. Therefore, an animation script is not an end-product that audiences can directly enjoy. Apparently, an animation script is written using technical jargon and tight, spare prose when describing stage directions. Unlike a novel or short story (Marx, 2021), an animation script focuses on describing the literal, visual aspects of the story, rather than on the internal thoughts of its characters. Hence, the aim of script writing is to evoke thoughts and emotions through subtext, action, and symbolism.

The general script writing theory is an approach to organize the structure, goals and techniques of writing a script in a systematic fashion. The common and widely used structure of a script is a 'three act structure' (Ladopoulou and Dimitriadis, 2022). The three acts include a setup (first act), a confrontation (second act) and a resolution (third act) (Khalili, 2018) . Ratio of the acts is 25:50:25 percent for setup act, confrontation act, and resolution act, respectively.

• The setup act is to establish and to introduce the main characters, their relationships to other characters, and the world and its setting. In the later of first act, an inciting incident, known as the first plot point, is launched to raise a dramatic question that will be answered in a later act for the climax.

• The confrontation act features the main characters and their attempt to resolve the problem from the first turning point. In this act, a character development is introduced with the aid of other characters.

• The resolution act gives the resolution of the story. In this act, climax is mentioned as a sequence of the story are brought to the most intense point and an answer to the given dramatic question.

Besides the common 'three act structure', four or five acts are also proposed, but they are more complex and not suitable for novice script writers. Hence, this work focuses on gathering concepts related to the 'three act structure'.

For elements in an animation script, this depends on the respective act. However, the common elements are such as characters and their characteristics, locations of the stage, time of the scene, and major incidents of the story. These elements should not be in conflict to prevent a mazy viewpoint, and they should be carefully designed to logically relate to one another. The quality of a script is normally assessed by scriptwriting rubric (Gutierrez, 2014). Components in 3-act script paradigm is shown in Figure 1.



Figure 1. Components in 3-act script paradigm

# 2.2 Writing Supporting System

To assist on education, several researchers developed a tool on supporting writing learning. They design a tool with the environment to help students and instructors on improving writing expertise. We survey writing supporting systems and found the information given in Table 1.

Table. 1	Related	work on	supporting	tool for im	proving a	writing skill

Tool	Correction	Suggestion	Environment	Criteria	Collaborative
Eagle	Grammar	Error Location, Error Type	- Teacher hints the error location - Students consider error type and correct them	Error Type Structure based on "mistake" or "missing" type	Yes
SWAN	Content	Rewriting Sentence	- Apply NLP Parser to understand sentence structure - Highlight color to represent Fluidity	Fluidity (Connect, Inverted Topic, Out-of- Sync, Disconnected)	No
ArgRewrite	Content/ Grammar	Rewriting Sentences which labeled type	- Give statistics of document composition - Students reconsider the suggested sentence	Revision Type :Text based/Surface based type	No
TBeW	Content/ Thinkingg	Writing Sentence which Content Type and Relational Sentence	- Write Sentence - Connect Sentence Declare Sentence Rational	Assist learner to understand their own thinking	Yes
TBeR	Analyze/ Thinking	Reading good academic articles	<ul> <li>Reading</li> <li>Sentence</li> <li>Analyzing logical content</li> <li>Analyzing relations between content</li> </ul>	Assist learner to understand author's thinking systematically	Yes

In summary, Eagle (Takhom et al., 2011) is a tool for learning English as a foreign language for Thai students. It aims to improve understanding of English gramma with a collaboration between teachers and students. The tool allows teacher users to point out mistakes in written essays by assigning types of error for students to realize the mistake as well as keeping a log of editions made by students. SWAN (Kinnunen et al., 2012) is an assisting tool for writing scientific English content by providing an interface to guide the separation of content into required sections. The tool aims to help novice writers to understand purposes and relevant information of each section. ArgRewrite (Zhang et al., 2016) is a tool for learning a writing skill. It provides grammatical structure and vocabulary checking assistance to label errors in a English written sentence. It allows users to rewrite based on found errors by aiming to change writing behavior for the better. TBeW (Na Chai et al., 2017) and TBeR (Na Chai et al., 2019) is a framework that allows user to write and read content regardless of language used to understand the intention of content. This will help users to train their cognitive skills to think about thinking (metacognitive skill). The framework focuses on the awareness of how to represent the idea correctly in writing and how to determine the strategy to represent thought of author in reading.

From the review, we found that the tools can be categorized into two types. First type is a tool that allows student to learn how to write while the tool provides supportive functions to assist on learning. Second type is a tool providing environment for students to collaborate with peers for sharing opinions and written content to inspire each other and learn from other's examples. The key concepts of these supporting tools are to let students become aware of their own mistakes and improve from it by getting instructor comment or realizing them from examples. These tools thus assist students to learn more effectively than traditional lectures. Furthermore, most of the tools are online-based application and can be accessed flexibly via a personal computer, smart phones, and tablets. The tools hence help to promote learning from anywhere-anytime and life-long learning.

#### 3. Methodology

#### 3.1 Design of the animation script writing tool

In this work, we design a tool to help on a task of animation script writing for students (WriterDuet, 2022). The tool allows students to describe the components in an animation script into distinguished objects and assigns relationship among them freely. The tool applies the existing ontology called animation script ontology to control semantic meaning of the script components. An overview of the tool is shown in Figure 2.

Based on the illustrated overview, the core components of the tool are the ontology, the database to store conceptualized objects of an animation script, and user interface for users to provide and retrieve script information. The tool is designed for 2 user roles as students and instructors. The students are to provide the idea of the animation story in details in the provided user interface while the instructors play a role to explore the given information to detect issues regarding theoretical details of script writing and give comments for students to correct the mistakes.



Figure 2. An overview of the proposed tool

## 3.1.1 Ontology as a knowledge base

The chosen ontology consists of ontological concepts about animation script components and relations among them. The core concepts are listed in Table 2. The ontology component type include a class and a property. The components must be named differently. As an ontology provides the knowledge regarding of animation script components, we gain a privilege of ontology benefits including ontology axioms to describe how concepts are related and connected semantically. However, we selected some part of knowledge from the ontology that matches the focus of this work. In this work, we focus on the common paradigm of animation scripts which consists of 3 acts according to the knowledge from the applied ontology as illustrated in Figure 1. From the figure mentioned, the 3 acts are a phase of the story while there are specific events which exist for a certain purpose in each phase. In summary, there are 7 essential components shown in Table 3 (Panadda et al., 2023).

Component Name	Туре	Definition	Properties
Script	Class	A main division of a script following structural theory	<ul> <li>OP: has_incident [Event]</li> <li>OP: has_setting [Setting]</li> </ul>
Character	Class	a live form in an animation	<ul> <li>DP: has_name {string}</li> <li>OP: has_characteristic [Characteristic]</li> <li>OP: has_family-relation to [Character]</li> <li>OP: has_relation_to [Character]</li> </ul>
Setting	Class	Place, Time, Condition, and Environment of a story	-None-
Event	Class	a noteworthy happening in a story	<ul> <li>OP: involved_character [character]</li> <li>DP: located_in {string}</li> <li>OP: caused_by_prior_event [Event]</li> </ul>

Table	2	Details	of the	applied	ontology	and i	its	classes
i abie.	~	Details		applieu	Unitology	anui	15	CI22222

has_family- relation_to	Property	Familial relationship among character	-None-
has_relation_to	Property	Relationship among character	-None-

Table. 3 Animation	script com	ponents and	their s	specification
--------------------	------------	-------------	---------	---------------

	Script component	Description
		Introduction of the story including location
Act 1		setting, main characters, and specific conditions
		of the created story
•	Begin	A starting of the story
•	The inciting incident	An event that stimulates the story
Act 2		Provision of details of problems and solutions
•	Progressive Complication	An event that increases more issues in the story
•	Crisis	An event that leads to the critical issue
Act 3		Solving issue and Conclusion
٠	Climax	The most intensive decisive moment of the story
•	Solution	An event that solves the issues
•	End	Conclusion of the story

As shown in Figure 1, the acts are in sequence; hence, the relation of the acts are important to relay the story, especially for students who are yet excel in composing story. Furthermore, another important aspect in this study is to appropriateness of using the benefit of being animation movie. Thus, aside for completeness of the script details based on components, the soundness of each components part and aspect of animation appropriateness are also considered as essential in this work.

#### 3.1.2 User interface

With the aforementioned aspects, we design the user interface (UI) for students to fill in information of their scripts. The information includes title, story theme, character details, setting, and script components. The UI provides slots for each type of components separately, so each piece of information is objectified and differentiated. This will help students to become aware of the flaw of ideas while using the tool as well as realizing of their own thought. The screen-captured UI for script components is exemplified in Figure 3.

s	creenwriting su	pport tools					占 Wasan	C Logout
		N	ame: เดี๋ยวก่อน					
		Author: Wasan	Running time: 4.36 Min	Premise: [เพื่อเพื่อ	1			
1	Theme		Synops	8	Artwork/Poster Q. (9819)			
	• A Relational		B Content		Cor	C	Log	D & Delete
2	)				เหตุการเร่งเร้า	1 (Inciting Incident		
	▲ 2 1 ▼				เปิดเรื่อง (AC	Τ 1)	•	

Figure 3. Screen-captured UI of the tool for providing script components

From the screen captured UI, the (1) part is where students provide a theme of a script with additional information including summary and image (optional). The (2) part is where students give script details including the content (B) of each part and assigning the script component (C) to indicate the type of the given content of each slot roll. The list of components for selection follows the information given in Table 2. Students are also asked to assign relationship (A) of each content. Last, students can reach an instructor's comment from (D).

The information given in each slot is collected in a database and able to be retrieved for self-corrected and examined by instructors. For the instructor role, the tool provides a space for commenting on each component, and the comments are shown to users to reflect on mistakes.

#### 3.2 Providing Goal of the animation scriptwriting

The tool is used in lecturing in a course of training writing an animation script. According to the design, instructors can examine the completeness of script components, soundness of each component, and appropriateness of being animation movie. For details, description of each examining is given in Table 4.

Examining aspect	Criteria
Completeness of script components	Checking if an animation script is completed according to theory of 7 components
Soundness of each component	Checking if each component is correct to its description or not
Interestingness of the script	Checking understanding of the theory and how it can be applied to creativity
Appropriateness of being animation movie	Checking overall story that it takes benefit of being animation into account or not

**Table. 4** Animation script components and their specification

#### 3.3 Feedback from Coach

From the criteria, instructors can realize that what mistakes students make and provide a comment sufficient for students to correct them accordingly. Moreover, the collected information of scripts is stored in a database which can be used further in understanding the issue in teaching material or finding what needs to be focused for different groups of students. The coach is asked to provide feedback on the annotations. The options for the feedback 'agree' and 'disagree' while the 'disagree' has the option to provide additional comment on how the coach disagrees on the annotation. Particularly, the coaches are not suggested to give the correct answer directly, but the guidance on how to think to reach that conclusion. This will help to instill the correct cognitive process and the chance for learners to think about their way of thinking.

## 4. Experiment

#### 4.1 Experimental Results

To evaluate the effect of the tool, we set up participants in this experiment were 42 students who studied in a course of principles of animation and directing animation. The participants were asked to write 3 short animation scripts using the tool. For each script, there were 2 rounds for providing information. Each round allotted for 4 hours per round, and there were 7-day interval time from previous round. For each script information providing, the first round was an initial round where the participants provided the script information by themselves, and 3 instructors as evaluators examined the information and provided feedback to students. The second round was a reflected round which was for students to correct the

mistakes following the comment and allowed to edit the script components by their free will.

In this experiment, there were 3 evaluators who evaluated the given information independently based on criteria given in Table 3. For evaluators, evaluation measurement metric of information was as given in Table 5.

Aspect	Measurement
Completeness of equipt compensate	Range of 1 to 7 according to existing components
Completeness of script components	in the given information from students
Soundhoos of each component (7 components)	1 to 5 where 1 is lowest and 5 is highest
Soundness of each component (7 components)	(subjective to evaluators)
Interactingness of the earint	1 to 5 where 1 is poor and 5 is excellent
interestingness of the script	(subjective to evaluators)
Appropriateness of being animation movie	ves or no

 Table. 5 Evaluation measurement metric for each aspect

From 3 evaluators, the voting method was used to finalize the evaluation for aspect of completeness of script components, and appropriateness of being animation movie. For soundness of each component, the mode score was used to represent the final score of each of 7 elements. The evaluation results are split in 6 rounds as given in Table 6.

	Round of first script	Round of second script	Round of third script						
Initial	I-1	I-2	I-3						
Reflected	R-1	R-2	R-3						

 Table. 6 The evaluation results

The evaluation results of the aspect of completeness of script components, soundness of each component, and appropriateness of being animation movie are given in Table 7 and Table 8, respectively.

From the results, we found that students show noticeable improvement after using the tool, especially on aspect of completeness. By comparing results of I-1, I-2 and I-3, the result of all aspects indicate that students realized the mistakes in their prior attempts and stably improved the score. For the aspect of soundness of component, the score of students were separated into 3 groups based on total score where score group of 29-35 indicated students who did at least 4 points in average in all components. The results showed that the number of students who were in this group was increased in 'reflected round' comparing to its own initial round. Furthermore, the number of students in higher point group were increased when comparing between first script, second script and third script. The results of the aspect of appropriateness of being animation also showed improvement.

	Rating S	core	I-1	R-1	I-2	R-2	I-3	R-3
	Begin	1-5	2	3	3	4	3	4
Act 1	The Inciting Incident	1-5	2	3	2	3	3	4
Act 2	Progressive Complication	1-5	1	2	2	3	3	4
	Crisis	1-5	0	2	0	3	3	3
	Climax	1-5	1	3	2	3	3	4
Act 3	Resolution	1-5	2	3	2	4	3	4
	End	1-5	1	3	2	4	3	4
Total score		0-21	39	12	33	9	7	6
(0-35 points)	Overall	22-28	3	24	9	25	33	21
		29-35	0	6	0	8	2	15

Number of missing components	I-1	R-1	I-2	R-2	I-3	R-3
Missing 2 or more	5	2	3	0	1	0
Missing 1	9	3	5	0	2	0
No-missing	28	37	34	42	39	42

**Table. 8** Evaluation results of the aspect of completeness of script components based on rounds

#### 4.2 Discussion

From the experimental results, we can conclude that the tool helps students to improve their understanding and scope their thoughts/idea on script details. Thus, we aim to see how much students improved. We thus compare the score of the same students from the initial test (I) and reflect (R). By analyzing, we found the improvement rate of students in writing the first script and second script via the proposed tool as shown in Figure 4 and 5 respectively.



Figure 4. Comparison graph of score improvement between 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> written script

The improvement of students from Figure.4 show that all students improved from the initial round to the reflect round in the first, the second and the third script, but the improved score varied individually. These students were mostly in 0-21 score group and made the improvement to the mid-score group (22-28 score) and high score group (29-35 score) after reflection. The highest improved score was 18 higher score in the first script. The score improvement on the second script was averagely lower than the first script since the score initial round of the second script was noticeably higher than the first script.

In terms of common mistakes, the experimental results show that most missing components were components in Act2 including 'Crisis' and 'Progressive Complication', respectively. However, the tool helped to clearly see the required components, and students did not make this missing component again in the later round. This issue was hard to teach in traditional lectures since it is difficult to point out the missing part. This thus shows that the tool can help students to realize what is required as the basis of script writing.

#### 5. Conclusion

This work proposes a tool to support the learning of animation script writing. By developing an environment with clear requirements of script components based on related theory, students can realize what is needed to compose a complete and sound script. The tool allows students to systematically design the story and them of a script by decomposing the thought/idea into distinguished parts and assigning relations among them. Instructors can hence examine the part separately to find mistakes and give feedback for students to correct and reflect on their issues. This will help students to remember how to design with error examples so they would not make such mistakes again. From the experiment results, the students showed improvement in all focused aspects including completeness of script components, soundness of each component, and appropriateness of being animation movie. All students show progressive improvement as the evaluated score on the second script was averagely higher than the first script. Moreover, the issue of missing mandatory components has been improved significantly as 34% of participating students who made these mistakes did not make the mistakes after the first time they use the tool.

#### References

- Demorest, A. P., & Alexander, I. E. (1992). Affective scripts as organizers of personal experience. *Journal of Personality*, *60*(3), 645-663.
- Gutierrez, P. (2014). *The Power of Scriptwriting!: Teaching Essential Writing Skills Through Podcasts, Graphic Novels, Movies, and More.* Teachers College Press.
- Khalili, S. (2018). Analysing the advantages of Aristotle's two-act structure in comparison with Syd Field's three-act structure in short comedic animation scriptwriting. *Journal of Screenwriting*, *9*(3), 265-277.
- Kinnunen, T., Leisma, H., Machunik, M., Kakkonen, T., & Lebrun, J. L. (2012, April). SWAN-scientific writing AssistaNt. a tool for helping scholars to write reader-friendly manuscripts. The The 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2012), Avignon, France.
- Ladopoulou, G., & Dimitriadis, G. (2022). *The Female Perspective: Stereotypes Vs. Realism in Film Character Portrayal and Development* [Diploma Thesis, Aristotle University of Thessaloniki]. IKEE / Aristotle University of Thessaloniki Library.
- Marx, C. (2021). *Writing for animation, comics, and games*. CRC Press. https://www.sciencedirect.com/book/9780240805825/writing-for-animation-comics-and-games
- Na Chai, W., Ruangrajitpakorn, T., & Supnithi, T. (2017, December 4-8). A Tool for data acquisition of *thinking processes through writing*. The 25th International Conference on Computers in Education (ICCE 2017), Christchurch, New Zealand.
- Na Chai, W., Ruangrajitpakorn, T., & Supnithi, T. (2019, December 2-6). *A Tool for Learning of Cognitive Process by Analysis from Exemplar Documents*. The 27th International Conference on Computers in Education (ICCE 2019), Kenting, Taiwan.
- Jaiboonlue, P., & Na Chai, W. (2023). A Design of an Ontology for Animation Scripts Writing Towards the Learning of Necessary Elements and Relationship Among Them. *Journal of Information Science*, *41*(1), 66-85.
- Redvall, E. N. (2009). Scriptwriting as a creative, collaborative learning process of problem finding and problem solving. *MedieKultur: Journal of media and communication research*, *25*(46), 34-55.
- Scott, J. (2003). How to write for animation. The Overlook Press.
- Takhom, A., Trakul Taweekoon, K., Chotimongkol, A., Porkaew, P., Na-Thalang, S. S., & Supnithi, T. (2011, November-December 28-2). *EAGLE: an Error tAGger for Learners of English*. The 19th International Conference on Computers in Education (ICCE 2011), Chiang Mai, Thailand.
- WriterDuet. (2022, January 5). Effortless Screenwriting. WriterDuet. https://www.writerduet.com Zhang, F., Hwa, R., Litman, D., & Hashemi, H. B. (2016, June 12-17). Argrewrite: A web-based revision assistant for argumentative writings. The 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL HLT 2016), San Diego California, USA.

# The Trends of Computer-Supported Collaborative Learning in Two Decades

Hui-Chun Chu<sup>a</sup>, Gwo- Haur Hwang<sup>b</sup>, Han-Chieh Chao<sup>c</sup>, Ching-Yi Chang<sup>d\*</sup> <sup>a</sup>Department of Computer Science and Information Management, Soochow University, Taiwan

<sup>b</sup>Bachelor Program in Industrial Technology, National Yunlin University of Science and Technology, Taiwan

<sup>c</sup>Department of Electrical Engineering, National Dong Hwa University, Hualien City, Taiwan <sup>d</sup>\*School of Nursing, College of Nursing, Taipei Medical University, Taiwan

#### \*frinng.cyc@gmail.com

**Abstract:** Computer-supported collaborative learning (CSCL) has undergone significant advancements in the past two decades, revolutionizing the way learners engage with educational content and peers. This abstract explores key trends in CSCL, including the evolution of online communication and collaboration platforms, the impact of mobile technologies, the integration of data analytics, and the emergence of virtual reality (VR) and augmented reality (AR) environments. Additionally, the ongoing pandemic has accelerated the adoption of online and blended learning models, further shaping CSCL. These trends have transformed collaborative learning, allowing learners to exchange ideas, co-create knowledge, and engage in meaningful discussions. The utilization of data analytics enables personalized instruction and targeted support, enhancing learners' engagement and motivation. Immersive VR/AR environments promote active participation and deeper learning. Looking ahead, hybrid models combining face-to-face and online collaboration are likely to shape the future of CSCL. As instructional practices adapt and technology advances, creating engaging and effective learning environments remains a crucial focus in CSCL.

**Keywords:** Collaborative learning, Computer-supported collaborative learning (CSCL), Cooperative learning, learning environments.

## 1. Introduction

Computer-supported collaborative learning (CSCL) is a prominent research area in 21stcentury education. It emphasizes using computer technologies to enhance collaborative learning among learners (Hod & Sagy, 2019). Through computer-based tools like online discussion forums, wikis, and virtual environments, learners engage in collaborative activities such as problem-solving and knowledge sharing (Jung, Shin, & Zumbach, 2021; Lämsä et al., 2021). CSCL recognizes the vital role of social communication and collaboration in effective learning and leverages technology to facilitate these processes (Hernández-Sellés, Muñoz-Carril, & González-Sanmamed, 2019). Drawing from education, psychology, and computer science, CSCL is implemented in diverse educational settings, including traditional classrooms and online platforms across various education levels (Radkowitsch, Vogel, & Fischer, 2020).

Technological advancements have underscored the significance of CSCL environments in education (Wang & Hwang, 2012). This has led to new research questions and perspectives in areas such as higher education, K-12 instruction, special education, and workplace training (Lämsä et al., 2021). CSCL in programming education, for example, promotes teamwork, communication, autonomous learning, and problem-solving skills, while acknowledging

challenges like technical limitations and students' social and learning abilities (Cheng et al., 2022).

## 2. Research results

By analyzing the top CSCL articles spanning from 1995 to 2023, this study has unveiled substantial findings, encompassing a total of 1068 studies. Notably, Ertmer et al. (2012) authored the most cited articles. The United States emerged as the leading contributor in terms of the number of publications, as depicted in Figure 1. Springer Nature stood out as the most prominent publisher, as indicated in Figure 2. The International Journal of Computer-Supported Collaborative Learning featured prominently with the highest number of studies, as illustrated in Figure 3. Within the realm of authors, Frank Fischer held the distinction of being the most cited, with co-authors such as Karsten Stegmann and Armin Weinberger also making notable contributions, as highlighted in Figure 4. The University of Hong Kong emerged as the most frequently affiliated institution, alongside co-affiliated institutions like Open University Netherlands, University of Oulu, and NTUST, as demonstrated in Figure 5. Notable keywords that gained popularity included computer-supported collaboration, cooperative learning, collaborative learning, computer-mediated communication, and interactive learning environments, as showcased in Figure 6. Quantitative research methods predominated throughout the analyzed articles. The primary focus of the studies was on higher education students as the most studied participants. Notably explored were collaborative learning strategies for curriculum design, while the prevalence of tablet computers as a subject of discussion was also evident across the articles.



Figure 1. The leading countries.

454 Springer Nature	259 Elsevier	67 Wiley	
		35 IEEE	24 Sage
	111		
	Taylor & Francis	13 INT FORUM EDUCATIONAL	12 Graz Techno Inst
		13 NATL TAIWAN NORMAL	9 NATL

Figure 2. The most publisher.



Figure 3. The most published Journal.



Figure 4. The most published authors.

18 UNIVERSITY OF HONG KONG	18 UNIVERSITY OF OSLO	13 NANYANG TECHNOLOGICAL UNIVERSITY NATIONAL INSTITUTE OF EDUCATION NIE SINGAPORE	13 NATIONAL INSTITUTE OF EDUCATION NIE SINGAPORE
16 CARNEGIE MELLON UNIVERSITY	15 EBERHARD KARLS UNIVERSITY OF TUBINGEN		
16		12 HEBREW UNIVERSITY OF JERUS	ALEM
LEIBNIZ INSTITUT FUR WISSENSMEDIEN	13 NANYANG TECHNOLOGICAL UNIVERSITY	12 UNIVERSITY OF MUNICH	

Figure 5. The most published institution.



Figure 6. The most published key topics.

# 3. Discussion and conclusion

Recent CSCL research has focused on the use of learning devices, particularly tablet computers, to enhance CSCL competencies and learning outcomes. It has also emphasized leveraging computer-mediated communication and various tools for collaboration and interactive education. Educational institutions must adapt to technological advancements to equip learners with the necessary information and communication technology skills. CSCL research in education highlights the importance of collaboration, technology, and the broader impact of CSCL across different disciplines.

# Acknowledgments

This study is supported in part by the Ministry of Science and Technology of Taiwan under contract number NSTC 111-2410-H-038-029-MY2.

## References

- Cheng, Y. P., Shen, P. D., Hung, M. L., Tsai, C. W., Lin, C. H., & Hsu, L. C. (2022). Applying online content-based knowledge awareness and team learning to develop students' programming skills, reduce their anxiety, and regulate cognitive load in a cloud classroom. Universal Access in the Information Society, 1-16. https://doi.org/10.1007/s10209-020-00789-6
- Hernández-Sellés, N., Muñoz-Carril, P. C., & González-Sanmamed, M. (2019). Computersupported collaborative learning: An analysis of the relationship between interaction,
emotional support and online collaborative tools. Computers & Education, 138, 1-12. https://doi.org/10.1016/j.compedu.2019.04.012

- Hod, Y., & Sagy, O. (2019). Conceptualizing the designs of authentic computer-supported collaborative learning environments in schools. International Journal of Computer-Supported Collaborative Learning, 14, 143-164. https://doi.org/10.1007/s11412-019-09300-7
- Jung, J., Shin, Y., & Zumbach, J. (2021). The effects of pre-training types on cognitive load, collaborative knowledge construction and deep learning in a computer-supported collaborative learning environment. Interactive Learning Environments, 29(7), 1163-1175. https://doi.org/10.1080/10494820.2019.1619592
- Lämsä, J., Hämäläinen, R., Koskinen, P., Viiri, J., & Lampi, E. (2021). What do we do when we analyse the temporal aspects of computer-supported collaborative learning? A systematic literature review. Educational Research Review, 33, 100387. https://doi.org/10.1016/j.edurev.2021.100387
- Radkowitsch, A., Vogel, F., & Fischer, F. (2020). Good for learning, bad for motivation? A meta-analysis on the effects of computer-supported collaboration scripts. International Journal of Computer-Supported Collaborative Learning, 15, 5-47. https://doi.org/10.1007/s11412-020-09316-4
- Wang, S. L., & Hwang, G. J. (2012). The role of collective efficacy, cognitive quality, and task cohesion in computer-supported collaborative learning (CSCL). Computers & Education, 58(2), 679-687. https://doi.org/10.1016/j.compedu.2011.09.003

# Facilitating nursing students' critical thinking and problem-solving competence in a computer supported collaborative learning environment

Zi-Ying Jiang<sup>a</sup>,Ching-Yi Chang<sup>a\*</sup> <sup>a</sup>School of Nursing, Taipei Medical University, Taiwan

#### \*frinng.cyc@gmail.com

Abstract: In the era of globalization, teachers primarily assist students in applying realistic learning environments, helping them integrate the knowledge acquired from textbooks with clinical practical issues and develop problem-solving skills. The advancement of mobile technology allows students to access learning resources and receive accurate guidance in a virtual environment, promoting safe and effective learning. This research proposes a computer-supported collaborative learning environment to support the learning of obstetrics and gynecology four-stage palpation professional courses. A quasi-experimental design is conducted to validate the impact of this method on students' learning achievements, engagement, and satisfaction. The experimental results demonstrate that the proposed approach enhances students' learning achievement, critical thinking, and problem-solving competence. Moreover, students engaged in scenario-based learning show more proactive learning behavior compared to the control group. Based on these findings, specific recommendations for developing effective learning strategies and incorporating computer-supported collaborative learning environments in medical and healthcare educational materials are suggested.

**Keywords:** Computer-supported collaborative learning environment, computer-supported, collaborative, learning environment.

#### 1. Introduction

In the era of globalization, the purpose and educational mission of nursing curriculum training is to cultivate nursing students with core competencies in professionalism, knowledge, skills, and information abilities. Adapting to the challenges of teaching during the pandemic and accurately addressing students' learning needs in real time has become a pressing issue for all educators worldwide. To mitigate the impact of the pandemic on education, many fields in education have turned to managing digital interactive learning platforms as an alternative solution to meet learners' needs. Teachers have been forced to transform their teaching approaches in response to the pandemic, and scholars Nuangchalerm, Prachagool, and Dostál (2020) suggested incorporating more interactive educational courses through digital platforms.

Creating a digital learning environment offers several advantages, including meeting the learning needs of students regardless of time and geographical limitations, addressing interdisciplinary learning needs, and fostering opportunities for interactive learning and a sense of achievement. Modern digitized educational materials should be diverse, incorporating technology options, interactivity, and high-quality content to demonstrate the specific value of specialized fields in digital textbooks. Cultivating nursing students through digital learning materials allows for learning in a safe environment, enhancing students'

professionalism, knowledge, skills, and information literacy, ultimately improving healthcare quality, which is the current educational goal. Moreover, guiding students to prioritize patient and learner safety and facilitating repetitive learning in a secure environment can enhance learning outcomes and make digital platform materials more appealing and appreciated.

Given the demand for digital education and the pandemic's impact, it has been observed that the main issue in nursing teaching lies in the lack of effective interactive teaching materials that allow students to learn anywhere and anytime (Geng, Huang, & Huang, 2021). Analyzing relevant empirical research literature both domestically and internationally, it is worth noting that apart from imparting professional knowledge, meeting learners' needs and using digital learning platforms to teach students to practice and master nursing basic skills through interactive learning strategies can foster students' engagement and improve learning outcomes, making it a crucial topic for teachers to continue researching and addressing (Supriatin, Rithpho, Asiah, & Hikhmat, 2022).

Nevertheless, the amalgamation of interactive learning strategies into education for obstetric assisting nursing skills, commonly referred to as the four-stage palpation education, remains an area that has garnered only minimal research attention. Thus, this study proposes a learning activity that applies the digital learning platform in combination with the computer-supported collaborative learning environment learning model to create an interactive learning environment, incorporating the situational learning theory. Additionally, Salminen et al. (2021) mentioned that the educational goals of the nursing curriculum include cultivating students' professional knowledge, skills, and attitudes. To evaluate the effectiveness of the proposed method, the computer-supported collaborative learning environment learning environment learning system is applied to the obstetrics four-stage palpation education course, and its impact on students' learning achievements, critical thinking, and problem-solving competence is examined. Therefore, the following research questions are proposed to verify and assess the effectiveness of the method:

(1) Does the computer-supported collaborative learning environment learning method significantly improve students' learning achievement compared to traditional educational methods?

(2) Does the computer-supported collaborative learning environment learning method enhance students' critical thinking more effectively than traditional teaching methods?(3) Does the computer-supported collaborative learning environment learning method improve students' problem-solving competence more effectively than traditional teaching methods?

# 2. Method

# 2.1 Experimental design

The purpose of this course is to train students in the basic skills of learning obstetric professional knowledge through computer-supported collaborative learning environment and scenario-based learning for the four-stage palpation. To validate the effectiveness of this method, a quasi-experimental design is employed to assess students' learning outcomes, critical thinking, and problem-solving abilities through questionnaires.

# 2.2 Participants

The participants of this experiment were second-year students from a specialized nursing school. Two classes, each consisting of 20 students, were used as the experimental group and the control group, respectively. The same teacher, who had 20 years of experience in obstetric clinical practice and teaching, conducted the instruction. To prevent potential mutual influence between students from different classes, the two classes received the instruction at different time points. During the learning process, the teacher provided obstetric learning tasks for the four-stage palpation, supplemented with teaching materials and assessment items. Additionally, the teacher addressed any relevant questions about the learning tasks raised by the students during the instructional sessions.

#### 2.3 Experimental procedure

The experimental design of this study lasted for five weeks. In the first week, the students received instructions on basic knowledge related to maternal and infant care. Subsequently, they underwent pre-tests and questionnaire surveys. In the second week, the experimental group received guidance on the four-stage palpation, utilizing computer-supported collaborative learning environment to facilitate scenario-based learning and understanding the objectives of the learning activities. On the other hand, the control group used traditional video learning for the relevant knowledge on the four-stage palpation. All participating students completed post-tests and questionnaires on their learning outcomes, critical thinking, and problem-solving abilities in the fifth week to validate their learning effectiveness, as shown in Figure 1.



#### 2.4 Measuring tools

The measurement tools used in this study included learning achievement tests and questionnaires to assess the learning status. The test questionnaires were designed by the obstetric course teacher and their validity was evaluated by two obstetric education experts. These tests aimed to understand the learning outcomes of the two groups of students. The pre-test was conducted to assess students' prior knowledge of the four-stage palpation in obstetrics before the learning activities. It consisted of 20 multiple-choice questions, with a maximum score of 100. The post-test included obstetric assessment skill tests derived from obstetric assessment performance proposed by two experienced teachers. It comprised 10 dimensions.

The Critical Thinking Scale was developed by Lai and Hwang (2014) and consisted of 6 items, using a Likert 5-point rating scale, with a Cronbach's alpha value of 0.80. One of the items was "I like to ask some questions that others do not think of."

The Problem-Solving Competence Scale, also proposed by Lai and Hwang (2014), consisted of 6 items, using a Likert 5-point rating scale, with a Cronbach's alpha value of 0.81. One of the items was "I believe that I have the ability to solve the problems I encounter," and "I believe that I can solve problems on my own."

#### 3. Results

#### 3.1. Learning achievement

The ACONVA analysis was used to compare the learning achievements of the experimental group and the control group, and to test the homogeneity assumption of ACONVA. The results showed that the pre-test F = 5.02 (p > 0.05) and the post-test F = 0.21 (p > 0.05), indicating

that ACONVA could be used to compare the learning achievements of the two groups before and after the intervention. The ACONVA analysis results also indicated that there was no significant difference in prior knowledge between the two groups, with F = 0.48 (p = 0.52).

Furthermore, there was a significant difference in the post-test scores between the two groups (F = 3.57, p < 0.05). The experimental group had an average score of 86.08 (standard error = 0.95), which was higher than the control group's average score of 80.80 (standard error = 0.95). This indicates that the learning method based on computer supported collaborative learning environment in the experimental group significantly improved students' learning achievement compared to the traditional learning method in the control group.

#### 3.2. Critical thinking tendency

To validate the rationality of ANCOVA analysis between the experimental and control groups' critical thinking tendency, the homogeneity of regression coefficients within each group was examined. The result showed that F = 0.58 did not violate the homogeneity assumption (p > 0.05). Subsequently, ANCOVA was used to analyze the variables, with the post-test questionnaire scores as the dependent variable and the pre-test questionnaire scores as the covariate. Table 2 presents the ANCOVA results for the critical thinking tendency after the questionnaire survey in both groups.

After controlling for the influence of pre-test scores, there was a significant difference in the post-test scores between the two groups (F = 3.35, p < 0.05). The adjusted mean for the experimental group was 4.69 (standard error = 0.69), which was higher than the control group's mean of 4.66 (standard error = 0.69). These results indicate that students in the experimental group, who adopted the learning method based on computer supported collaborative learning environment, demonstrated a significant improvement in critical thinking tendency.

# 3.3. Problem-solving tendency

This study utilized a pre-questionnaire to measure whether the experimental and control groups had similar experiences in the classroom. The mean and standard deviation of the pretest scores for the experimental group were 4.64 and 0.32, respectively, while the mean and standard deviation of the pre-test scores for the control group were 4.63 and 0.37, respectively. The results of the t-test indicated that there was no significant difference between the two groups in the pre-test scores (t = 1.03, p > 0.05). Therefore, it can be inferred that both groups had similar learning experiences.

After the learning activities concluded, a problem-solving competence questionnaire survey was conducted to measure the learning satisfaction of both groups. As shown in Table 3, the t-test results revealed that the problem-solving competence scores in the experimental group were higher than those in the control group (t = 2.78, p < 0.05). Based on these results, it can be concluded that students using the learning method based on computer supported collaborative learning environment demonstrated an effective improvement in their problem-solving tendency.

# 4. Discussion and conclusion

Within the context of this investigation, with the aim of elevating the proficiency of nursing students in the domain of obstetric nursing skills, we introduced an innovative computer-supported collaborative learning system. This pioneering approach was subsequently subjected to a quasi-experimental study conducted within a specialized obstetric course, serving as the foundation for an in-depth assessment of the efficacy of this novel pedagogical technique. The empirical outcomes of this study unequivocally demonstrated that the utilization of the computer-supported collaborative context learning method resulted in a noteworthy enhancement of students' learning accomplishments, heightened critical thinking

capacities, and an augmented inclination toward effective problem-solving, all of which stood in marked contrast to the outcomes associated with traditional learning methodologies.

We integrated the clinical learning environment of prenatal care into the computer-supported collaborative learning environment system and designed obstetric context learning, highlighting the key learning activities in the clinical setting. Students in the experimental group were able to practice the learning tasks repeatedly and build essential knowledge and operational skills through the computer-supported collaborative learning environment-based context learning system. As a result, the students in the experimental group demonstrated more concrete nursing concepts and better interaction with the teacher during the learning process. In contrast, students in the control group, who had not experienced obstetric context learning, found the context and operational skills more abstract in the classroom. Therefore, the results indicated that the students who used the computer-supported collaborative learning environment-based context learning method achieved better learning outcomes, learning participation, and overall satisfaction with the teaching.

Based on the findings, we suggest that future research on computer-supported collaborative learning environment-based context learning should incorporate not only context-specific questions, tests, or concepts but also interactive and engaging guidance. These questions should include essential concepts and learning units related to the course. Additionally, educators can present core questions related to the learning units for further exploration and discussion in the classroom.

Nevertheless, this study has some limitations. Firstly, it employed the computer-supported collaborative learning environment-based context learning technology, and its generalizability may be affected by the actual cases in the specialized field. Secondly, the experiment was conducted in a specific school and for a particular course, which may limit the generalizability of the research results to students with different professional backgrounds or in other medical fields. Thirdly, the sample size of this study was limited, and the statistical differences in the results may not be sufficient for inference. Future research could consider implementing additional learning behavior assessments or learning strategies in the computer-supported collaborative learning environment context learning and conducting interviews to understand the impact on student's performance and learning experiences.

In conclusion, the primary contribution of this study is the proposal of the computer-supported collaborative learning environment-based context learning method and its integration with the context-promoting mechanism in the computer-supported collaborative learning environment learning system, which innovatively transforms traditional obstetric teaching design. The experimental results further demonstrate that the learning strategy of guiding students to face contextual problems and seek answers according to the context is applicable to other medical and health education fields, educators, or researchers to promote teaching objectives in various clinical contexts and develop innovative digital teaching materials tailored to different professional medical and health education contexts.

#### Acknowledgements

This study is supported in part by the Ministry of Science and Technology of Taiwan under contract number NSTC 111-2410-H-038-029-MY2.

# References

- Geng, Y., Huang, P. S., & Huang, Y. M. (2021). Crowdsourcing in nursing education: A possibility of creating a personalized online learning environment for student nurses in the post-COVID Era. *Sustainability*, *13*(6), 3413.
- Lai, C. L., & Hwang, G. J. (2014). Effects of mobile learning time on students' conception of collaboration, communication, complex problem-solving, meta-cognitive awareness and creativity. *International Journal of Mobile Learning and Organisation*, *8*(3), 276-291.

- Nuangchalerm, P., Prachagool, V., & Dostál, J. (2020). Digital learning of pre-service teachers during COVID-19 outbreak. *Journal of Technology and Information Education*, *12*(2), 143-151.
- Salminen, L., Tuukkanen, M., Clever, K., Fuster, P., Kelly, M., Kielé, V., ... & PROCOMPNurse-Consortium. (2021). The competence of nurse educators and graduating nurse students. *Nurse education today, 98,* 104769.
- Supriatin, S., Rithpho, P., Asiah, A., & Hikhmat, R. (2022). Blended learning to improve the physical examination ability of nursing students. *International Journal of Educational Qualitative Quantitative Research*, *1*(2), 23-30.

# Motivating Knowledge Seeking and Sharing Behavior in the Online Discussion Forum

Yue HU<sup>a</sup>, Baoping Ll<sup>a,b\*</sup>, Xiang ZHANG<sup>b</sup> & Ling CHEN<sup>a</sup>

<sup>a</sup>Faculty of Education, Beijing Normal University, China <sup>b</sup>Advanced Innovation Center of Future Education, Beijing Normal University, China \*libp@bnu.edu.cn

**Abstract:** This study developed an online discussion forum that motivate students' knowledge seeking and sharing. The forum enables students post questions anonymously and provides a question or response prompt that invite students to engage in the discussion after completing a quiz. A quasi-experimental approach was conducted with a sample of 100 students. The results indicated that the students in the experimental group exhibited more active participation in knowledge seeking and sharing, as well as showed a higher density of group's social network. The findings provide some evidence for the value of the motivational design, in terms of anonymous questioning and prompt features, to address the challenge of under-contribution in online discussion forums.

**Keywords:** Knowledge seeking, knowledge sharing, motivation, online discussion forum

#### 1. Introduction

With the advancement of information and communication technology, the Internet now provides students new channels such as email, search engines to seek academic help online. Online discussion forums have already demonstrated their potential to facilitate collaborative problem-solving among students. However, the effectiveness of these forums is intrinsically tied to how actively students participate and how deeply they engage with one another. Additionally, previous research has shown that only a minority of students would like to post their questions or to respond to others' queries (Chao et al., 2018). As a result, it is essential to create a supportive environment for online discussions in order to motivate student engagement in actively seeking and sharing knowledge.

# 2. Literature Review

#### 2.1 Knowledge Seeking and Sharing Behavior in the Online Discussion forum

Online discussion forums provide a place where students and teachers can have discussions regarding course materials and assignments online, regardless of time and place (Demir et al., 2023). In these forums, students interact with peers and teachers to seek and share knowledge. Knowledge seeking behaviors involve looking for information, help, or advice in response to a need. It's like learning from what others have experienced. Knowledge sharing behaviors include offering information and/or knowledge about personal experiences and making it available to others (Kumi & Sabherwal, 2019). Active participation and engagement in these discussions are crucial for facilitating effective knowledge exchange (Arazy et al., 2016). However, the well-known dilemma of the "tragedy of commons" manifests in online discussion forums, where many individuals tend to free ride or under-contribute rather than take the time to continually share their knowledge (Chen et al., 2019), leading to a decrease in participants posting questions. Therefore, motivating students to participate in and

contribute to online discussions for the purpose of seeking and sharing knowledge is challenging.

# 2.2 Motivation to Seek and Share Knowledge

Motivation refers to those methods that attribute efforts, direction, or purpose to the requisite behaviors, and plays an important role in the knowledge seeking and sharing process (Ergün & Avcı, 2018). In such situation, the self-determination theory differentiates two primary types of motivations: intrinsic motivation, driven by inherent satisfaction and enjoyment, and extrinsic motivation, characterized by the inclination to perform for external rewards.

Previous studies have explored the internal and external motivating factors that support individuals' engagement in knowledge seeking and sharing in various contexts. For example, Kwahk and Park (2016) underscored the positive correlation between knowledge self-efficacy and knowledge sharing, as well as continued sharing. Individuals with high knowledge selfefficacy demonstrate a heightened sense of responsibility, motivating them contribute highquality contents to the online discussions (Zhang et al., 2019). Therefore, considering interventions that invite individuals with high knowledge self-efficacy, such as students with extensive experience or a profound grasp of the subject to engage in knowledge-seeking processes, may enhance the value of knowledge exchange and improve students' willingness to actively participate in knowledge acquisition in an online discussion forum.

Additionally, reputation serves as a significant motivator for active engagement in online discussions. Karabenick and Dembo (2011) proposed that the absence or avoidance of knowledge seeking may not be directly related to the degree of help a student may need. Instead, when students refrain from seeking help, it may mean that they are embarrassed to seek academic help (Koc & Liu, 2016). Hence, it is recommended that the fundamental design of any online discussion forum should consider concealing the questioners' identities to safeguard self-esteem and thereby enhance students' participation in knowledge seeking.

Based on the aforementioned concerns, the design of an online discussion forum system for enhancing knowledge seeking and sharing should consider the motivational factors that encourage meaningful interaction among students. To address this need, the current study modified an online discussion forum to enable students to post questions anonymously and actively engage in online discussions. Furthermore, this study investigates the impact of the modified online discussion forum on students' engagement in both knowledge seeking and sharing activities. The research questions in this study are as follows: (1) What impact does the modified online discussion forum have on students' engagement in knowledge seeking and sharing? (2) What impact does the online discussion forum have on students' social network?

#### 3. Methods

#### 3.1 Participants

The study included 100 graduate students (24 males and 76 females) enrolled in the C programming course at a university in Beijing. All students are all non-computer majors, including psychology, physics, and other majors. Each participant was assigned to either the experimental or control group at random.

#### 3.2 Online Discussion Forum

Daily Quiz is a practice system that offers a sequence of multiple-choice questions and supports distributed practice for novice programming (Li et al., 2021). This system enables students to take quizzes, view correct answers, as well as read the related materials. Below each daily quiz is an online discussion area where students can post their questions or provide responses to others' queries. As a question is posted, the system generates timely

notifications to inform the students of any comments from peers or teachers, facilitating rapid checking and response to the received messages. Furthermore, they can also express their perspectives on the replies received using the 'like' feature.

Daily Quiz provides two discussion forum-related features exclusively for the experimental group. The first feature is the anonymous questioning, which allows students to post questions without revealing their identities (see Figure 1.). The second is a question or response prompt that invite students to engage in the discussion. For example, after successfully completing a quiz, the system creates a prompt encouraging participation in the discussion forum, along with a message like "Congratulations! Please enter the discussion forum to help the 5 students who are experiencing trouble with this quiz."



*Figure 1.* The Online Discussion Forum for Experimental Group

# 3.3 Procedure

The study lasted for eight weeks. In the first week, the teacher introduced students to the Daily Quiz and provided them a video tutorial outlining the basic features of the system. Students were encouraged to use the discussion forum to post their questions or provide comments to the questioners. In the remaining weeks, the teacher continued teaching the C programming course, and students used Daily Quiz for practice and engaging in interaction during their available spare time. Notably, students in the experimental group would remain anonymous when they ask questions and receive the prompt after completing a quiz. In contrast, the functions of anonymous questioning feature and prompts were turned off for the control group within Daily Quiz.

# 3.4 Data Collection and Analysis

During the course, the two groups of students were free to use the features of Daily Quiz to do quizzes and resolve their questions in a convenient location and at a convenient time. All students' activities of taking quizzes, posting questions, making comments on questions, giving likes were collected. To address the research questions raised in this study, the frequency of these activities was computed to compare the differences between the experimental group and the control group.

Social network could represent social connections between people and reflect the qualitative aspect of participation in collective activities. Social Network Analysis (SNA) is a valuable method for revealing relationship patterns among members of social systems. Various studies have indicated that SNA can map interactions in text-based asynchronous online discussions in an effective and analyzable manner (Norz et al., 2023). For example, Lim (2023) used SNA to investigate the features and patterns of student interactions on online discussion boards during a course and found the critical role of social presence in enhancing students' online learning experiences and outcomes. Through SNA, we could gain insight into students' activity levels and participation patterns participation patterns in online discussions. Thus, SNA was employed to explore the degree of student participation in interaction in this study. The network data which comprise a set of nodes and ties was derived from the interactions among students through exchanged postings. SNA was then created with UCINET 6.0 software and visualized with Netdraw (Borgatti et al., 2002).

# 4. Results

# 4.1 The Differences in Knowledge Seeking and Sharing Behavior

A series of independent sample t-tests was used to examine the difference between the experimental and control groups in terms of students' behavior in the online discussion forum, as shown in Table 1. The results revealed significant differences in the participation of asking questions, answering, and liking responses. Specifically, students in the experimental group had significantly higher frequencies than those in the control group for posting questions (t = 6.585, p < 0.001). In terms of answering questions, the experimental group made significantly more comments on others' questions (t = 9.969, p < 0.001) than the control group did. Moreover, there was significant differences in giving likes on the responses with the experimental group had higher frequencies of liking than the control group (t = 4,362, p < 0.001). These results indicate that the modified online discussion forum could potentially facilitate students' knowledge seeking and sharing activities as well as motivate them to interact with others.

			1
Variables	Experimental group	Control group	t
	(N=52)	(N=48)	
Questioning	4.44±5.09	0.00±0.00	6.585***
Answering	23.79±17.33	0.67±0.58	9.969***
Liking	4.96±7.86	0.15±0.55	4.362***

Table 1. Comparison of Students' Behavior of the Experimental and Control Groups

Note. \*\*\**p*<0.001, \*\**p*<0.01, \**p*<0.05.

# 4.2 The Differences in Social Network

SNA was employed in this study to explore the social connections within both experimental and control groups. Figure 2. illustrates the social network of the control group, which has just 3 nodes (students) and 2 ties (connections between students), showing that there is little interaction among them. As shown in Figure 3., the social network of the experimental group exhibits a higher density and increased connectivity. A significant majority of students in the experimental group actively engaged in the online discussions and responded to peers' questions. These findings demonstrate that the value of the design strategy that prompt students to enter the discussion forum to post or answer questions. This strategy not only facilitate their engagement in meaningful conversations but also serve as a motivation for their knowledge seeking and sharing behaviors.



Figure 2. The Social Network of the Control Group



Figure 3. The Social Network of the Experimental Group

# 5. Conclusion

This study introduced an online discussion forum to motivate students' active engagement in knowledge seeking and sharing activities. The effectiveness of this forum was further explored by analyzing students' participation in knowledge seeking and sharing, as well as the social network of student groups. The results indicated that the students in the experimental group have significant higher frequencies than the those in the control group for questioning, answering and giving positive feedback on those valuable comments. Knowledge construction mainly happens in reading and commenting on others' posts by internalizing and integrating external useful information from others(Ye & Pennisi, 2022). The active participation of students in the experimental group in the discussion forum may help them achieve better knowledge construction. In terms of the social network, the experimental group exhibited a closer relationship among students than the control group. The overall results revealed the efficacy of the modified online discussion forum in promoting active participation and knowledge exchange. The anonymous questioning feature in the online discussion forum fostered a comfortable environment where students felt at ease posting their queries publicly. Moreover, the prompt inviting students to participate in the discussions instilled a sense of responsibility among those proficient in the guiz content, motivating their help to peers in need. Consequently, this study provides innovative approach to address the challenge of undercontribution in online discussion forums from the motivational perspective.

This study has some limitations that offer opportunities for future research. Specifically, the current study explored the combined effects of two features. Future research could adopt a segmentation approach, segregating participants into distinct groups to investigate specific impacts on knowledge seeking and sharing. Furthermore, it should be noted that a minority of students displayed limited engagement in discussions, with some forming only minimal ties (2 or 3) in the social network of the experimental group. Future studies can explore strategies more adept at stimulating students' participation and engagement in knowledge exchange in the online discussion forums.

# Acknowledgements

The study is supported by the Chinese National Social Science Foundation (BCA220209).

#### References

- Arazy, O., Gellatly, I., Brainin, E., & Nov, O. (2016). Motivation to share knowledge using wiki technology and the moderating effect of role perceptions. *Journal of the Association for Information Science and Technology*, 67(10), 2362–2378. https://doi.org/10.1002/asi.23579
- Borgatti, S., Everett, M., & Freeman, L. (2002). UCINET for Windows: Software for social network analysis. Harvard, MA: Analytic Technologies.
- Chao, P.-Y., Lai, K. R., Liu, C.-C., & Lin, H.-M. (2018). Strengthening social networks in online discussion forums to facilitate help seeking for solving problems. *Journal of Educational Technology & Society*, *21*(4), 39–50.
- Chen, L., Baird, A., & Straub, D. (2019). Why do participants continue to contribute? Evaluation of usefulness voting and commenting motivational affordances within an online knowledge community. *Decision Support Systems*, *118*, 21–32. https://doi.org/10.1016/j.dss.2018.12.008
- Demir, Ö., Cinar, M., & Keskin, S. (2023). Participation style and social anxiety as predictors of active participation in asynchronous discussion forums and academic achievement. *Education and Information Technologies*. https://doi.org/10.1007/s10639-022-11517-3
- Ergün, E., & Avcı, Ü. (2018). Knowledge sharing self-efficacy, motivation and sense of community as predictors of knowledge receiving and giving behaviors. *Journal of Educational Technology & Society*, *21*(3), 60–73.
- Karabenick, S. A., & Dembo, M. H. (2011). Understanding and facilitating self-regulated help seeking. *New Directions for Teaching and Learning*, *126*, 33–43. https://doi.org/10.1002/tl.442
- Koc, S., & Liu, X. (2016). An investigation of graduate students' help-seeking experiences, preferences and attitudes in online learning. *The Turkish Online Journal of Educational Technology*, *15*(3).
- Kumi, R., & Sabherwal, R. (2019). Knowledge sharing behavior in online discussion communities: Examining behavior motivation from social and individual perspectives. *Knowledge and Process Management*, 26(2), 110–122. https://doi.org/10.1002/kpm.1574
- Kwahk, K.-Y., & Park, D.-H. (2016). The effects of network sharing on knowledge-sharing activities and job performance in enterprise social media environments. *Computers in Human Behavior*, 55, 826– 839. https://doi.org/10.1016/j.chb.2015.09.044
- Li, B., Ning, F., Zhang, L., Yang, B., & Zhang, L. (2021). Evaluation of a practice system supporting distributed practice for novice programming students. *Journal of Pacific Rim Psychology*, *15*, 18344909211008264. https://doi.org/10.1177/18344909211008264
- Lim, J. (2023). Exploring the relationships between interaction measures and learning outcomes through social network analysis: The mediating role of social presence. *International Journal of Educational Technology in Higher Education*, *20*(1), 14. https://doi.org/10.1186/s41239-023-00384-8
- Ye, D., & Pennisi, S. (2022). Analysing interactions in online discussions through social network analysis. *Journal of Computer Assisted Learning*, *38*(3), 784–796. https://doi.org/10.1111/jcal.12648
- Zhang, Y., Zhang, M., Luo, N., Wang, Y., & Niu, T. (2019). Understanding the formation mechanism of high-quality knowledge in social question and answer communities: A knowledge co-creation perspective. *International Journal of Information Management*, 48, 72–84. https://doi.org/10.1016/j.ijinfomgt.2019.01.022

# Personalized Hands-on Training Via a Hybrid Intelligent Teacher System

#### Dongkun HAN<sup>a\*</sup>

<sup>a</sup>Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China \*dkhan@mae.cuhk.edu.hk

**Abstract:** In this paper, we propose a student-centered hybrid intelligent teacher system, and show how this system advances in a case study in teaching robotics. The key innovation of this system is to focus on the combination of a human teacher and a cloud teacher, which is a textual-based conversational agent for answering the questions from students in Mechanical Engineering via machine learning technologies. The basic underlying idea is to train the agent by utilizing open-source artificial intelligence tools from Google's DeepMind, such that the agent can understand and answer the questions raised by students with an acceptable confidence value. Specifically, it is expected to customize a training database combining VEX hardware assembling, Auduino and Solidworks programming. Moreover, Big Data Analytics can be conducted accordingly based on students' historical questions. An Android-based mobile application is produced. As a result, students can raise their questions conveniently and get instantaneous feedback with a 24-7 service. The method has demonstrated its effectiveness in a university general education course (UGEB2303 Robots in Action).

**Keywords:** Personalized education, Chatbot development, Big data analytics, Robotics teaching

#### 1. Introduction

Nowadays, technology in education is playing a growingly essential role in teaching and learning. Many studies have demonstrated that machines perform better in modern education than human teachers in certain areas, like automatic assessment, functional chatting services, and big data analysis (Selwyn, N., 2019). Technology in education is expected to bring fundamental changes to the relationship between students and teachers, and to promote learning outputs.

Still, it is important to note that technology is a tool used in education and not an end in itself. The promise of educational technology lies in what educators do with it and how it is used to best support their students' needs. Although machines surpass human teachers in some specific tasks, they need to be carefully placed by their designer (Wajcman, J., 2017). Thus, an optimal teaching and learning system relies on an elegant collaboration of technologies, their designer, and the executor.

# 2. Hybrid Intelligent Teacher System

#### 2.1 System Description and Main Component

As a teacher of engineering, I firmly believe in a superlative combination of teacher and technology, which enables students to take a more active role in learning, interacting with their peers, and creatively making their contributions to society. Based on the Johnson's Technology Integration Model (Johnson, D. L., and Liu, L., 2000) and the Technology Team-Teaching Model (Jong, J. S., 2008), a student-centered hybrid teacher system is innovatively developed with students as the centre (see Figure 1)

In the proposed system, the following features are worth clarifying:

- **Human teacher**: A real person helps students acquire knowledge and competence effectively and efficiently. S/he could develop or apply various educational technologies and devices to accelerate students' learning. The role of a teacher can go beyond teaching, such as being a counsellor, mentor, role model, and so on. His/Her basic functions include: 1) Design and manage the cloud teacher; 2) Maintain a desirable relationship between students and the cloud teacher; 3) Breed curiosity and cultivate creativity; 4) Provide nuanced feedback; 5) Expedite students' access to innovations, knowledge, experience, peers, and community; and 6) Supply humanistic care such as empathy, trust, inspiration, and encouragement.
- Cloud teacher: The instrumental element or educational technologies could be any format of carriers that facilitate teaching and learning, bringing cause-effect results through its implementations in the system. Basic features of the cloud teacher include:
   1) 24/7 functional Q&A service to students; 2) Big data analysis for individualized education (hence, precision education) (Zhang, K., and Ayse, B. A., 2021); 3) Accessible and online experiential training; 4) Automatic assessment and invigilation; 5) Intelligent teaching assistance based on robots and Al; and 6) Related software like APPs, websites, micro-modules, or integration with hardware.
- **Collaboration** between human teacher and cloud teacher in a hybrid mode: 1) Human teacher designs and manages cloud teacher; 2) Cloud teacher gives human teacher a better way of teaching and learning in various aspects, like automatic and formative assessments; 3) Cloud teacher is able to establish students' profiling and helps human teacher in making predictions and course preparation; 4) Cloud teacher analyzes students' on-the-spot inputs, and provides instant feedback to human teacher, who can give timely interventions to students; 5) Cloud teacher saves time and energy of human teacher such that s/he can perform more effectively; 6) Human teacher fine-tunes cloud teacher based on user experience and feedbacks.



Figure 1: A diagram of the student-centered hybrid intelligent teacher system

- Creativity: The capacity of students to produce original ideas or objects. The created item may be intangible (such as an idea, a scientific theory, a musical composition, or a joke) or a physical object (such as an invention, a printed literary work, or a painting).
- **Experience**: Personal reflection or skills students learn by doing. Acquisition of experience entails a hands-on approach to learning. It makes learning an experience that moves beyond the classroom and strives to bring a more involved way of learning (Vosniadou, S., and Brewer, W. F., 1987).

- **Knowledge**: It is the combination of data and information, in addition to expert opinion, skills, and apprehension, which results in a valuable asset that aids decision making.
- **Peers**: A social group of people with similar background, social status and learning objectives. Members of the group are likely to influence each other's beliefs, behavior and learning efficiency (Scardamalia, M., and Bereiter, C., 1994).
- **Community:** A group of people who share common academic goals, and collaborate on learning materials. In a learning community, the goal is to advance collective knowledge, and meanwhile, support the growth of individual knowledge (Martin Nunezi, J. L., and Lantada, A. D., 2020).

# 2.2 The goal of Hybrid Intelligent Teacher System

The main idea of this system is to take advantages of both human teacher and cloud teacher to maximize students' learning outcomes. Our efforts are devoted to achieving the following objectives:

- to boost students' self-motivation, engagement, curiosity and passion for learning;
- to bridge the gap amongst knowledge, creativity, and real implementations;
- to establish an ecosystem for multi-dimensional theoretical and experiential peer learning with individual-to-individual, individual-to-group, group-to-group learning;
- to enhance learning outcomes and train passionate professionals who can contribute to the community and society.

The effectiveness of this system will be illustrated by a real case study in teaching robotics in the following section.

# 3. Individualized Hands-on Training in Computer Programming

### 3.1 Chatbot Development for Arduino Programming in Teaching Robotics

We have built a "cloud teacher", which is a *textual-based conversational agent* for answering questions from students in Mechanical Engineering via machine learning technologies. The main idea is to train the agent by utilizing open-source artificial intelligence tools from Google's DeepMind, such that the agent can understand and answer questions raised by students with an acceptable confidence value (Figure 2). In addition, the answers with associated confidence value will be reviewed by a panel of teachers in Mechanical Engineering while further training on the agent could be launched based on review results. Big Data Analytics can be conducted accordingly based on students' historical questions, and prescribed teaching can be provided. Specifically, different reading references and exercise quiz questions will be provided according to students' historical questions. In this way, personalized hands-on training materials are obtained such that each student could learn the computer language Arduino in a suitable way.

We combined the cloud teacher with the flipped online laboratory in teaching robotics (General Education Course UGEB2303 Robots in Action in the academic year 2020-2021, 2021-2022, and 2021-2023, 30 students per a semester).



Figure 2: Front-end and flow editing of "cloud teacher" in answering different questions for Arduino programming

# 3.2 Impact on Teaching and Learning

To show the positive influence on the hands-on learning, we have conducted the following methods as evaluation plan:

- a) Survey on the online laboratory learning experience towards the end of the course UGEB2303.
- b) Survey on the user-experience on the conversational intelligent agent towards the end of the course UGEB2303.
- c) Focus group interview with a small group of volunteer students of the courses UGEB2303.
- d) Feedbacks and discussions from course website of UGEB2303 and small group forum.
- e) Weekly reflection meetings with technicians and student helpers to monitor the progress and propose future developments and improvements.
- f) Presentation of the project and summarized feedbacks at seminars and conferences.

The monitoring data have been collected by the evaluation methods above and illustrates the usefulness of this work:

- a) The results of feedback from students have been collected. Rate of positive feedback is 83.3% (25/30) in the Survey on the online laboratory learning experience for the course UGEB2303.
- b) Rate of positive feedback is 77.4% in the Survey on the user-experience with the conversational intelligent agent towards the end of the courses UGEB2303.

The following table shows the elevated learning outcome of UGEB2303 by using the proposed method. It provides 21.2%, 33.8%, and 11% boost of performance in Quiz1, Quiz2 and Final project.

	Quiz1	Quiz2	Final Project
UGEB2303 in 2019-20 without the hybrid teacher system	9.72/15	9.13/15	34.79/50
UGEB2303 in 2022-23 with the hybrid teacher system	11.78/15	12.22/15	38.62/50

Table 1.	Impact on	the Learning	Outcomes in	<b>UGEB2303</b>	Robots in Actions
		· · · · J			

#### 3.3 Influence and Sustainability of this Work

The ability of programming with C, MATLAB and Arduino is assumed and required in a wide range of courses across different engineering programmes. The developed mobile app can be used directly in these courses such that the course instructors do not need to review or re-teach basic computer programming. The general approach of AI-powered conversational agent can also be explored in other subject areas across different faculties for sustaining effective virtual teaching and learning in the new normal.

- a) Impact in the Department of Mechanical and Automation Engineering: Besides the course UGEB2303 that already adopted the cloud teacher, more colleagues and courses will try the innovative pedagogy in the coming semester, including some courses in the same department: MAEG3080 (expected 30 students) and EEEN2020 (expected 40 students).
- b) Impact in the Faculty of Engineering: Some Faculty foundation courses with computer language programming training would adopt this pedagogy, like programming with C, MATLAB, SolidWork, etc. ENGG1110 Problem Solving By Programming will be one of the first Faculty foundation course adopt this method in the following semester.

- c) Local impact in Hong Kong: The Center for Artificial intelligence Research (CAiRe), Hong Kong University of Science and Technology, displayed interests in implementation of the developed cloud teacher in some courses for possible collaboration.
- d) Impact worldwide: Based on the developed cloud teacher, we further collaborate with Cardiff Metropolitan University and made a proposal together called "AI-assisted Automated assessment system for remote robot laboratory platform".
- e) Prize received: Poster Silver Award Pedagogical Innovation in CUHK Teaching and Learning Innovation Expo.

# 4. Conclusion

A student-centered hybrid intelligent teacher system is proposed and illustrated in this paper. It shows how this system displays its advantages in teaching robotics. The key innovation of this system is to fuse a human teacher and a cloud teacher, which is a textual-based conversational agent for answering the questions from students in *Mechanical Engineering* via machine learning technologies. The basic underlying idea is to train the agent by utilizing open-source artificial intelligence tools from Google's DeepMind, such that the agent can understand and answer the questions raised by students with an acceptable confidence value. Specifically, it is expected to customize a training database combining VEX hardware assembling, Auduino and Solidworks programming. Moreover, *Big Data Analytics* can be conducted accordingly based on students' historical questions. An Android-based mobile application is produced. As a result, students can raise their questions conveniently and get instantaneous feedback with a 24-7 service. The effectiveness of this method is given by a case study in a university general education course.

# Acknowledgements

We would like to appreciate the time and efforts of editor and reviewers on this manuscript.

# References

- Selwyn, N., (2019), "Should robots replace teachers?: AI and the future of education. " John Wiley & Sons.
- Wajcman, J., (2017), "Automation: Is it really different this time?." *The British Journal of Sociology* 68(1), 119-127.
- Johnson, D. L., and Liu, L., (2000), "First steps toward a statistically generated information technology integration model". Computers in the Schools, 16(2), 3-12.
- Jong, J. S., (2008), "Innovations in science teacher education: Effects of integrating technology and team-teaching strategies." *Computers & Education*, 51, 646–659.
- Zhang, K., and Ayse, B. A., (2021), "Al technologies for education: Recent research & future directions." *Computers and Education: Artificial Intelligence*, 100025.
- Vosniadou, S., and Brewer, W. F., (1987), "Theories of knowledge restructuring in development." *Review of Educational Research*, 57(1), 51-67.
- Blumenfeld, P. C., Marx, R. W., Soloway, E., and Krajcik, J., (1996), "Learning with peers: From small group cooperation to collaborative communities." *Educational Researcher*, 25(8), 37-39.
- Scardamalia, M., and Bereiter, C., (1994), "Computer support for knowledge-building communities." *Journal of the Learning Sciences*, 3(3), 265–283.
- Martin Nunezi, J. L., and Lantada, A. D., (2020), "Artificial Intelligence Aided Engineering Education: State of the Art, Potentials and Challenges." *International Journal of Engineering Education*, 36(6), 1740–1751.

# Analysis of Students' Action Decisions about SDGs Issues in Open Inquiry Activity with Real Open Data

#### Yu-Hao LU\*, Chia-Chun TSENG, Ju-Ling SHIH

Graduate Institute of Network Learning Technology, National Central University, Taiwan \*yuhao.lu@g.ncu.edu.tw

**Abstract:** The inquiry-based learning model is an instructional approach tailored to students' exploration of both the natural world and society. By integrating prominent Sustainable Development Goals (SDGs) and leveraging open data, this approach nurtures students' scientific literacy and data analysis proficiencies. In this context, the current study devised an online inquiry learning system named <City Auncel>, enabling students to engage with data and independently delve into socioscientific issues through an open inquiry learning model. The findings underscore that students exhibited a high level of immersion throughout the activity, along with adept data collection and summarization skills.

**Keywords:** Open-ended inquiry, Inquiry-based learning, Real open data, Action Decisions, Sustainable Development Goals (SDGs)

### 1. Introduction

<City Auncel> is a web system for inquiry-based learning embedded with real open data, which promotes inquiry learning in students' action decisions towards the Sustainable Development Goals (SDGs). Bangura and Ally (2020) highlighted that online learning has the potential to significantly broaden educational access, thereby contributing to the realization of the SDG's objective of achieving universal education. The main purpose of this study is to promote the inquiry-based learning to the interdisciplinary socioscientific issues. It aims to improve the problem of physical activities' inability to expand knowledge content and course resources in a timely manner. On the contrary, although the <City Auncel> system already has designated data from the real open data, the advantage of online learning allows students to freely explore related knowledge beyond the data offered in the system. Therefore, through promoting scientific literacy and active learning strategies, an inquirybased learning model was developed to encourage students for free exploration of problems, unimpeded collection, analysis and interpretation of exploratory content, and writing their ideas about social issues. In the end, students present their exploration results and content in the system.

This study utilizes the background of socioscientific issues to generate open-ended questions, serving as a guide for students to engage in initial open-ended inquiry activities. The system provides students with various forms of data, such as images, text, and data visualization, to facilitate an in-depth exploration of social science issues related to Taiwan. This allows students to actively explore and discover the complex relationships inherent in social science issues. The purpose of this study is to investigate how students can conduct open-ended inquiries in social science issues and what the content and outcomes of such inquiries are.

# 2. Related Work

# 2.1 Inquiry-Based Learning

Inquiry-based learning is a learning method based on problem-solving and exploration. This approach encourages active student engagement in the learning process through steps such as posing questions, gathering data, conducting experiments, analyzing, and evaluating to gain a deeper understanding of the subject matter. According to Prince (2004), inquiry-based learning enhances students' learning motivation and interest, encouraging them to explore the learned knowledge in depth. Banchi and Bell (2008) also suggest that inquiry-based learning is a teaching model designed around students' exploration of the natural world and society, aiming to help them construct their own knowledge by encouraging questioning. investigation, knowledge discovery, and application. Research by Kuhn (2007) indicates the benefits of inquiry-based learning for developing students' critical thinking and problemsolving abilities. It has also been shown to increase students' learning motivation and interest (Blumenfeld et al., 1991; Prince, 2004). However, implementing inquiry-based learning requires teachers to have sufficient time and resources for design and implementation, as well as continuous assessment and adjustment of instructional strategies to ensure that students achieve the learning goals (Banchi & Bell, 2008). While some scholars believe that inquiry-based learning is less effective than direct instruction, in practice, the effectiveness and applicability of inquiry-based learning largely depend on specific contexts and learning objectives, the extent of students' foundational learning skills and background knowledge, and the need for teachers to provide appropriate guidance and support. Therefore, further exploration and research are needed to fully understand the effectiveness and applicability of inquiry-based learning.

Based on the variations in the structure, guidance, and level of student and teacher involvement, inquiry-based learning can be categorized into structured, guided, and open inquiry learning models.

- 1. Structured Inquiry: Structured inquiry refers to the explicit definition and explanation of learning objectives by the teacher, who provides specific guidance and support to assist students in their inquiries. According to Hmelo-Silver et al. (2007), structured inquiry helps students grasp learning objectives and improves learning outcomes. Furthermore, structured inquiry can enhance students' learning outcomes and efficiency, enabling them to quickly master learning content and skills. However, structured inquiry may limit students' creativity and autonomy in learning (Kirschner et al., 2006).
- 2. Guided Inquiry: Guided inquiry involves the teacher providing a certain degree of guidance and support while allowing students some autonomy in their investigations. Research by Hung et al. (2005) indicates that guided inquiry helps students discover problems, enhance learning motivation and interest, and engage in interesting and challenging learning activities. Guided inquiry also promotes students' critical thinking and problem-solving abilities. However, guided inquiry may require teachers to invest more time and effort in designing and guiding learning activities.
- 3. Open inquiry: Open inquiry refers to the teacher not explicitly defining or explaining learning objectives, but instead allowing students to freely explore and discover knowledge. Studies by Hmelo-Silver et al. (2007) show that open inquiry enhances students' autonomy and creativity in learning. Students are free to explore and discover new knowledge, and it helps them develop critical thinking and problem-solving skills. However, open inquiry may require more time and resources to support students' learning.

However, in inquiry-based learning, structured, guided, and open inquiry are commonly used approaches, each with its own advantages, disadvantages, and suitable contexts. According to Kirschner et al. (2006), structured inquiry emphasizes students learning in a structured environment, where teachers provide clear learning objectives, guidance, and detailed learning paths. This approach helps students acquire important knowledge and skills but may limit their creativity and independent thinking abilities. On the other hand, guided inquiry offers more freedom and flexibility while providing students with appropriate guidance and support. Hung et al. (2005) found that guided inquiry emphasizes students discovering and solving problems during the inquiry process. Teachers provide relatively ambiguous learning objectives and guidance to stimulate students' learning motivation and creativity. This approach enhances students' autonomy and sense of responsibility in learning but may require them to face more challenges and difficulties. According to Lazonder & Harmsen (2016), guided inquiry often leads to higher learning effectiveness because it provides appropriate guidance, making it easier for students to understand the goals and process of inquiry. It is suitable for relatively simple and clear-cut problems.

Lastly, open inquiry is a completely independent learning approach suitable for complex problems that require students to engage in deeper exploration and analysis. Students have the freedom to choose problems and solutions and independently explore and learn during the process. Hmelo-Silver et al. (2007) found that open inquiry can stimulate students' learning interests and motivation, enhance their creativity and independent thinking abilities, as well as improve their autonomy and sense of responsibility in learning (Hung et al., 2005). However, open inquiry may also present challenges and difficulties. It requires students to have sufficient prior knowledge and skills; otherwise, it may lead to poor learning outcomes. Hmelo-Silver et al. (2007) noted that open inquiry can be challenging for students, particularly when they lack the necessary prior knowledge and skills. Therefore, teachers need to provide appropriate support and guidance in open inquiry to help students overcome difficulties and improve learning outcomes.

### 2.2 Real Open Data

Real open data can facilitate the effective utilization and management of resources. As highlighted by Davies et al. (2017), real open data can provide valuable information on energy, water resources, land use, waste management, and other aspects, thereby assisting governments and businesses in better managing these resources. Governments can gain a better understanding of energy consumption in each region through open data, enabling them to formulate more effective energy-saving policies. Real open data also promotes knowledge sharing and technological innovation. As articulated by Giest et al. (2017), real open data supports communication and collaboration among different knowledge domains, thereby fostering technological innovation and development. For example, open data allows scientists and researchers from various fields to share and utilize data, advancing scientific research. It can even drive sustainable investment and the development of a green economy. Pandey et al. (2017) underscore that real open data can provide information on environmental and social benefits, helping investors and businesses make more informed investment decisions. Additionally, it can promote green industries and innovation, thus propelling the growth of a green economy.

In addition to these issues, open data also has profound social and economic impacts. Firstly, open data can enhance transparency and trust in government and other institutions. Open data can increase public trust in the government, leading to increased public participation and support (Bertot et al., 2010). Secondly, open data can foster innovation and economic growth. Research by Kuk et al. (2016) suggests that open data can generate new business models and employment opportunities, thus promoting economic growth and innovation. Lastly, open data can also facilitate citizen science and social engagement. Wiggins and Crowston (2011) found that open data can enhance citizen science and social participation, increasing public understanding and involvement in scientific and social issues. It can even help governments and businesses better understand the problems they face and find more effective solutions (Janssen et al., 2012). In urban planning, real open data can provide valuable information on city traffic flow, pollution, water usage, and other aspects, helping urban planners gain a better understanding of the city's situation and make more informed decisions.

Furthermore, real open data can also promote public engagement and social innovation. Hui et al. (2014) stated that real open data enables the public to have a better

understanding of the operations of governments and businesses, thus facilitating public participation and oversight. Additionally, real open data can foster social innovation, such as through the innovative use of open data for new products and services.

In today's society, sustainable development has become a global issue due to the significant impact of human lifestyles and economic activities on the Earth's environment. To achieve sustainable development, innovation and improvement are required in various fields. In this regard, real open data has been proven to be a highly valuable tool with significant application value and advantages in the context of sustainable development.

However, there are also challenges that need to be addressed, such as privacy and security concerns, as well as potential shortcomings in data quality and reliability. The research of Halevy et al. (2009) found that many open datasets suffer from missing values, errors, and inconsistencies with reality, which can negatively impact data analysis and applications, affecting data effectiveness. Additionally, issues related to privacy and security may lead to the exposure of personal identities and sensitive information, as well as potential misuse. Moreover, open data can potentially have negative social impacts, such as perpetuating racial discrimination and implicit biases. Therefore, effective measures need to be implemented in the application and promotion of real open data to address these issues and achieve more effective and sustainable development.

Although open data is a powerful tool that can play a crucial role in various domains, but there are also challenges and issues that need to be addressed. To maximize the value of open data, efforts should be focused on resolving privacy and security concerns, improving data quality and reliability. Additionally, further exploration is needed to understand the social and economic impacts of open data better, in order to grasp its value and potential more effectively.

#### 2.3 Action Decision

Action decision is a situation that humans often face in daily life, and many scholars have conducted in-depth research on this type of decision-making. For example, Dawes and Corrigan (1974) pointed out that "in action decision-making, individuals typically refer to their experience and knowledge, but at the same time are also influenced by non-rational factors such as emotions, motives, and situations." In other words, when individuals make decisions, they not only consider their own knowledge and experience but also are influenced by non-rational factors such as emotions and motives. In the actual decisionmaking process, people can use various methods. For example, Simon (1956) proposed the concept of "bounded rationality," stating that people only consider local factors when making decisions and cannot consider all possible choices. In addition, people can also use heuristics and rules to simplify the decision-making process (Tversky & Kahneman, 1974).

In the field of education, action decision refers to the action choices made by learners when facing various situations, which is an important process of self-adjustment and reflection for learners. Action decision refers to the action choices made by learners in a specific context and is made after considering multiple factors (Muis et al., 2015). These factors can be the learner's learning goals, understanding of the task, time constraints, learning experience, and so on. Zimmerman and Risemberg (1997) pointed out in their research that learners' action decisions can reflect their learning motivation and learning strategies. At the same time, action decisions can also affect learners' learning outcomes. If learners can choose the most suitable action plan, their learning effect may be better. Helping learners improve their ability to make action decisions is an important teaching goal.

#### 3. Online System Design and Activity Process for Socioscientific Issues

This study focuses on the socioscientific issue of the "Endangered Status of The Leopard Cat," encompassing SDGs 11 (Sustainable Cities and Communities), 13 (Climate Action), and 15 (Life on Land). The study's exploration data primarily originates from the Miaoli area of Taiwan, where the leopard cat's survival and activity rates are relatively high. Students

utilize the <City Auncel> system to access data pertaining to designated land based on information about the leopard cat's endangered status. They gather information from diverse sources to address this socioscientific issue.

The system categorizes the data into five sections (*Figure 1*): (1) Water resources (latest & historical river water quality monitoring data, river distribution, etc.). (2) Land information (1984-2015 land use change, contour lines, terrain distribution, slope, etc.). (3) Introduction to the leopard cat (characteristics, habits, habitat, observed locations, road kill records, etc.). (4) Forestland (forest type and distribution), and (5) Demographics (population distribution, age distribution, etc.).

The online system offers a diverse array of data and records the outcomes of students' explorations, facilitating the review of each student's findings post-exploration. Through the online system, various forms of data presentation are available, including interactive images, text, and statistical data tables. These aids help learners quickly comprehend the activity content and engage in a comprehensive inquiry-based learning process. The insights gained from the inquiry can also be incorporated back into the learning materials.



Figure 1. Real Open Data Domains <City Auncel>

This activity is structured around three roles, each representing a country's government, conservation groups, and corporations. Through role-playing, students initially delve into the endangered status of leopard cats and subsequently analyze appropriate regions for development or conservation efforts. They then engage in discussions and coordination to determine potential overlaps between areas earmarked for development and those requiring conservation, ultimately crafting a cohesive national plan. In this activity, students encounter three open-ended questions presented from distinct role perspectives (*Figure 2*). This approach empowers them to engage in open-ended inquiry, fostering the generation of distinct insights into socioscientific matters. Through meaningful discussions and thoughtful reflections, they gain exposure to their peers' research findings and outcomes, thereby enhancing their comprehension of the significance and intricacy inherent in socioscientific issues.



Figure 2. Inquiry of respective role

The activity comprises four stages (*Figure 3*) During Step 1, students receive a 15minute introduction to the activity's theme and system operation. In Step 2, students delve

into answering three open-ended questions: (1) the survival crisis of the leopard cats, (2) identifying suitable areas for development or conservation, and (3) examining whether areas chosen for development and conservation align across different roles. This stage takes approximately 90 minutes. Moving to Step 3, students explore the reasons behind overlapping areas for development and conservation and propose potential solutions. Lastly, Step 4 involves voluntary sharing of exploration content, reflections on the topic, and overall sentiments about the activity.





# 4. Research Design

The study's experimental subjects consisted of students from a university's in-service graduate program in Taiwan. The participants were grouped randomly, with each group comprising 3 students, resulting in a total of 18 students divided among 6 groups. The majority of these students possessed educational and information backgrounds. Their ages ranged from 24 to 50 years on average.

The open-ended inquiry process and the content of <City Auncel> are presented in Table 1. The initial step aims to help students comprehend the issue's significance and the contributing factors to the survival crisis. For this purpose, the question "What causes the survival crisis of the leopard cat?" is posed. In <City Auncel>, students have the opportunity to explore the central topic and pinpoint the factors that have led to the crisis.

The second question caters to the unique inquiries from different roles, specifically addressing the need for conservation or development in certain areas. This question encourages students to generate their own crisis factors, requiring them to focus on their designated roles and explore outcomes that align with their interests as corporations, government representatives, or conservation groups.

The third question is strategically designed to stimulate collaboration and discussions among students. It centers on determining whether any overlap exists between the areas targeted for conservation by conservation groups and those earmarked for development by corporations. To accomplish this, the query "Is there an overlap between areas that can be developed and areas that need to be conserved?" is presented.

Through the formulation of these three open-ended questions, students are prompted to delve deeper into investigating and addressing the issues at hand.

Step	Purpose	Question Type	<city auncel=""> example</city>
Step 1	Discover Problems		What causes the survival crisis of the leopard cat?
Step 2	Solve / Generate Crisis	open-ended	Which area need to be conserved/developed? Why does it need?
Step 3	Consultations	-	Is there an overlap between areas that can be developed and areas that need to be conserved? How to solve it?

#### Table 1. Open Inquiry Model

Considering the previously outlined design of open-ended questions, this study examines students' textual responses to these three inquiries. The hypothesis posits that students derive their answers from the data made available through the system. Table 2 presents analyzed sample sentences extracted from student responses. The categorization of students' answers is structured around five dimensions: (1) Water Resources, (2) Land Information, (3) Animal Conservation, (4) Vegetation Distribution, (5) Population, and (6)

Other Aspects. In cases where a student's writing addresses multiple dimensions, their text is subjected to a multidimensional textual analysis approach.

Table 2 Text Anal	lysis	Exam	ple
-------------------	-------	------	-----

Role	Example Sentences	Data Domain	Multiple or Single
Consortium	Water quality: RPI=7.25, severe pollution (6+10+3+10). Land use map: Mostly developed area. The region is an important habitat for leopard cats.	Water Resources Land Information Animal Conservation	multiple
Consortium	According to water quality detection, the data indicates that the pollution index of the "Nangang Creek Bridge River" reaches a high level of 7.25, classified as severe pollution.	Water Resources	single
Conservation Groups	From the roadkill statistics of leopard cats in various sections of Miaoli County, it can be observed that Section 128 of the county is a high-risk roadkill zone for the leopard cats. Furthermore, based on the highway network map of the Miaoli region, it is evident that this section constitutes an expressway. Consequently, the construction of this expressway would place the leopard cats at a heightened risk.	Land Information Animal Conservation	multiple

# 5. Results

While exploring the question "What causes the survival crisis of the leopard cat?", 18 students divided into six groups undertook diverse perspectives, yielding 56 distinct inquiry outcomes (Table 3). Upon analyzing and categorizing phrases and statements featured in students' reports, it becomes evident that a significant portion of the students centered their inquiries around water resources, land resources, and animal data.

Upon analyzing the textual results of students' inquiries, it was evident that nearly all students recognized the significant issue of roadkill as a primary contributor to the leopard cats' survival crisis. Moreover, students identified areas affected by water pollution and acknowledged the consequences of historical land development changes on the survival of leopard cats. These factors collectively led to extensive habitat destruction for the leopard cats, resulting in a continual decline in population and worsening living conditions.

However, aspects such as vegetation distribution and population structure were less frequently acknowledged in the students' inquiry outcomes. The researcher of this study suggests that students might have encountered challenges in linking data related to vegetation and population structure with the leopard cats' survival crisis. This discrepancy could arise from the fact that, within the system, issues concerning water resources, land information, and animal data were more immediate and easily discernible matters that could be addressed using accessible data. On the other hand, comprehending the correlation between vegetation and population structure required more intricate connections with other data points. Put simply, students may have identified issues stemming from land development but perhaps did not extend their thoughts to encompass the relationship between these issues and population structure. The same logic applies to vegetation data, where although vegetation holds relevance for the leopard cats' habitat, its connection to the specific question might not have been as apparent. Consequently, the students' inquiry results indicate that during the discussion phase of the inquiry process, students predominantly emphasized interpreting data tied to pivotal factors. Essentially, within the context of the inquiry question, students gravitated towards interpreting and investigating data directly pertinent to water, land, and animal information. This encompassed addressing concerns such as roadkill, water quality pollution, and modifications in land development. These factors stood out as the most directly influential elements affecting the survival environment of the leopard cats.

Data Domain	Total Frequencies of Inquiry Content Items
Water Resources	13
Land Information	17
Animal Conservation	21
Vegetation Distribution	2
Population	3
Total Items	56

Table 3	Total Fred	nuencies	of Inc	uirv	Items	in I	Data	Cateo	ories
	1010111100	Juchicics	or my	uny	1101113		Jaia	Calcy	51103

Regarding the second question, "Which area needs to be conserved/developed?" The outcomes derived from students' inquiries reveal that, upon analyzing content items mentioned in their reports and categorizing them by different roles, the Conservation group generated 23 inquiry results, while the Consortium produced 27, resulting in a total of 50 (refer to Table 4). In comparison to the first question, students' responses to this question exhibit greater depth and diversity. However, the overall focus of inquiry remains aligned with that of the first question, with a higher emphasis on land data and animal information in contrast to other data aspects.

Additionally, when inspecting the outcomes through the lens of role differentiation, students assuming the role of Conservation groups predominantly concentrated on animal data for this question. Most Conservation groups cited areas marked by pronounced leopard cat roadkill and habitat concerns as key areas for conservation.

Conversely, drawing from the outcomes of students in the Consortium role, it becomes apparent that their exploration and explanation of answers to this question unfolded in a more multifaceted manner. Their primary inquiries revolved around geographical location and population distribution driving the tourism sector, among other factors. This diverges from the content of the first question, thereby displaying a lower degree of similarity. Textual analysis outcomes demonstrate that the primary basis for development centered on the convenience of transportation due to geographical location or population distribution. These research results underscore a high level of engagement among students and reveal distinctly varied inquiry focuses based on the roles they undertook.

Data Domain	<b>Conservation Groups</b>	Consortium	Total
Water Resources	4	3	7
Land Information	5	9	14
Animal Conservation	10	5	15
Vegetation Distribution	4	0	4
Population	0	7	7
Others	0	3	3
Total Items	23	27	50

Table 1. Differences in Content of the Inquiry

# 6. Conclusion

This study developed an online platform <City Auncel> for exploring social issues utilizing real data. Through an inquiry-based learning approach, it aims to foster students' scientific literacy and underscore sustainability concerns tied to the Sustainable Development Goals (SDGs). The platform offers a variety of interactive visualized data. Educators can craft learning materials or activities to facilitate three modes of inquiry-based learning: structured, guided, and open-ended.

Within this study, an open-ended inquiry-based learning model was introduced, affording students the autonomy to explore a range of authentic and open data. Throughout the inquiry-based learning process, spanning from content investigation and data collection to final deduction and interpretation, students construct their own knowledge to address socioscientific issues.

The study's outcomes demonstrate that on the whole, students adeptly executed inquiries using the system. They identified multiple factors contributing to the survival crisis of the leopard cats. Textual analysis unveiled that students' inquiry directions tended to concentrate on key data closely linked to the issue. However, data with weaker associations were less likely to emerge in students' inquiries. This matter warrants further investigation and analysis. Researchers could classify the multidimensional data provided into levels of simplicity, moderate complexity, and challenge, thus better gauging students' inquiry capabilities.

Moreover, whether evaluating the overall inquiry outcomes or individual student inquiries, uneven distributions in data exploration were observed. This suggests that while the data and inquiry questions formulated in this study were open-ended, students' roleplaying potentially constrained their comprehensive data exploration and interpretation. This bias might result from incomplete data exploration dimensions. Despite presenting diverse data for student exploration, the study uncovered skewed outcomes in the broader inquiry process. To broaden students' access to data and information, the relevance of roleplaying needs consideration. Alternatively, enhancing interactions and debates among roles could help counteract bias in data exploration.

Nonetheless, the study highlights a high level of engagement among students. Each role exhibited distinct inquiry foci, and students made logical action decisions based on their inquiry findings. However, situations requiring physical coordination and interaction were influenced by real interpersonal relationships, impacting assessment outcomes. From these findings, it's evident that online open-ended inquiry, shaped by role-playing, can guide students' attention toward specific key data while providing them with an immersive learning experience.

# Acknowledgements

This study is supported in part by the National Science and Technology Council (previously known as Ministry of Science and Technology) of Taiwan, under MOST 108-2511-H-008 -016 -MY4 and NSTC 112-2410-H-008 -010 -MY3.

# References

Banchi, H., & Bell, R. (2008). The many levels of inquiry. Science and children, 46(2), 26.

- Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government information quarterly*, 27(3), 264-271.
- Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational psychologist*, 26(3-4), 369-398.
- Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological bulletin*, 81(2), 95.

Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., ... & Leach, A. R. (2017). The ChEMBL database in 2017. *Nucleic acids research*, *45*(D1), D945D954.

Giest, S. (2017). Big data for policymaking: fad or fasttrack? *Policy Sciences*, 50(3), 367382.

- Halevy, A., Norvig, P., & Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE intelligent systems*, *24*(2), 8-12.
- Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and. *Educational psychologist*, 42(2), 99-107.
- Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and. *Educational psychologist*, 42(2), 99-107.
- Hung, W., Jonassen, D. H., & Liu, R. (2008). Problem-based learning. In *Handbook of research on educational communications and technology* (pp. 485-506). Routledge.
- Janssen, M., & Kuk, G. (2016). The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly*, *33*(3), 371-377.
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information systems management*, *29*(4), 258-268.
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, *41*(2), 75-86.
- Kuhn, D. (2007). Is direct instruction an answer to the right question? *Educational psychologist*, *42*(2), 109-113.
- Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of educational research*, *86*(3), 681-718.
- Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G., Meier, E., & Heddy, B. C. (2015). The curious case of climate change: Testing a theoretical model of epistemic beliefs, epistemic emotions, and complex learning. *Learning and Instruction*, *39*, 168183.
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of engineering education*, *93*(3), 223-231.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological review*, *63*(2), 129.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, *185*(4157), 1124-1131.
- Wiggins, A., & Crowston, K. (2011, January). From conservation to crowdsourcing: A typology of citizen science. In *2011 44th Hawaii international conference on system sciences* (pp. 1-10). IEEE.
- Zimmerman, B. J., & Risemberg, R. (1997). Becoming a self-regulated writer: A social cognitive perspective. *Contemporary educational psychology*.

# Empowering Language Learners: Harnessing Computer-Based Writing for Enhanced Chinese Language Proficiency

Lung-Hsiang WONG<sup>a\*</sup>, Guat Poh AW<sup>a</sup>, Wenli Chen<sup>a</sup>, Yin Ling CHEUNG<sup>a</sup>, & Seok Hwa SIM<sup>b</sup>

<sup>a</sup>National Institute of Education, Nanyang Technological University, Singapore <sup>b</sup>Curriculum Planning and Development Division, Ministry of Education, Singapore \*Ihwong.acad@gmail.com

**Abstract:** Computer-based writing (CBW) has become integral in language education, enabling personalised learning and online collaboration. In Singapore, CBW is being introduced in Chinese Language (CL) classes, yet challenges persist, particularly for students using the phonetic-based *pinyin* input method. This study investigated the impact of partially replacing paper-and-pen-based writing (PPBW) with CBW in primary schools. Using mixed-effects modeling, we compared post-CBW and post-PPBW scores, considering CL achievement levels. Results indicated CBW's positive influence on writing quality regardless of prior competence. Furthermore, CBW significantly benefits low and medium-achieving students' performance. Typing speed and handwriting speed are influential for these students, while different writing strategies play distinct roles. Informed by the findings, we propose an online writing community platform, WeeWrite, incorporating generative AI to enhance CL CBW skills, offering personalised and collaborative learning experiences.

**Keywords:** Computer-assisted language learning; Computer-based writing; Chinese as a second language; ChatGPT in language learning

# 1. Introduction

Computer-based writing (CBW) has become the primary mode of writing in our daily life. The proliferation of Information and Communication Technology (ICT) has also opened up opportunities for language students to leverage digital tools that can personalise their learning and join online collaborative learning communities with text-based communication. This implies the need for them to be adept in CBW skills to optimise their technologyenhanced learning experience and outcomes. In Singapore, CBW is gradually being infused into the Chinese Language (CL) classes in primary and secondary schools, as well as national examinations in secondary schools (Tushara, 2023). Yet the students, most having a second language (L2) standard, continue to grapple with the challenge of making the 'indirect' keyboard-based input method (the phonetic-based pinyin) an efficient alternative to paper-and-pen-based writing (PPBW). Furthermore, decades of comparative studies in CBW versus PPBW in L2 classrooms have yielded inconsistent findings on which writing mode would be more conducive for young students to produce writing of higher quality (Cheung, 2012). While we do not advocate a full replacement of PPBW with CBW in L2 classes, we see the need to infuse more CBW activities, not only to reflect the real-world use of CBW, but also to prepare students to communicate electronically when they grow up.

The purpose of this baseline study is to discover the conditions under which students can benefit from infusing more CBW in primary school classrooms. Conditions may vary for students with different CL achievement levels and characteristics. We intend to derive differentiated ICT-aided means from the research findings to support students of different profiles in improving CL CBW competences. Informed by the findings, we propose a design model of an online writing community platform, WeeWrite (We e-Write; "共写未来"). The

proposed platform will harness generative Artificial Intelligence (AI) and other advanced technologies to deliver writing pedagogy and facilitate individual, personalised and collaborative learning out of class.

# 2. Literature Review

# 2.1 Rationale and challenges for promoting CBW in language classes

E-writing is replacing PPBW to become our primary mode of writing (Li, 2006). Authoring for the purpose of connecting with people and sharing knowledge is becoming part of digital natives' way of life and they exert subtle influences on learners' holistic development (Greenhow et al., 2009). Thus, a necessary language pedagogical shift, without hindering the development of basic handwriting and PPBW skills, is to situate learners more intensively in the e-writing space (Bolter, 1991). However, as CL scripts are not based on the Latin alphabet system but are logographic, there are barriers for students to adopt the 'indirect' phonetic-based *pinyin* input method. The use of *pinyin* input requires L2 writers' additional mental processing (Xie, 2001), on top of the "double translation" effect (Wolfe & Manalo, 2004; Wong et al., 2011) where they may perform a mental translation from their first language to L2, and then translate from L2 into keystrokes.

### 2.1.1 Prior Studies on CBW versus PPBW for compositions in language classes

Prior research comparing CBW and PPBW in English Language (EL) classes has yielded inconsistent findings. The mixed results may be due to variations in sample sizes, students' age groups and language and computer proficiencies, task types and software used, etc. (Cheung, 2012). the quality of CBW compositions has been found to be higher than PPBW compositions in some studies (e.g., Bernhardt et al., 1989; Lam & Pennington, 1995; Li, 2006). Other researchers, however, have found no significant differences in this aspect (e.g., Burley, 1994; Dalton & Hannafin, 1987; Wolfe & Manalo, 2004).

Some studies collected (e.g., with screen capturing tools) and analysed the writing process data to unveil the students' writing patterns in both modes. For example, Li (2006) discovered that during PPBW, adult English as second language (ESL) learners spent more time in pre-planning; whereas for CBW, they carried out more higher-order thinking activities while evaluating their written texts. Zoghi and Reshadi (2013) unpacked ESL students' writing process and found that those engaged in CBW and PPBW did not show significant differences in the employment of rhetorical, metacognitive, cognitive, communicative and social/affective strategies. Yet students in the CBW group had tapped on the technological affordances to either perform certain productive strategies more frequently (e.g., formatting/modelling, which is made easy by the technology), or be offloaded from performing unproductive strategies (e.g., less rehearsing, due to the ease of editing). Elsewhere, focusing on analysing L2 learners' revision behaviour in CBW, Barkaoui (2016) discovered that it was L2 proficiency and task type, not keyboarding skills, that affected such behaviour. The subjects with low L2 proficiency performed more low-level (e.g., typographic) revisions at the expense of higher-level revisions.

Scattered research investigated student perceptions on CBW versus PPBW. For example, three studies focused on measuring the students' writing anxiety through administering Second Language Writing Anxiety Inventory (L2WAI) (Cheng, 2004) to their respective samples (college ESL students) in CBW and PPBW groups before and after the writing tasks. L2WAI encompasses three sub-components: cognitive anxiety (mental aspect of anxiety, i.e., negative expectations, preoccupation with performance, etc.), somatic anxiety (autonomic responses, i.e., nervousness and tensions), and behavioural anxiety (procrastination, withdrawal, and avoidance behaviour). Bailey et al. (2017) found that the students in the CBW group increased significantly in behavioural anxiety, and the students in PPBW group increased significantly in somatic anxiety, while there is no decrease in any form of anxiety. The results contradicted the findings of Kurt and Atay (2007), and Foroutan and Noordin (2012), which indicated significant decreases in such anxiety. Compared to the abundant prior research in EL CBW, the research on Chinese CBW (particularly in comparison with those of PPBW) is relatively scarce. For example, Kang (2011) and Zhu et al. (2015) studied Chinese as Foreign Language college students in China respectively. Kang et al. (2011) found that CBW helped intermediate students more (compared to beginners) as they produced longer articles and in shorter time. Zhu et al.'s (2015) study, on the contrary, showed that CBW benefited students of low proficiency more as the scores and lengths of their CBW compositions were significantly greater than those of their PPBW compositions. Instead, the differences between the CBW and PPBW compositions written by students of high proficiency were not significant.

Again, all these conflicting findings have shown that the research evidence on CBW versus PPBW is context dependent. Thus, Chinese as L2 educators would be baffled should they consult the literature on how and to what extent they could employ CBW in writing classrooms. Our study is intended to overcome most of the limitations in the prior research, especially that our focus is on younger CL as L2 students.

#### 3. Method

#### 3.1 Participants and data collection

This study aimed to investigate the impacts of partially incorporating CBW compared to PPBW (status quo) in CL classes in primary schools in Singapore. The focus was to compare the effects of the writing media used, not the writing pedagogy. This was accomplished through measuring project participants' writing products and relevant perceptions. 127 Primary 5 (11-year-old) students from four public primary schools participated in this study. All students began writing full compositions with paper and pen in CL classes a year ago. Yet they had no prior experience in writing full compositions on computer. The research questions (RQs) are:

(RQ1) What is the comparative performance of students in the post-CBW test when accounting for the influence of the pre-CBW test?

(RQ2) Is there a significant performance disparity between the students' post-CBW test and their post-PPBW test?

(RQ3) What factors predict the post-CBW scores of high-achieving (HA), mediumachieving (MA), and low-achieving (LA) students, regardless of the treatment conditions?

Three types of data were collected over one academic year, namely, CL compositions, pre- and post-surveys, and computer writing and handwriting speed tests.

(1) CL compositions

We tapped on regular CL in-class composition sessions to collect CBW and PPBW artefacts across an entire year. The participating students were randomly assigned to two treatment groups with varied ratios of CBW and PPBW sessions, namely, **(G1)** 5 CBW + 3 PPBW sessions; **(G2)** 3 CBW + 5 PPBW sessions. The total number of students (randomly) assigned to G1 and G2 across the participating schools were 63 and 64 respectively. Notebook computers with Microsoft Word installed were used for the sessions. All schools conducted each writing session for 50 minutes, with the task type (pictorial compositions / narrative essays) and teacher scaffolding designs modelled on the national CL syllabus. However, as each school has its own school-based elements in their writing curriculum, we were not able to control their writing prompts (the pictures) across the sessions.

To mitigate the order effect, the PPBW and CBW sessions in each group were evenly spread across the academic year. The first and the last (eighth) sessions were PPBW sessions and were treated as pre-PPBW and post-PPBW tests respectively. The second and the seventh sessions were CBW sessions and were treated as pre-CBW and post-CBW tests respectively. We randomly selected 10% (52) of the compositions collected from these four sessions and asked two raters (who were retired CL master teachers) to score them independently using a validated rubric derived from the Ministry of Education's teachers' guide to primary school CL curriculum. Each composition was scored by the following dimensions (in the Likert scale of 1 (poor) to 5 (excellent)): material, theme, narration, grammar, structure, and vocabulary. Interrater reliabilities were then computed, with the

absolute agreement average measures of intraclass consistency coefficients (ICC) ranging from .86 to .90 (i.e., all >.75), indicating excellent reliability across all dimensions (Cicchetti, 1994). The rest of the compositions were divided in half and assigned to the raters to score.

(2) Pre- and post-surveys

All participants were administered pre- and post-surveys before the first writing session and after the last writing session respectively. The survey, in a Likert scale of 1 to 5, combines three validated instruments, namely, Motivation and Learning Strategy Questionnaire (MSLQ) (Pintrich et al., 1991) (to measure their intrinsic goal orientation, extrinsic goal orientation, and task value in CL writing), L2 Writing Anxiety Inventory (L2WAI) (Cheng, 2004) (to measure their cognitive anxiety, somatic anxiety, and avoidance behaviour in CL writing), and Writing strategy questionnaire (Bai et al., 2014) (to measure their frequencies of applying planning, monitoring/evaluation, revising, text-generating, and resourcing strategies in their CL PPBW). There are additional items the in post-survey to measure their strategies in CL CBW.

(3) Pre- and post-computer input or handwriting speed test

All participants were administered pre- and post-CL input speed tests and EL input speed tests. The EL input speed is a proxy for their general keyboarding speed as EL typing is more direct, compared to the indirect *pinyin* input. In each test, they were required to copy a recently taught CL or EL textbook passage. The time limit of each test was two minutes. The correct number of Chinese character or English words input per minute was determined.

Furthermore, we are interested in identifying the varied conditions that may affect HA, MA and LA students' post-CBW test scores. The results may derive differentiated, personalised writing instruction for each student band. Thus, we used the pre-PPBW score as a proxy for their CL writing achievement level and divided them into the three bands (HA – 39 students; MA – 48 students; LA – 40 students; the uneven student counts are due to the same scores of "borderline" students).

# 3.2 Data analysis

For the statistical analysis, the post-CBW test scores are the dependent variable to address all three RQs. The pre-CBW scores are the covariate to address RQ1. The post-PPBW scores are the predictor to address RQ2. All the variables related to the pre- and post-survey and speed tests are predictors to address RQ3.

Mixed-effects modeling (Laird & Ware, 1982) was used for most of our analyses. Compared to traditional approaches such as *t*-test and ANCOVA, mixed-effects modelling is more flexible as we can combine and analyse data collected from multiple schools into a single model while accounting for the inherent variability and dependencies within each school. In essence, a mixed-effects model consists of fixed effects and random effects. Fixed effects represent the variables that are of primary interest in our analysis, while random effects capture the variability across the different schools. In addition, we performed stepwise regression analysis to determine the best fit models of the three student bands.

# 4. Findings and Preliminary Discussion

# 4.1 Comparison of pre- and post-test scores (to address RQ1 and RQ2)

We performed mixed effects modelling on the pre- and post-CBW scores and the post-PPBW score in place of paired *t*-tests or ANCOVA to account for the school factor as the random effects. Tables 1 and 2 show the descriptive statistics and the results of the analyses respectively. Due to the space limit, only the *t*-values or the *F*-values, and the statistical significance of the score differences, are presented in Table 2. We excluded pre-PPBW test score from the analysis as it was only treated as the proxy for the students' prior CL competence and was used for dividing them into the three bands.

Table 1. Descriptive statistics (mean (standard deviation))

	All students	G1	G2	HA	MA	LA
Pre-CBW	16.44 (4.68)	16.64 (4.48)	16.24 (4.90)	20.13 (3.47)	16.56 (3.77)	12.70 (3.71)
Post-CBW	17.69 (3.81)	17.91 (3.32)	17.46 (4.26)	20.33 (3.56)	17.75 (2.56)	15.03 (3.52)
Post-PPBW	17.11 (3.34)	17.06 (3.64)	17.02 (3.23)	18.85 (2.93)	17.25 (2.73)	14.89 (2.89)

Table 2. Mixed effects modeling analyses to compare the test sc	cores
---	-------

	Comparison of the two scores within each group/band (t-value)					Comparison of the two scores across groups or bands ( <i>F</i> -value)		
	All students	G1	G2	HA	MA	LA	G1 vs G2	HA vs MA vs LA
Pre- vs post- CBW	13.308**	10.042**	12.406**	4.413**	6.770**	5.445**	.855	.477
Post-CBW vs post-PPBW	-1.488	912	-1.848	825	-2.250	-1.700	.967	27.167**

\* p<.05; \*\* p<.001

According to Table 2, all students improved significantly in the post-CBW scores compared to the pre-CBW scores, regardless of treatment conditions/groups, or student bands. No significant difference was observed when comparing post-CBW score across the groups or bands while controlling for their pre-CBW score.

On the contrary, there was no significant difference in the students' post-CBW and post-PPBW scores across the treatment conditions or student bands. When comparing post-CBW scores across treatment groups, no significant difference was found either. However, a significant difference emerged when comparing post-CBW scores across student bands. A post-hoc test of pairwise comparison with Bonferroni correction indicates that the HA students showed a significantly larger difference in scores between post-CBW and post-PPBW compared to both MA and LA students. Additionally, LA students exhibited a significantly greater difference in scores between the two tests than the MA students.

These findings seem to contradict early common assumptions that the "unnatural" CL typing would hinder beginners' CL writing performance and writing skill development (Zhang, 2021). By the end of the study, we not only observed students' significant improvement in their CBW writing regardless of their prior CL competences and CL typing experience, they produced CBW and PPBW artefacts with similar quality. Henceforth, the concern about the possible detrimental effect on students' writing development of partially replacing PPBW with CBW in CL classes may be less pertinent.

#### 4.2 The predictors of post-CBW scores of individual student bands (to address RQ3)

We conducted stepwise regression analyses to yield the best-fit models for predicting post-CBW scores within each student band. Due to the space limit and for ease of comparison, we present the three models in Table 3 in a summarising view.

			· · · · · · · · · · · · · · · · · · ·
	HA	MA	LA
Pre-CBW score	2.524*	5.074**	3.652*
Pre-PPBW score			-3.120*
Post-PPBW score		4.109**	10.125**
CL typing speed		2.648*	
CL handwriting speed			10.760*
Somatic anxiety (post-survey)	-9.402*		
Avoidance behaviour (post-survey)	-3.737*		
Planning strategies in PPBW (post-survey)		2.723*	
Planning strategies in CBW (post-survey)			-6.601**
Monitoring/evaluation strategies in PPBW (post-survey)	5.009**		
Text generation strategies in CBW (post-survey)		4.476**	
Resourcing strategies in CBW (post-survey)			8.244**

 Table 3: Best-fit models for predicting post-CBW scores within each student band (t-values)

\* p<.05; \*\* p<.001

Based on the respective predictor sets of the three models, we determined the characteristics of the students' CL competences in each band, which are triangulated by the

feedback from participating CL teachers of the study and our own extensive relevant experience and observations (three of the co-authors are educational researchers and teacher educators with vast experience in conducting school-based research in authentic CL classes; two of them are former teachers).

#### 4.2.1 The CL typing speed and handwriting speed

Do the students' CL handwriting and typing speeds affect their CBW performance? According to the models, both variables did not predict HA students' post-CBW scores, while MA and LA students' post-CBW scores were positively predicted by their CL typing speed and handwriting speed respectively. The plausible explanation is as follows.

For HA students, the switching of writing modes did not affect their general CL writing performance – they were likely to perform equally well when writing in both modes.

For LA students, based on the teachers' feedback and our team members' experience, we posit that they were generally weaker in vocabulary, grammar and content development. Furthermore, they often encountered the challenges of recalling the visual structure of individual CL characters to handwrite, after subvocalising the text internally (i.e., silently articulating the text in their mind, simulating the act of speaking without vocalising words aloud). We posit that these were likely to be the more prominent factors affecting their writing performance compared to their CL typing speed. The handwriting speed as a predictor of their CBW performance is seemingly counter-intuitive. Yet this can be explained as follows: slower handwriting speed may indeed reflect their difficulties in both composing the text in their heads and recalling character 'shapes'.

For MA students, their CL typing speed did matter. A possible reason is that they also faced the difficulty of recalling character 'shapes', whereas their overall linguistic skills were stronger than LA students (i.e., less challenging in subvocalising the text before transcribing it on computer). During CBW, pinyin input did help them get around the former difficulty.

Henceforth, we posit that switching to CBW would benefit both LA and MA students more in terms of producing higher quality written artefacts. The key is to provide them with further training in CL typing. We posit that if they are more adept at CL typing, their cognitive load in recalling character 'shapes' could be reduced. They could then devote more cognitive effort to subvocalising the text when writing (or developing the skills to do so). LA students are generally also weaker in the *pinyin* scheme. Additional training is needed in this area. For HA students, CL typing training may also be helpful, but need not be the focus.

#### 4.2.2 The writing strategies

In both the pre- and post-surveys, students were asked to indicate the frequency with which they used planning, monitoring/evaluation, resourcing, revising, and text generation strategies in PPBW. In the post-survey, they also indicated the frequency with which they used these strategies in CBW. Our analysis showed that not all the strategies could predict their post-CBW scores.

The frequency of using resourcing strategies (checking e-dictionary or other language references when encountering difficulties in writing) during CBW positively predicted LA students' scores. In Singapore's CL composition classes and examinations, students are allowed to check e-dictionary. Resourcing strategies as a predictor could mean that some LA students did not bother to make good use of the e-dictionary to compensate for their linguistic deficits, thus affecting their CBW performance.

The frequency of using text generation strategies (recalling segments of text or content ideas in other reading materials for use in their own writing) during CBW positively predicted MA students' scores. There is indeed a writing strategy known as "utilising impressive phrases and sentences" (好词好句) which has been a common strategy promoted by many Singapore primary CL teachers for composition writing. Some students deliberately memorise such phrases/sentences from other sources for this purpose. While we are cautious about using this strategy in the long term, we recognise that it is helpful for beginners, as it can be seen as a basis for imitation and modelling before they gain

confidence and move on to producing text largely in their own words. Combining this finding with the discussion of MA students' challenges in PPBW in section 4.2.1, we posit that MA students who produced better CBW articles may have had prior knowledge of writing, such as "impressive phrases and sentences", but were not confident in handwriting them because of the difficulty of retrieving the character 'shapes'. Switching from PPBW to CBW would optimise the effect of text generation strategies, as they would retrieve characters by pronunciation through *pinyin* input, which is directly linked to their subvocalised text.

A surprising finding is that the frequency of applying planning strategies (e.g., outlining, thinking or discussing with others before starting to write) in CBW negatively predicted LA students' post-CBW scores. A possible explanation is that LA students' ability to comprehend the writing instructions and the writing prompts (pictures for pictorial composition) is weaker. Thus, those who used the strategies tended to plan for content that was off-topic or of lower linguistic quality.

Two PPBW strategies emerged as predictors of post-CBW scores (planning strategies for MA students, and monitoring/evaluation strategies for HA students). We suspect that this is because the students filled out the PPBW-related survey items before the CBW items, and some of them did not consciously distinguish between the strategies they used in PPBW and in CBW. With the expectation that the PPBW strategies can be transferred to CBW when students are given more CBW opportunities, we can assume that these strategies are also conducive to CBW for the respective student bands.

All in all, the five categories of strategies could be effective for improving writing quality if they are properly used. However, it may be a tall order for young students to pick all of them up at once. Our findings might suggest strategies to prioritise the development of these strategies for students of different bands – monitoring/evaluation strategies for HA students, planning strategies and text generation strategies for MA students, and resourcing strategies for LA students. Also, given that some LA students may need to be rectified in their planning strategies (as these are a negative predictor), it may also be beneficial to provide explicit coaching to LA students to help them use such strategies in the right way.

#### 4.2.3 Other predictors (and non-predictors)

Somatic anxiety (e.g., nervousness and tensions) and avoidance behaviour in CL writing negatively predicted the post-CBW scores of HA students, but not students in other bands. However, cognitive anxiety (e.g., worrying about being evaluated or getting bad grades), related to the anxiety caused by extrinsic goal orientation, did not predict the scores. The three learning motivation subscales taken from the MSLQ instrument, namely, intrinsic goal orientation, extrinsic goal orientation, and task value in CL writing, also did not predict scores. Thus, the students' level of motivation to learn CL writing might not have played a notable role in their CBW performance.

The group factor (with varying numbers of CBW sessions over the year) is not a predictor of the post-CBW scores of any student band. This finding echoes the results of our analysis in section 4.1, implying that the assumption that "(more) practice makes perfect (better scores)" does not hold. It is more important to practice under the right conditions - and these conditions should be determined holistically, for example through our analyses to identify the right combination of predictors of eventual CBW performance.

# 5. Proposed design model for an online writing community platform

Informed by our findings, we propose a design model for an online writing community platform known as WeeWrite  $(\# {\it S} {\it I} {\it R} {\it R})$ . We envision an online learning environment that provides systematic coaching and opportunities for CL e-writing for communication. The proposed platform will harness generative AI (particularly chatbots empowered by Large Language Models (LLMs)) and other technologies to deliver writing pedagogy and facilitate personalised and collaborative learning.

The WeeWrite design model consists of a personal learning space and a social learning space. A natural class of students (or beyond) can form a learning community, which is configurable by their teacher. When a student registers for an account, a chatbot known as MeWrite pal is created for them, and will keep them company when they stay on the platform. The MeWrite pal is an educational version of an LLM-empowered chatbot such as ChatGPT, which is reconfigured to be proactive in coaching and supporting the student to learn and e-write in the platform (rather than ChatGPT or alike which always "sit back" and wait for user's next prompt), and providing instant evaluation on their e-writing. Furthermore, it is reconfigured in the way that it will only respond to student's prompts written in CL (with a grace period at the beginning which allows students to occasionally "switch code" to EL while building their motivation to stay in the environment), and might refuse to entertain certain types of student requests which do not conducive to their learning of e-writing, such as auto-generating a composition from scratch.

The personal learning space is meant for the student to be coached in various skills involved in e-writing and for them to draft their e-essays or e-messages with the MeWrite pal's guidance or support. For example, amid the student's e-writing process, the MeWrite pal could recommend relevant words to include based on the writing context (akin to such affordances in Microsoft Office CoPilot) – this can be considered as an automated version of **text generation strategies**, which may be most beneficial to MA students according to our findings. Yet such recommendations should not be too frequent (or they will prevent the student from developing the ability to e-write in their own words) and should be treated as a scaffold that should fade out when the student's e-writing competence has improved.

Furthermore, various e-coaching modules for the pinyin scheme, CL typing, and the five types of writing strategies will be developed and incorporated into the platform. Yet they will be activated by the MeWrite pal to the student based on different priorities and orders according to the student's band (given the implications drawn in section 4.2, informed by our findings). For example, a LA student might first be tested in **the** *pinyin* **scheme and CL typing speed**. If the results indicate that additional coaching and practice are needed, the MeWrite pal will activate the corresponding e-coaching packages. Next, the student will be coached in **resourcing strategies** (and the coaching package will be linked to an online dictionary with the same content as the e-dictionary device they use in their physical composition sessions) and followed by planning strategies in e-writing. For example, they might be asked to draft a short email by applying specific strategies, and instant formative feedback on their draft (e.g., in terms of the six criteria in the rubric we used for scoring student compositions in our study) will be given; and they will be encouraged to improve it accordingly. Additionally, the MeWrite pal could be their vocabulary and grammar coach, helping them catch up with their peers in other bands in these areas.

For MA students, the learning pathway could start with **CL typing**, followed by **text generation** and **planning strategies**. For HA students, the initial focus could be placed on **monitoring/evaluation strategies**.

Depending on their learning progress, all students may eventually be coached in all strategy types, but the learning pathways could be personalised. Furthermore, as HA students are adept at e-writing as suggested by our findings, we advocate going one step further by developing their critical and creative e-writing skills – which can be guided by the meWrite pal with more advanced coaching modules.

Social learning activities can take place and be intertwined with individual learning tasks. Students can enter the social learning space at any time to read and comment on their peers' e-writing. They can create and share social media, perhaps with a requirement to use certain target words or sentence structures, or on certain topics – such requirements could be set by the teacher to link to the formal teaching; and the requirements would be delivered by the MeWrite pal to individual students. Furthermore, after being evaluated by the MeWrite pal, students' practice write-up in their individual spaces may be selected for sharing in a special spotlight bulletin in the social space. The goal is to foster their sense of pride and create a 'reader effect' that encourages them to further enhance their e-writing skills. The MeWrite pal may also alert the student to check the social learning space when there are interesting posts or teacher-facilitated collaborative e-writing events going on.
The proposed online platform is not meant to replace the formal writing instruction in CL lessons but to complement it. It serves the purpose of extending the students' learning to e-write beyond the school sessions, and offering more opportunities for students to access a greater range of ICT affordances, many of which are adaptive, thus facilitating self-directed and personalised learning (as well as collaborative learning in the social learning space). We envisage that such a platform will realise what the **negative predictors of "avoidance behaviour" and "somatic anxiety"** imply, as the environment is low-stakes – their e-writing products on the platform are not formally assessed, and they have no time pressure to write, unlike in time-limited class composition sessions or examinations.

#### 6. Conclusion and future directions

This study advances our understanding of the impact of incorporating CBW in CL education and sheds light on the varying conditions under which students benefit from such integration. Our findings suggest that CBW can significantly improve writing quality, challenging traditional assumptions. For MA and LA students, enhancing CL typing skills is crucial to harness CBW's potential. Different writing strategies are vital for distinct student bands, offering a roadmap for targeted instructional strategies.

Moreover, our proposed design model for WeeWrite as an online writing community platform capitalises on advanced technologies, particularly generative AI, to provide a personalised and collaborative learning environment. By leveraging AI-driven chatbots, students can engage in continuous practice, receive real-time feedback, and collaborate with peers. This platform aligns with the 21st-century learning landscape, fostering CL CBW competences that extend beyond the classroom.

The study, however, is not without limitations. Firstly, it was conducted within specific schools in Singapore, resulting in relatively limited sample sizes when the subjects were split into the three student bands for analyses. Factors beyond our scope, such as variations in teaching methods, teacher expertise, and students' individual writing practice outside class, were not addressed. Thus, the generalisability of the findings may be constrained, both in terms of the research context and the extent to which the results can be extrapolated.

Moreover, the study did not control the writing prompts, which could have exerted diverse influences on students' performance. A critical challenge arose from the reluctance of participating schools to relinquish their school-based components of the writing curriculum. This precluded us from controlling year-round composition sessions. The schools would likely have shown greater willingness to follow our design if data collection were confined to one or two CBW sessions.

Nonetheless, the study's merits lie in demonstrating that young students can perform well in both CBW and PPBW given exposure to the former. As a result, future investigations should shift their focus from comparing CBW and PPBW to exploring the conditions (predictors) conducive to improved CBW performance. Instead of pursuing longitudinal interventions, an alternative approach could involve a substantially larger sample size, encompassing students with varying levels of competence in CL typing or CBW skills, across more schools. This modified strategy, incorporating only one or two rigorously controlled composition sessions, would circumvent schools' resistance to year-round control. Such an approach would offer an avenue to verify or refine our findings concerning the optimal conditions for enhanced CBW performance among different student bands.

## References

Bai, R., Hu, G., & Gu, P. Y. (2014). The relationship between use of writing strategies and English proficiency in Singapore primary schools. *Asia-Pacific Education Researcher*, 23(3), 355-365.

Bailey, D. R., Lee, A. R., Vorst, T. M., & Crosthwaite, P. (2017). An investigation of differences and changes in L2 writing anxiety between blended and conventional English Language learning context. *CALL-EJ*, 18(1), 22-39. Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Journal*, *100*(1), 320-340.

Bernhardt, S. A., Wojahn, P. G., & Edwards, P. R. (1989). Teaching college composition with computers: A timed observation study. *Written Communication*, *6*, 108-133.

Bolter, D. (1991). Topographic writing: hypertext and the electronic writing space. In P. Delany & G. Landow (Eds.), *Hypermedia and Literary Studies* (pp. 105-118). MIT Press.

Burley, H. (1994). *Postsecondary novice and better than novice writers: The effects of word processing and a very special computer assisted writing lab* Southwestern Educational Research Association Meeting, San Antonio, TX.

Cheng, Y. S. (2004). A measurement of second language writing anxiety: Scale development and preliminary validation *Second Language Writing*, *13*, 313-335.

Cheung, Y. L. (2012). Critical review of recent studies investigating effects of word processingassisted writing and pen-and-paper writing on the quality of writing and higher level revisions. *Procedia - Social and Behavioral Sciences*, *46*, 1047-1050.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290.

Dalton, D. W., & Hannafin, M. J. (1987). The effect of word processing on written composition. *Journal* of *Educational Research*, *80*, 338-342.

Foroutan, M., & Noordin, N. (2012). Effect of dialogue journal writing through the use of conventional tools and e-mail on writing anxiety in the ESL context. *English Language Teaching*, 5(1), 10-19.

Greenhow, C., Robelia, B., & Hughes, J. E. (2009). Web 2.0 and classroom research: What path should we take now? *Educational Researchers*, *38*(4), 246-259.

Kang, H. (2011). Computer-based writing and paper-based writing: A study of beginning-level and intermediate-level Chinese learners' writing The Ohio State University]. Columbus, USA.

Kurt, G., & Atay, D. (2007). The effects of peer feedback on the writing anxiety of prospective Turkish teachers of EFL. Journal of Theory and Practice in Education & Treatment of Children, 3(12-23).

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38(4), 963-974.

Lam, F. S., & Pennington, M. C. (1995). The computer vs. the pen: A comparative study of word processing in a Hong Kong secondary classroom. *Computer Assisted Language Learning*, 8, 75-92.

Li, J. (2006). The mediation of technology in ESL writing and its implications for writing assessment. *Assessing Writing*, *11*, 5-21.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ). University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning.

Tushara, E. (2023). Laptops replace pen and paper as S'pore shifts towards electronic examinations. *The Straits Times*. <u>https://www.straitstimes.com/singapore/laptops-replace-pen-and-paper-as-</u> <u>s-pore-shifts-towards-electronic-examinations</u>

Wolfe, E. W., & Manalo, J. R. (2004). Composition medium comparability in a direct writing assessment of non-native English speakers. *Language Learning & Technology*, 8(1), 53-65.

Wong, L.-H., Chai, C.-S., & Gao, P. (2011). The Chinese input challenges for Chinese as second language learners in computer mediated writing: An exploratory study. *The Turkish Online Journal of Educational Technology*, 10(3), 233-248.

Xie, T. (2001). *The digitalization of Chinese teaching and learning: The use of vocabulary processing software and its impact in teaching and learning* International Conference on Internet Chinese Education 2001, Taipei, Taiwan.

Zhang, P. N. (2021). Typing to Replace Handwriting: Effectiveness of the Typing-Primary Approach for L2 Chinese Beginners. *Journal of Technology & Chinese Language Teaching*, *12*(2), 1-28.

Zhu, Y., Shum, S.-K. M., Tse, S.-K., & Liu, J. J. (2015). Word-processor or pencil-and-paper? A comparison of students' writing in Chinese as a foreign language. *Computer Assisted Language Learning*, 29(3), 1-22.

Zoghi, M., & Reshadi, E. (2013). Computer-based vs. paper-based writing tests: A comparative study of test taking strategies. *International Journal of English Language Education*, *1*(3), 43-56.

# Reader Characteristics, Reading Order, and Facial Emotions Expressed in Reading Science Texts

Yu-San HSIANG<sup>a</sup>, Zheng-Hong GUAN<sup>a</sup> & Sunny S.J. LIN<sup>a\*</sup> <sup>a</sup>Institute of Education, National Yang Ming Chiao Tung University, Taiwan \*sunnylin.nctu@gmail.com

Abstract: Emotion plays an important role in online reading and self-regulated learning. It could influence reading processing and even the outcomes through motivation functions. Since previous studies mostly adopted questionnaires to access readers' emotions, research that applied facial emotion detection non-intrusively during reading process is still limited. Therefore, the purpose of this study was to investigate how readers' on-going emotions fluctuated under different reading conditions, i.e., with different reader characteristics and reading orders, using a facial emotion analysis system. 21 participants were recruited and presented with four conflicting scientific texts that included positive or negative arguments on genome editing. These participants were divided into the high and low engagements and then were assigned different reading orders of positive-first or negative-first articles. The results showed that participants who read negative-first expressed more sadness and less happiness. Meanwhile, the high engagements showed more anger when reading positive-first, but the low-engagements showed more anger when reading the negative-first. Reading negative-first performed better, but writing performance did not differ according to the group membership or reading order. These findings suggest that (1) the influence of emotions on online reading multiple and conflict science texts and performance is complex; negative activating emotions like sadness and anger could be beneficial to reading performance, while positive activating emotions like happiness could be harmful for reproductive performance. (2) Using facial expression to non-intrusively detect readers' momentary emotions provides meaningful insights into the fluctuation of emotion during reading science texts. With only six basic emotions, it may not be enough to fully show the motivational transitions in reading process. The interpretation of these facial emotions should be considered beyond their surface meaning.

**Keywords:** Facial expressions, emotions, reader characteristics, reading order, scientific article

## 1. Introduction

Self-regulated learning (SRL) has been found an important element in distance or online learning situations (You & Kang, 2014). Reading multiple texts online was also one of these situations that required learners' self-regulation to persist and to overcome challenges during learning processes. Emotions, which was relating to learners' goal setting, motivation, and cognitive strategies used, can be vital to self-regulation and persistence in SRL (Muis et al, 2018). Meanwhile, when reading multiple scientific articles with conflicting arguments and controversial topics, emotion also profoundly influence learners' understanding through the allocation of cognition resources during processing (Danielson et al., 2022; Pekrun, 2022; Trevors, 2021). Thus, in this study, we aimed to understand how students' emotional states change during a scientific reading task in an online environment.

Traditionally, emotions could be categorized according to two dimensions: valence (positive or negative) and activation (activating or deactivating). Believed to be a subjective experience, previous studies adopted self-report questionnaires to investigate students'

emotions in learning situations. For instance, the Achievement Emotions Questionnaire (AEQ) as well as the Epistemically-Related Emotion Scales (EES) both developed by Pekrun and colleagues (Pekrun et al., 2005; Pekrun et al., 2017b) were two renowned and frequently used questionnaires. Based on self-report data of emotions, previous research found an average positive effect of positive emotions on learning outcomes, and effect of negative emotions were generally negative (Lajoie et al., 2021; Pekrun, 2022).

However, despite the numerous findings, self-report questionnaires have several limitations. On the one hand, although questionnaires are able to measure students' anticipatory and retrospective feelings, their on-going emotion states and the changes of emotions during the tasks were hard to collect. On the other hand, sometimes participants may not be able to tell how they feel or they may conceal their real opinion, so there may be discrepancies between the data we obtained and the real situations (Mason et al., 2018; Wang et al. 2021). While some researchers incorporated multiple time point measurement to capture the changes of emotions through questionnaire (Tulis & Fulmer, 2013), others adopted physiological or behavioral measures like hart rate (HR) and facial expression recognition tools to collect participants' expressive emotional states (Mason et al., 2018; Lajoie et al., 2021; Wang et al, 2021). Through these multi-modal data, different patterns of emotion fluctuations have been observed. For example, in Tulis & Fulmer (2013), they found that a moderate level of anxiety may motivate learners to exert more effort and to persist in a challenge task, which suggested that negative emotions could also be beneficial in learning process. Meanwhile, research also found that positive emotions may be harmful in some situations. Experiencing happy mood may distract students from the contents, but rather focused on irrelevant aspects of the learning task (Mensink, 2021; Trevors et al., 2017). That is to say, to examine how exactly emotions change during students' learning process, as well as to understand its relation with students' motivation and learning outcomes, tools that could capture on-going emotional data were essential. Previous research suggested that facial expressions were significantly correlated to students' real emotion, which could further be use to recognize their comprehension toward the learning content (Sathik & Jonathan, 2013).

Thus, in this study by combining multimodal analytic tools that can reveal temporal changes with those that could compare group differences, we aimed for a closer look at the roles of emotions playing in reading conflict scientific articles. Specifically, we utilized an automated facial emotion recognition system, the Facial Emotion Analysis Tool (FEAT, Lin et al., 2019; Wu & Lin, 2018) to record participants' emotions during reading tasks. FEAT detects different facial muscle movements to identify facial expressions based on the Facial Action Coding System (FACS, Ekman and Friesen, 1978) and categorized these expressions into six basic emotions (i.e., happiness, sadness, anger, disgust, fear, and surprise, Ekman, 1992). Readers were grouped according to their characteristics (reading interest and self-efficacy) and were assigned with articles of various orders (negative argument first or positive first) so that group comparisons could be observed.

We proposed the following two research questions to examine the interplay of students' characteristics, their expressed emotions, and text features. The relations between participants' expressed emotion and post reading performance will also be discussed.

RQ1: When reading positive or negative texts in different order, what types and patterns of emotions do readers with different characteristics exhibit?

RQ2: Do readers' expressed emotions differ significantly while reading negative or positive argument articles, according to their personal characteristics and the reading order?

## 2. Methods

#### 2.1 Participants

Twenty-one undergraduate and graduate students (16 female, 76%,  $M_{age}$  = 24, SD = 2.55) in a university in the northern part of Taiwan participated in the study. Participants were instructed to read four articles about genome editing babies. A facial expression analysis tool (FEAT) was used to record participants' emotion during reading task. Two participants' data

were excluded from the following analysis since their emotional responses were too low (most of the intensity of emotion were below the 0.1 criteria). The final sample consisted of three male and 16 female participants with an average age of 23.89 (SD = 2.66). The age of the participants followed a normal distribution and there was no significant difference between the ages of the four groups (F(3,15) = 1.905, p = 0.17).

#### 2.2 Procedures

The entire experiment was consisted of four sessions. In the first session, participants' prior knowledge, interest toward genome editing technologies, self-efficacy in reading multiple documents, and basic demographic information were accessed. Then, the camera was calibrated to ensure that participants' facial expression could be captured before they started reading the articles on a computer. Each article could only be read once. Participants were allowed to read each article according to their speed. While they were reading, their facial expressions were collected automatically. In the next session, participants completed a writing task and a multiple-choice questions test. Finally, an interview has been conducted to explore the cause of participants' emotional changes during the reading task. Line graphs depicting participant's emotional changes during the reading task were presented to participants, and they were asked to recall retrospectively what they were thinking or feeling during the task. Participants also reported the difficulty of each article in the interview.

## 2.3 Reading Materials

All participants need to read four articles about genome editing technology, including two expository texts, and two stories. In the expository texts, benefits and drawbacks of applying genome editing were discussed; while successful and failed cases of genome editing babies were presented in stories. The words in each text were around 351 words. Participants were randomly assigned to two groups which read these articles with different orders. The PN group read the positive expository or positive story first, and the NP group were first presented with negative expository or negative story.

#### 2.4 Measures

#### 2.4.1 Interest

Participant's interest was measured with self-developed questionnaire. Participants responded to statements from 1 (strongly disagree) to 6 (strongly agree). A higher score indicated a stronger interest/ higher self-efficacy. The interest questionnaire contains six questions (e.g., "I am interested in genome editing related topics", and "learning more about genome editing related topics is meaningful to me"). The reliability coefficients for overall questionnaire was acceptable (Cronbach  $\alpha = .79$ ).

## 2.4.2 Self-efficacy

Participant's self-efficacy was measured with self-developed questionnaire. Participants responded to statements from 1 (strongly disagree) to 6 (strongly agree). A higher score indicated a higher self-efficacy. The self-efficacy questionnaire contains eight questions (e.g., "I can understand the hardest part while reading multiple documents", "I am confident that I can understand the contents of article from different sources"). The internal consistency for the self-efficacy questionnaire was good (Cronbach  $\alpha$  = .93).

## 2.4.3 Behavioral Emotion Responses

Participants' facial expression were recorded and analyzed using Facial Emotion Analysis Tool (FEAT) (Lin et al., 2019; Wu & Lin, 2018). FEAT is a non-intrusive tool which can detect participants' facial expressions and automatically classify them into one of the six

basic emotions. Since it could learn in batches, this facial recognition tool was appropriate for real-time systems. The average accuracy of FEAT is around 90%. While comparing with other deep learning architectures (i.e., GoogLeNet, AlexNet, and CNN), our recognition model had better performance (Wu & Lin, 2018).

The sampling rate of FEAT was 30 frames per second, which means that emotion data would be generated once every 30 milliseconds. The log data generated by the system that has been used in this study includes (1) timestamp, and (2) intensity of six emotions. The intensity of the six emotions ranged from 0 to 1.

#### 2.4.4 Reading Comprehension Performance

In this study, two post reading assessments, including a multiple-choice test (post-test) and a writing task, were used. The post-test contains six multiple-choice questions, through which participants' memory toward articles were accessed. The writing task ask the participants to summarize what they have read about genome editing babies, whether they support or is against the use of genome editing technics and provide arguments for their position. Three researchers separately rated the written responses considering whether they demonstrated two-sided reasoning, elaborated from personal point of view, integrated points from different articles, and the evidences used. A total score was provided to each participant's writing response. The inter-rater reliability (the Kendall's coefficient of concordance; W) of the ratings was 0.93. The discrepancies in ratings were discussed to reach a consensus.

## 2.5 Statistical Analysis

Data analysis was performed in R environment (version 4.1.0; R Core Team, 2021) and divided into four phases. Firstly, the hierarchical cluster analysis with Ward method was used to group participants based on their characteristics. The dendrogram and elbow method were employed to determine the number of the groups. In the second phase, participants' emotion data generated by FEAT were examined and cleaned to form main dataset. Next, emotion states from main dataset have been used to draw line graphs according to article types and participants' characteristics. From the line graph, changes of participants' emotion during reading task can been seen. Although FEAT identified six basic emotions (i.e., happiness, sadness, anger, surprise, disgust, and fear), the intensity for surprise, disgust and fear were low in this study, which means participants rarely showed these emotions while reading. Therefore, in the following sessions, the results presented and the discussion would focus on happiness, sadness, and anger. Finally, we used the 'Ime4' package (Bates et al., 2015) to analyzed the intensity of emotion because LMM could include the random effects and fixed effects. In this study, the participants were regarded as random effect while the emotion, different groups and reading order were regard as fixed effects (Baayen et al., 2008). All the fixed effects were categorical variables. The 'anova' function was used to get F test result and the 'emmeans' package was used to conduct post-hoc comparison for simple main effect (Lenth, 2021).

## 3. Results

## 3.1 Cluster Analysis to Group Readers as the High and Low Engagements

Participants were divided into two clusters through the hierarchical cluster analysis with Ward method according to two variables, i.e., participants' interest and self-efficacy. Based on elbow method and dendrogram (see Figure 1 and Figure 2), the results suggested a solution of two clusters. The first cluster was made up of 11 participants who have lower interest toward the topic of genome editing and lower self-efficacy on reading multiple documents. The second cluster consisted of 8 participants who showed higher interest and self-efficacy. Therefore, the first cluster was named low engagement group while the second cluster was named high engagement group.



Figure 2. Dendrogram for clusters

Except for interest and self-efficacy ( $F_{\text{interest}}(1,17) = 19.68$ , p < .01;  $F_{\text{self-efficacy}}(1,17) = 25.04$ , p < .01), participants' age, prior knowledge, and performance in post-test did not differ between the two groups. Descriptive statistics were presented in Table 1.

	Clu	Clusters of Reader Characteristics				
	Cluster 2	1 (N = 11)	Cluster 2	2 (N = 8)	ANO	/A
	Low engag	ement group	High engage	ement group		
Variables	Mean	SD	Mean	SD	F	р
Age	24.45	2.91	23.12	2.23	1.16	= .30
Interest	16.00	3.13	22.00	2.56	19.68	< .01
Self-efficacy	28.91	4.37	38.25	3.45	25.04	< .01
Prior Knowledge	41.82	14.01	47.50	10.35	0.94	= .35
Writing task	46.36	12.06	46.25	10.61	0.00	= .98
Post-test	7.64	2.16	8.38	2.67	0.45	= .51

Table 1. Descriptive statistics results by cluster

Moreover, participants were also compared according to their reading order, participants who read negative article first (i.e., negative expository/story) were labeled NP group (N = 8), while participants who read positive article first (i.e., positive expository/story) were labeled PN group (N = 11). Participants from different reading order groups did not differ in terms of their interest, self-efficacy, prior knowledge and performance on writing task, but they differed significantly on post-test score (F(1,14) = 10.93, p < .01). Descriptive statistics were presented in Table 2.

 Table 2. Descriptive statistics results by reading order

	Reading order					
	Negative a NP (N	article first I = 8)	Positive a PN (N	rticle first = 11)	ANG	AVC
Variables	Mean	SD	Mean	SD	F	р
Age	24.00	3.30	23.82	2.27	0.08	= .79
Interest	20.25	3.65	17.27	4.20	2.59	= .13

Self-efficacy	34.50	5.15	31.64	6.74	1.01	= .33
Prior Knowledge	43.75	11.88	44.55	13.68	0.02	= .90
Writing task	8.12	2.64	7.82	2.23	0.08	= .79
Post-test	53.75	5.18	40.91	11.36	10.96	< .01



*Figure 3.* Over Time Changes of Six Basic Emotions (in Various Colors) for the Low and High Engagements in the Reading Order of NP (Negative Argument Article First) versus PN (Positive Argument Article First).

#### 3.2 Emotion Transitions Shown During Reading

In order to explore what kinds of emotions do reader expressed during reading task, and the fluctuation of these emotions, participants' emotion states were aggregated by groups and presented with time and intensity. Different patterns of emotion changes shown by different groups were shown in Figure 1. Reading order were labeled NP and PN, presented on the left vs. right side. The low engagement group (low interest/self-efficacy) was presented at the upper part while the high engagement group (high interest/self-efficacy) was presented at the lower part in Figure 3.

If we compare the differences between reading order (i.e., NP or PN), participants who read negative articles first exhibited greater fluctuation in their emotions. To be more specific, readers from NP group showed emotions with intensity greater than 0.25, while PN group readers' emotion states were within the level of 0.25. In addition, readers from NP group expressed obviously more sadness and less happiness while reading.

On the other hand, if readers with different characteristics were compared, low engagement group showed more anger and less happiness. Specifically, these readers continued to show anger throughout the experiment, and the intensity of their anger was above 0.1 thresholds at most of the time; on the contrary, they rarely showed happiness, and the intensity was mostly within the level of 0.1.

#### 3.3 Group Differences on Expressed Anger, Happiness, and Sadness

To further investigate whether participants' expressed emotion significantly differed with each other, a linear mixed model was conducted with reader characteristics (2: high vs.

low engagement), reading order (2: Negative first, NP vs. Positive first, PN), and emotions (6: happiness, sadness, anger, surprise, disgust, and fear) as independent variables. The three-way interaction effect was significant (F = 1414.87, p < .01). Table  $3\sim5$  shows the simple main effect for each emotion.

As indicated in Table 3, when fixing the level of reading order, low engagement group reading the order of NP expressed more anger than high engagement group (b = -0.127, z = -10.48, p < .01). However, if reading positive article first (PN), the result shows the opposite: high engagement group showed more anger than low engagement group (b = 0.024, z = 2.15, p < .05). On the other hand, when fixing the level of reader characteristics, high engagement group showed more anger in reading the positive article first compared to reading the negative article first (b = -0.070, z = -5.75, p < .01); while low engagement group expressed more anger when the reading order was NP (b = 0.081, z = 7.22, p < .01).

Fixed level	Comparison		Ang	ger	
		b	SE	Z	р
NP (Negative article first)	High-Low	-0.127	0.012	-10.48	< .01
PN (Positive article first)	High-Low	0.024	0.011	2.15	< .05
High engagement group	NP-PN	-0.070	0.012	-5.75	< .01
Low engagement group	NP-PN	0.081	0.011	7.22	< .01

Table 3. Model of 3-factor interactions on Anger

Statistic results of participants' expressed happiness were presented in Table 4. When fixing the reading order as NP, high engagement group showed more happiness than low engagement group (b = 0.026, z = 2.16, p = .03). Similarly, high engagement group also showed more happiness when the reading order was PN (b = 0.202, z = 17.86, p < .01). In addition, for both high and low engagement readers showed more expression of happiness when reading positive article first ( $b_{high} = -0.235$ , z = -19.21, p < .01;  $b_{low} = -0.059$ , z = -5.34, p < .01).

Fixed level Comparison		Happiness			
		b	SE	Z	р
NP (Negative article first)	High-Low	0.026	0.012	2.16	< .05
PN (Positive article first)	High-Low	0.202	0.011	17.86	< .01
High engagement group	NP-PN	-0.235	0.012	-19.21	< .01
Low engagement group	NP-PN	-0.059	0.011	-5.34	< .01

 Table 4. Model of 3-factor interactions on Happiness

Finally, Table 5 shows the results of participants' expressed sadness. When reading negative article first, low engagement group showed more sadness (b = -0.025, z = -2.02, p = .04). Likewise, they also showed more sadness than the high engagement group when the reading order was PN (b = -0.132, z = -11.76, p < .01). Meanwhile, both student groups showed more sadness when they read negative article first ( $b_{high} = 0.149$ , z = 12.23, p < .01;  $b_{low} = 0.041$ , z = 3.68, p < .01).

Fixed level	Fixed level Comparison		Sadness			
		b	SE	Z	р	
NP (Negative article first)	High-Low	-0.025	0.012	-2.02	< .05	
PN (Positive article first)	High-Low	-0.132	0.011	-11.76	< .01	
High engagement group	NP-PN	0.149	0.012	12.23	< .01	
Low engagement group	NP-PN	0.041	0.011	3.68	< .01	

## 4. Discussion and Conclusion

#### 4.1 The Interplay of Reader Characteristics, Reading Order, and Emotions

In this study, we aimed to investigate how students' emotional states changed during self-regulated reading process. Self-directed learning in an online situation has become more and more common for today's learners. Thus, understanding how students' emotion fluctuated during learning task is important since emotion was critical to students' persistence and self-regulation in online learning environments. Moreover, in an online situation, students usually found controversial information with conflicting arguments. Learners' emotion toward these different articles, especially the scientific ones, largely influence how they construct meaning.

In our first research question, we examined the types and patterns of emotion changes caused by the interplay between reader characteristics and reading order, while the text topic was "genome editing." The text encompassed two subtexts: some descriptions and arguments to positively support the genome editing and others, negatively. An interesting result regarding reading order has been found. Readers who read negative articles first tended to exhibit a larger fluctuation in their emotions, especially in sadness. As reported in the interview session, when readers were suddenly exposed to the overwhelming negative impact of genome editing techniques, they felt "worried" about the applications to new generation. Some also reported that they felt an incongruence between their previous knowledge and the new information presented in the text, which made them "confused." Thus, the fluctuation may be due to the negative descriptions and arguments from the article. Previous studies (Mason, 2018; Mensink, 2021; Pekrun, 2022) suggested that the types of emotions, levels of arousal (intensity of these emotions), as well as the number of emotions experienced, all influence students' reading process and reading performance. Participants' performance in this study was partially confirmed with the previous results. On the one hand, the PN group, which experienced a higher number of emotions, performed lower than the participants in the NP group. On the other hand, the NP group, which experienced more intense negative emotions, outperformed the PN group. Our results confirmed the results of previous epistemic emotion studies. For example, Pekrun et al. (2017b) found that negative activating epistemic emotions such as anxiety, confusion, and frustration were positively correlated with learning strategies (e.g., critical thinking, metacognitive self-regulation, etc.). In addition, Tulis & Fulmer (2013) found that both positive and negative emotions could be beneficial for learning outcomes. That is, experiencing a low to moderate level of negative activating emotions may contribute to continuous engagement during difficult tasks. In contrast, Dever and colleagues (2022) found that experiencing more emotions during reading and assessment was associated with lower performance in reproductive tests due to the limited resources that have been spent to manage emotions. Thus, although the NP group experienced intense negative emotions, it could be that they adopted more analytical strategies, and since they spent less cognitive resources on emotion management, they performed better in the post test.

In our second research question, we compared whether the emotions expressed by participants differed by student group. As the result suggested, readers who read negative text first expressed more sadness and less happiness. It's probable that, after being exposed to the terrible outcomes caused by genome technology, readers would read the subsequent paragraphs with more worries, which would be portrayed as sadness. In the meantime, readers' happiness can decrease because of concerns about the application of this technique. On the contrary, readers who read positive texts first may maintain a neutral or even positive attitude toward the reading task and the genome editing technique; thus, they may be relaxed while reading, and their emotional states would also fluctuate with the ups and downs in the texts. As for anger, participants in the low engagement group expressed more anger if they read negative text first, while readers in the high engagement group expressed more anger when they read positive text first. According to participants' reflections in the interview, most of them reported that they did not feel angry but were concentrated on the reading task. Based on this interpretation, it's reasonable to see that readers from the low engagement group showed more anger if the reading order is NP, since they were prompted to be more serious. However, readers with high interest and self-efficacy showed more anger when the reading

order was PN. This result was contrary to our expectations. A possible explanation could be that high engagement readers in the PN group were relatively conservative when talking about the application of genome editing to humans. When reading articles on the benefits of this technique, they may feel more serious since the content is against their beliefs.

According to the results of this study, one thing could be inferred is that maybe the interpretation of facial expression should be done with more caution. As suggested above, participants' anger and sadness should not be explained literally. Most of the time, participants did not feel anger or sadness, although the facial expression has been identified as that. According to the reports in the interviews with participants, anger was more like a focused state or feeling serious about the reading task, while sadness could be interpreted as participants' reflection on the content of the reading. A distinctive difference between facial expression and participants' feelings showed in anger. Anger was usually categorized as a negative activating emotion, but seriousness was a relatively positive deactivating emotion. The function of these emotions in learning was also absolutely different. Thus, when using facial emotion detection tools, multi-modal data is highly recommended. By combining self-report questionnaires with physiological mechanisms (e.g., pupil dilation, skin response, heart rate, etc.) or behavioral responses (facial expressions, body position, etc.), a clearer and more accurate understanding of the emotions expressed by the participants could be gained.

#### 4.2 Conclusion and Limitations

Overall, this study addressed students' momentary emotions during reading conflicting scientific articles. Line graphs were used to depict the fluctuations of emotions in temporal dynamics. We discovered that reader characteristics (interest and self-efficacy) would interact with text features (positive-negative argument order), causing different patterns of emotional responses. By combining the multimodal analytic tools that can reveal temporal changes with those that could compare group differences, we can have a closer look at the roles of emotions playing in reading conflict science articles. Despite the findings of this study, there are several limitations to be considered.

On the one hand, the sample size of this study is small, and most of them were graduate students. This may be the main reason why we did not find a significant difference in readers' writing performance, since graduate students were familiar with the process of integrating new information with prior knowledge and creating arguments to support their position. Future research can expand the participants to a more diverse sample and, with a larger sample size, reexamine whether the results still hold true. On the other hand, according to readers' reports in the interview session, the articles used in this study were not very difficult. Most of the participants only gave 1~2 points on a 6-point scale. Future research should adopt material with different levels of difficulty to see if students' emotion patterns are more apparent or if there is a different pattern when it comes to challenging tasks.

Finally, in this study, we found that techniques that can automatically capture students' facial expressions and record their emotion changes can provide unique insights into students' learning processes. The Facial Emotion Analysis Tool (FEAT) is one of these tools. However, the limitation of FEAT is that it currently only identifies the six basic emotions. The recognition of emotions should at least be expanded to include concentration and confusion. By doing so, we can get a clear picture of students' learning processes.

## References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59(4), 390-412. https://doi.org/10.1016/j.jml.2007.12.005

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Usinglme4. Journal of Statistical Software, 67(1). <u>https://doi.org/10.18637/jss.v067.i01</u>

- Dever, D. A., Wiedbush, M. D., Cloude, E. B., Lester, J., & Azevedo, R. (2022). Emotions and the comprehension of single versus multiple texts during game-based learning. *Discourse Processes*, *59*(1-2), 94-115. <u>https://doi.org/10.1080/0163853X.2021.1950450</u>
- Lajoie, S. P., Zheng, J., Li, S., Jarrell, A., & Gube, M. (2021). Examining the interplay of affect and self regulation in the context of clinical reasoning. *Learning and Instruction*, 72: 101219. https://doi.org/10.1016/j.learninstruc.2019.101219
- Lenth, R. V. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.6.1. <u>https://doi.org/10.1080/00031305.1980.10483031</u>
- Lin, S. S. J., Chen, W., Lin, C.-H., & Wu, B.-F. (2019). Building a Chinese Facial Expression Database for Automatically Detecting Academic Emotions to Support Instruction in Blended and Digital Learning Environments. In *Innovative Technologies and Learning* (pp. 155-162). https://doi.org/10.1007/978-3-030-35343-8 17
- Mason, L., Scrimin, S., Zaccoletti, S., Tornatora, M. C., & Goetz, T. (2018). Webpage reading: Psychophysiological correlates of emotional arousal and regulation predict multiple-text comprehension. *Computers in Human Behavior*, 87, 317–326. <u>https://doi.org/10.1016/j.chb.2018.05.020</u>
- Mensink, M. C. (2022). Emotional responses to seductive scientific texts during online and offline reading tasks. *Discourse Processes, 59*(1-2), 76-93. https://doi.org/10.1080/0163853X.2021.1918492
- Muis, K. R., Chevrier, M., & Singh, C. A. (2018). The role of epistemic emotions in personal epistemology and self-regulated learning. *Educational Psychologist, 53*(3), 165-184. https://doi.org/10.1080/00461520.2017.1421465
- Pekrun, R. (2022). Emotions in reading and learning from texts: Progress and open problems. *Discourse Processes, 59*(1-2), 116-125. <u>https://doi.org/10.1080/0163853X.2021.1938878</u>
- Pekrun, R., Goetz, T., & Perry, R. P. (2005). Academic Emotions Questionnaire (AEQ) User's manual. Munich, Germany: Department of Psychology, University of Munich.
- Pekrun, R., Lichtenfeld, E., Marsh, W. H., Murayama, K., & Goetz, T. (2017)a. Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development, 88*(5), 1653-1670. <u>https://doi.org/10.1111/cdev.12704</u>
- Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017)b. Measuring emotions during epistemic activities: the Epistemically-Related Emotion Scales. *Cognition and Emotion, 31*(6), 1268-1276. https://doi.org/10.1080/02699931.2016.1204989
- Sathik, M., & Jonathan, S. G. (2013). Effect of facial expressions on student's comprehension recognition in virtual educational environments. *SpringerPlus, 2*:455. https://doi.org/10.1186/2193-1801-2-455
- Tonguç, G. & Ozkara, B. O. (2020). Automatic recognition of student emotions from facial expressions during a lecture. *Computers & Education, 148*: 103797. https://doi.org/10.1016/j.compedu.2019.103797
- Trevors, G. J., Muis, K. R., Pekrun, R., Sinatra, G. M., & Muijselaar, M. M. L. (2017). Exploring the relations between epistemic beliefs, emotions, and learning from texts. *Contemporary Educational Psychology*, *48*, 116–132. https://doi.org/10.1016/j.cedpsych.2016.10.001
- Trevors, G. J. (2021). The roles of identity conflict, emotion, and threat in learning from refutation texts on vaccination and immigration. *Discourse Processes, 59*(1-2), 36-51. https://doi.org/10.1080/0163853X.2021.1917950
- Tulis, M. & Fulmer, S. M. (2013). Students' motivational and emotional experiences and their relationship to persistence during academic challenge in mathematics and reading. *Learning and Individual Differences*, 27, 35-46. <u>https://doi.org/10.1016/j.lindif.2013.06.003</u>
- Wang, Y., Zhang, J., & Lee, H. (2021). An online experiment during COVID-19: Testing the influences of autonomy support toward emotions and academic persistence. *Frontiers in Psychology*, 12: 747209. <u>https://doi.org/10.3389/fpsyg.2021.747209</u>
- Wu, B.-F., & Lin, C.-H. (2018). Adaptive Feature Mapping for Customizing Deep Learning Based Facial Expression Recognition Model. *IEEE Access*, 6, 12451-12461. https://doi.org/10.1109/access.2018.2805861
- You, J. W., & Kang, M. (2014). The role of academic emotions in the relationship between perceived academic control and self-regulated learning in online learning. *Computers & Education*, 77, 125–133. <u>https://doi.org/10.1016/j.compedu.2014.04.018</u>

# Educational Use of SVVR: A Scoping Review on Harnessing EduVenture-VR in Learning and Teaching

#### Youwen SHI

Faculty of Education, the Chinese University of Hong Kong, Hong Kong China \*youwenshi@link.cuhk.edu.hk

**Abstract:** Virtual reality (VR) technology creates immersive user environments and has diverse applications, including education. A subset of VR, called spherical video-based virtual reality (SVVR), uses 360-degree videos to simulate realistic scenarios and can be easily accessed through mobile phones and low-cost head-mounted displays (HMDs). EduVenture-VR is an integrated SVVR learning platform used for SVVR-supported teaching in classrooms. While some studies have examined SVVR in education, there is a lack of research on the SVVR production platforms like EduVenture-VR. This scoping review aims to explore EduVenture-VR's current state and prospects in education, summarizing existing research design paradigms with collaborative learning elements, and proposing new research directions for future studies.

**Keywords:** SVVR, collaborative learning, EduVenture-VR

## 1. Introduction

Virtual reality involves interactive computer simulations detecting user input and modifying sensory feedback, creating a profound sense of immersion in a virtual environment (McGovern et al., 2020; Mihelj et al., 2014). Immersion, interaction, and imagination (3ls) are the three significant features of virtual reality (Burdea and Coiffet, 2003). Di Natale et al. (2020) distinguished between different levels of immersion, from non-immersive (desktop) VR, to semi-immersive (partial field of vision, eg, Fulldome, EMRELE, Smart Glasses) to immersive (eg, Cardboard, HMDs, CAVE). VR has gradually been integrated into education, for instance, computer science education (Agbo et al., 2021), patient education (Pandrangi et al., 2019), and physical education (Ding et al., 2020). However, there are numerous obstacles to the promotion and application of VR in education, with overhead expenses and the need for specialized training being the primary hindrances (Kavanagh et al., 2017). SVVR, a form of immersive video-based virtual reality, offers convenient accessibility through mobile phones and inexpensive cardboard goggles (Jong, 2022). Besides, advancements in technology, mobile applications, and 360-degree cameras enable the public to easily create their own 360-degree spherical images or videos with speed and convenience (Wu et al., 2021). It is characterized by its affordability, ease of use, and user-friendly nature, making it a suitable option for teachers seeking to provide immersive learning experiences with technological support, even if they are not highly proficient with technology (Chien et al., 2020).

Collaborative learning serves as an overarching framework encompassing diverse educational methodologies that entail collective cognitive endeavors among students, or between students and teachers. Typically, learners operate within groups, engaging in cooperative endeavors to jointly seek comprehension, resolutions, interpretations, or to collaboratively create a product (Laal & Ghodsi, 2012). The degree of active learner engagement during both the instructional design and implementation phases frequently influences the effectiveness of a learning intervention (Trentin, 2010). Thus, emerging technologies and platforms that facilitate online collaborative and interactive learning among students deserve further attention. EduVenture-VR is an integrated location-oriented

context-aware learning system launched in Hong Kong in 2016, consisting of EduVenture-VR Composer and EduVenture-VR App. The former enables teachers to design interactive outdoor learning materials without coding, using various question formats. Students use EduVenture-VR App to access the teaching materials in SVVR. Studies have shown positive impacts on student learning through SVVR-based teaching strategies on EduVenture-VR (Li, 2023; Liu, 2023; Chen, 2022; Huang, 2022; Li, 2022; Ruan, 2022; Wu, 2022; Chen, 2021; Huang, 2021; Lin, 2021; Chang, 2020). The paper, utilizing EduVenture-VR platform as a focal point, conducts a scoping review to encapsulate the research trends of EduVenture-VR and also briefly summarizes the design paradigms with collaborative learning elements in EduVenture-VR classes.

## 2. Methods

## 2.1 Protocol

This scoping review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for scoping reviews (PRISMA-ScR) guidelines (Tricco et al., 2018) and Arksey and O'Malley (2005)'s methodological framework to gather, select, summarize, analyze, and interpret evidence related to the review questions.

## 2.2 Eligibility Criteria

Included papers used EduVenture-VR for research, published between 2016-2023 in English, and related to education. It considered quantitative, qualitative, and mixed-method studies for assessing SVVR-supported approaches. Conference papers using EduVenture-VR and related to education were also included, with duplicate conference papers excluded.

## 2.3 Information Sources

Searches used two electronic databases: Scopus and Web of Science, to capture a comprehensive sample of literature for records published from 2016 to March 2023 in the English language.

## 2.4 Search terms

Search terms included 'eduventure-vr', 'spherical video based virtual', 'svvr', '360-degree video', '360° video', and 'panoramic video' with "OR" (Boolean relationship) to ensure that the studies met the inclusion criteria captured.

## 2.5 Charting, Collating, Summarizing, and Reporting the Results

After screening, 28 journal articles and conference papers are included for synthesis (see Figure 1 & Table 1). EduVenture-VR's usage for educational research has steadily increased since its launch in 2016. Current users are mainly from Mainland, Taiwan, and Hong Kong, and it has extended to international high school students in Thailand (Zhao & Yang, 2023). Implementation occurs across various education levels, particularly in universities and secondary schools, with Chinese language education and natural science being the main disciplines benefiting. EduVenture-VR is commonly used for short-term learning activities in schools, covering different SVVR experimental designs, mainly employing quasi-experimental methods (Wu et al., 2021). Experiential learning is the most common pedagogical use.



Figure 1. PRISMA-ScR flowchart of the study selection process.

Table 1. Literature	e List
---------------------	--------

	Title
First Author (Year)	I ITIE
Chang (2020)	The effects of spherical video-based virtual reality implementation on students' natural science learning effectiveness
Chen (2021)	Teachers' Conceptions of Teaching Chinese Descriptive Composition with Interactive Spherical Video-Based Virtual Reality
Geng (2021)	Understanding the pedagogical potential of Interactive Spherical Video-based Virtual Reality from the teachers' perspective through the ACE framework
Zhao (2023)	Promoting international high-school students' Chinese language learning achievements and perceptions: A mind mapping-based spherical video-based virtual reality learning system in Chinese language courses
Wu (2021)	A Spherical Video-Based Immersive Virtual Reality Learning System to Support Landscape Architecture Students' Learning Performance during the COVID-19 Era
Lin (2021)	Engaging university students in a library guide through wearable spherical video-based virtual reality: effects on situational interest and cognitive load, Interactive Learning

	Environments
Wu (2021)	Effects of the self-regulated strategy within the context of spherical video-based virtual reality on students' learning performances in an art history class
Jong (2022)	Flipped classroom: motivational affordances of spherical video-based immersive virtual reality in support of pre-lecture individual learning in pre-service teacher education
Wu (2022)	Effects of Multisensory Integration through Spherical Video-Based Immersive Virtual Reality on Students' Learning Performances in a Landscape Architecture Conservation Course
Chien (2020)	Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions
Jong (2020)	Integrating interactive learner-immersed video-based virtual reality into learning and teaching of physical geography
Li (2022)	Investigating Learners' Engagement and Chinese Writing Learning Outcomes with Different Designs of SVVR-Based Activities
Chen (2022)	Promoting deep writing with immersive technologies: An SVVR-supported Chinese composition writing approach for primary schools
Yang (2021)	From experiencing to expressing: A virtual reality approach to facilitating pupils' descriptive paper writing performance and learning behavior engagement
Huang (2021)	Facilitating decision making in authentic contexts: an SVVR-based experiential flipped learning approach for professional training
Huang (2022)	Technological solutions for promoting employees' knowledge levels and practical skills: An SVVR-based blended learning approach for professional training
Chen (2022)	Effects of experiencing authentic contexts on English speaking performances, anxiety and motivation of EFL students with different cognitive styles
Ye (2021)	Classroom misbehaviour management: an SVVR-based training system for preservice teachers
Ruan (2022)	VR-Assisted Environmental Education for Undergraduates
Lin (2021)	Supporting dyadic learning of English for tourism purposes with scenery-based virtual reality
Chen (2021)	Modeling learners' self-concept in Chinese descriptive writing based on the affordances of a virtual reality-supported environment
Chang (2020)	Integration of the peer assessment approach with a virtual reality design system for learning

earth science
Effects of applying a VR-based two-tier test
strategy to promote elementary students
learning performance in a Geology class
Effects of an article-structure strategy-based
spherical video-based virtual reality approach
on EFL learners' English reading
comprehension and learning conceptions
Motivating Ethnic Minority Students in Hong
Kong to Learn Chinese Culture with
EduVenture-VR
Pilot Study of Using ARCS Model to Investigate
Motivational Affordances of EduVenture-VR
from Perspective of Construction Management
Students
Students' Conceptions of Interactive Spherical
Video-based Virtual Reality Supported Chinese
Writing Learning
Educational Use of Spherical Video-based
Virtual Reality: A Preliminary Study from the
Teacher Perspective

The review identifies three design paradigms. Some studies use SVVR as an independent variable and assess its effectiveness by comparing an experimental group with SVVR and a control group without SVVR. Other studies use a teaching strategy as a variable and compare an SVVR approach with and without this teaching method. The remaining studies explore conceptions of SVVR in specific disciplines, particularly Chinese language education. For example, Chen et al. (2021) examined teachers' conceptions of SVVR in Chinese descriptive writing, while Chen et al. (2021) investigated students' self-concept in Chinese writing with SVVR. Li et al. (2022) found that a double-loop SVVR-based learning approach improved students' Chinese writing acquisition. Yang et al. (2021) found that SVVR enhanced writing performance but not creative thinking. Jong et al. (2021) explored ethnic minority students' motivation to learn Chinese culture with SVVR. Zhao and Yang (2023) applied a mind-mapping strategy to teach Chinese to non-native speakers and reported positive outcomes. Apart from that, in the most of SVVR-supported classes, students typically engage with SVVR-based learning materials in groups. During or prior to the experiential phase, teachers commonly assign pertinent questions aligned with the class objectives. Subsequently, during the discussion phase, students have the opportunity to share their experiential insights and collaboratively explore solutions to the inquiries with peers and teachers. In comparison to traditional teacher-centered classrooms, encouraging students to take a more proactive role in their learning endeavors.

## 3. Discussion and Future Research

The results show that EduVenture-VR is applied in diverse subjects and age groups, indicating broad development potential. However, in Chinese language education, research mainly focuses on descriptive writing for first-language speakers. Exploring teaching strategies for Chinese language education and expanding to non-native Chinese-speaking regions require further investigation. Additionally, there is a need for a new framework for formal education using SVVR, as most studies focus on informal learning. The LIVIE framework by Jong et al. (2020) exemplifies SVVR integration in formal geography education. Future research should design SVVR-supported teaching methods for specific disciplines, promote long-term classroom application, and engage students actively in the SVVR material design process. In addition, SVVR primarily serves as a means of delivering content and

factual knowledge through direct instruction, limiting its potential for interactive scaffolding and assessment within the video (Li et al., 2023). Furthermore, while some studies have systematically reviewed the educational value of SVVR (Rosendahl & Wagner, 2023), there remains a dearth of research exploring the interactive use of educational SVVR.

## 4. Limitation and Conclusion

The limitation of this paper is that this study only focuses on the EduVenture-VR platform and cannot comprehensively summarize the SVVR learning in education.

The aim of this scoping review is to explore the current research status of EduVenture-VR in education, briefly summarize the research design paradigms with collaborative learning elements, and identify potential research directions. The paper examines 28 journal articles and conference papers to synthesize the existing research on EduVenture-VR in education. Most studies use quasi-experimental designs to develop SVVR-based teaching approaches and evaluate their effectiveness in Chinese language education or natural science through short-term school activities. EduVenture-VR is applied across various education levels, with universities, secondary schools, and primary schools being the most common. The review summarizes the experimental design paradigms and proposes new research directions for EduVenture-VR, providing a basis for future empirical research in this area.

## References

- Agbo, F. J., Sanusi, I. T., Oyelere, S. S., & Suhonen, J. (2021). Application of Virtual Reality in Computer Science Education: A Systemic Review Based on Bibliometric and Content Analysis Methods. *Education sciences*, 11(3), 142.
- Arksey, H., & O' Malley, L. (2005). Scoping studies: towards a methodological framework. International *Journal of Social Research Methodology, 8*(1), 19-32.
- Burdea, G., & Coiffet, P. (2003). Virtual reality technology (2nd ed.). John Wiley.
- Chang, S. C., Hsu, T. C., Chen, Y. N., & Jong, M. S.Y. (2020). The effects of spherical video-based virtual reality implementation on students' natural science learning effectiveness. *Interactive learning environments*, *28*(7), 915-929.
- Chang, S. C., Hsu, T. C., & Jong, M. S. Y. (2020). Integration of the peer assessment approach with a virtual reality design system for learning earth science. *Computers and education*, 146, 103758.
- Chang, S. C., Hsu, T. C., Kuo, W. C., & Jong, M. S. Y. (2020). Effects of applying a VR-based two-tier test strategy to promote elementary students' learning performance in a Geology class. British *journal of educational technology*, *51*(1), 148-165.
- Chen, M., Chai, C. S., Jong, M. S. Y., & Chao, G. C. N. (2021). Modeling learners' self-concept in Chinese descriptive writing based on the affordances of a virtual reality-supported environment. *Education and information technologies*, *26*(5), 6013-6032.
- Chen, M., Chai, C. S., Jong, M. S. Y., & Jiang, M. Y. C. (2021). Teachers' Conceptions of Teaching Chinese Descriptive Composition with Interactive Spherical Video-Based Virtual Reality. *Frontiers in psychology*, 12, 591708-591708.
- Chen, M. R. A., & Hwang, G. J. (2022). Effects of experiencing authentic contexts on English speaking performances, anxiety and motivation of EFL students with different cognitive styles. *Interactive learning environments, 30*(9), 1619-1639.
- Chen, Y. T., Li, M., Huang, C. Q., Han, Z. M., Hwang, G. J., & Yang, G. (2022). Promoting deep writing with immersive technologies: An SVVR-supported Chinese composition writing approach for primary schools. *British journal of educational technology, 53*(6), 2071-2091.
- Chien, S. Y., Hwang, G. J., & Jong, M. S. Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Computers and education*, 146, 103751.
- Di Natale, A. F., Repetto, C., Riva, G., & Villani, D. (2020). Immersive virtual reality in K-12 and higher

education: A 10 - year systematic review of empirical research. *British journal of educational technology*, *51*(6), 2006-2033.

- Ding, Y., Li, Y., & Cheng, L. (2020). Application of Internet of Things and Virtual Reality Technology in College Physical Education. *IEEE Access*, 8, 96065-96074.
- Huang, H., Hwang, G. J., & Chang, S. C. (2021). Facilitating decision making in authentic contexts: an SVVR-based experiential flipped learning approach for professional training. *Interactive learning environments*, 1-17.
- Huang, H., Hwang, G. J., & Jong, M. S. Y. (2022). Technological solutions for promoting employees' knowledge levels and practical skills: An SVVR-based blended learning approach for professional training. *Computers and education*, 189.
- Jong, M. S. Y. (2022). Flipped classroom: motivational affordances of spherical video-based immersive virtual reality in support of pre-lecture individual learning in pre-service teacher education. *Journal of computing in higher education*, 1-22.
- Jong, M. S. Y., Jiang, M. Y. C., & Wong, P. S. (2022). Pilot Study of Using ARCS Model to Investigate Motivational Affordances of EduVenture VR from Perspective of Construction Management Students. 2022 IEEE 5th Eurasian Conference on Educational Innovation (ECEI).
- Jong, M. S. Y., Ng, N., Luk, E., Leung, J., Jiang, M. Y. C., Lau, D., & Tsai, C. C. (2021). Motivating Ethnic Minority Students in Hong Kong to Learn Chinese Culture with EduVenture VR. 29th International Conference on Computers in Education (ICCE).
- Jong, M. S. Y., Tsai, C. C., Xie, H., & Wong, K. K. F. (2020). Integrating interactive learner immersed video based virtual reality into learning and teaching of physical geography. *British journal of educational technology*, *51*(6), 2063-2078.
- Kavanagh, S., Luxton-Reilly, A., Wuensche, B., & Plimmer, B. (2017). A systematic review of virtual reality in education. *Themes in Science and Technology Education*, *10*(2), 85-119.
- Laal, M., & Ghodsi, S. M. (2012). Benefits of collaborative learning. *Procedia-social and behavioral sciences*, 31, 486-490.
- Li, F., Jiang, J., Qin, Q., Wang, X., Zeng, G., Gu, Y., & Guo, W. (2023). Application of Sustainable Development of Teaching in Engineering Education: A Case Study of Undergraduate Course Design of Raman Spectroscopy Based on Virtual Reality (VR) Technology. *Sustainability*, 15(3), 1782.
- Li, M., Chen, Y., Zhang, L., Wu, X., & Huang, C. (2022). Investigating Learners' Engagement and Chinese Writing Learning Outcomes with Different Designs of SVVR-Based Activities. *Sustainability*, 14(8), 4767.
- Li, W., Qian, L., Feng, Q., & Luo, H. (2023). Panoramic video in education: A systematic literature review from 2011 to 2021. *Journal of Computer Assisted Learning*, *39*(1), 1-19.
- Lin, H. C. S., Yu, S. J., Sun, J. C. Y., & Jong, M. S. Y. (2021). Engaging university students in a library guide through wearable spherical video-based virtual reality: effects on situational interest and cognitive load. *Interactive learning environments*, 29(8), 1272-1287.
- Lin, V., Barrett, N. E., Liu, G. Z., Chen, N. S., & Jong, M. S. Y. (2021). Supporting dyadic learning of English for tourism purposes with scenery-based virtual reality. *Computer assisted language learning*, 1-37.
- Liu, C. C., Guo, Y., Hwang, G. J., Tu, Y. F., & Wang, Z. (2023). Effects of an article-structure strategy-based spherical video-based virtual reality approach on EFL learners' English reading comprehension and learning conceptions. *Interactive learning environments*, 1-18.
- Mihelj, M., Novak, D., & Begu, S. (2014). Virtual Reality Technology and Applications. *Springer Netherlands.*
- Pandrangi, V. C., Gaston, B., Appelbaum, N. P., Albuquerque, F. C., Levy, M. M., & Larson, R. A. (2019). The Application of Virtual Reality in Patient Education. *Annals of vascular surgery*, 59, 184-189.
- Rosendahl, P., & Wagner, I. (2023). 360° videos in education A systematic literature review on application areas and future potentials. *Education and information technologies*.
- Ruan, B. (2022). VR-Assisted Environmental Education for Undergraduates. *Advances in Multimedia*, 2022, 3721301.
- Trentin, G. (2010). Networked collaborative learning: social interaction and active learning. Chandos.
- Tricco, A. C. et al. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine, 169*(7), 467-473.
- Wu, W. L., Hsu, Y., Yang, Q. F., & Chen, J. J. (2021). A Spherical Video-Based Immersive Virtual Reality Learning System to Support Landscape Architecture Students' Learning Performance during the COVID-19 Era. Land, 10(6), 561.
- Wu, W., Zhao, Z., Du, A., & Lin, J. (2022). Effects of Multisensory Integration through Spherical Video-Based Immersive Virtual Reality on Students' Learning Performances in a Landscape Architecture Conservation Course. *Sustainability*, 14(24), 16891.

- Wu, W. L., Hsu, Y., Yang, Q. F., Chen, J. J., & Jong, M. S. Y. (2021). Effects of the self-regulated strategy within the context of spherical video-based virtual reality on students' learning performances in an art history class. *Interactive learning environments*, 1-24.
- Yang, G., Chen, Y. T., Zheng, X. L., & Hwang, G. J. (2021). From experiencing to expressing: A virtual reality approach to facilitating pupils' descriptive paper writing performance and learning behavior engagement. *British journal of educational technology*, 52(2), 807-823.
- Zhao, J. H., & Yang, Q. F. (2023). Promoting international high-school students' Chinese language learning achievements and perceptions: A mind mapping-based spherical video-based virtual reality learning system in Chinese language courses. *Journal of Computer Assisted Learning*, 1-15.

## Exploring Student Engagement and Teacher-Student Interaction Patterns in Collaborative STEM PBL courses through Epistemic Network Analysis

**Ting-Yen KUO<sup>a\*</sup>, Ssu-Kuang CHEN<sup>a</sup> & Sunny S. J. LIN<sup>a</sup>** <sup>a</sup> Institute of Education, National Yang Ming Chiao Tung University, Taiwan <sup>1</sup> ats088266.hs11@nycu.edu.tw<sup>2</sup> schen75025@gmail.com<sup>3</sup>sunnylin.nctu@gmail.com

**Abstract:** To assess junior high school students' engagement in collaborative STEM project-based learning courses, this study employed a web-based observation protocol called COPIE-STEM to collect classroom observation data. This protocol coded data based on the ICAP theory and teacher instructional approaches. Over the course of six weeks, data on student engagement behaviors and teacher behaviors were gathered. In addition to using coding and counting methods, Epistemic Network Analysis (ENA) was employed to explore the trajectory of the STEM-PBL course during these six weeks. The results revealed that passive student engagement behaviors and teacher-centered behaviors clustered on the left side, while active student engagement behaviors and student-centered teacher behaviors clustered on the right side. Furthermore, examining the network graphs for each week's courses showed that as the course progressed, instruction shifted from teacher-centered to student-centered, and students exhibited higher levels of engagement. These findings provided insights into the dynamics of STEM-PBL classrooms.

**keywords:** STEM project-based learning, student engagement,instructional approach, Epistemic Network Analysis

## 1. Introduction

STEM project-based learning (STEM-PBL) is an educational approach that involves engaging students in real-world problems and encouraging them to investigate and solve these problems through multidisciplinary knowledge integration. By working collaboratively in groups, students can develop their problem-solving and critical-thinking skills while producing solutions or works (El Sayary et al., 2015). The effectiveness of STEM-PBL has been proven, and using this model for teaching can enhance learning outcomes (Saraç, 2018). In order to better understand the behavior of teachers and students, many classroom observation protocols have been developed for STEM-PBL courses (e.g., Smith et al., 2013; Hsiao et al., 2022). An important next step is analyzing data and depict a realistic picture that reflects what is happening in the classroom, assisting teachers and students to adapt in the PBL environment.

In the past, most of the data collected in classroom observation were coded and counted for descriptive analysis that unfortunately omits critical insights into how variables interacted and evolved (Csanadi et al., 2018; Rojas-Drummond et al., 2013). The dynamic aspect of how teachers conduct their teaching and how students engage in learning was not explored much. Therefore, this case study adopted the epistemic network analysis (ENA; Shaffer et al. 2009; Shaffer, 2017), which is an analytical method modeling behaviors in temporal framework, to understand STEM-PBL, particularly the interplay between instructional approaches and learning engagements.

## 2. Literature Review

## 2.1 Student Engagement and Instructional Approach

Student engagement includes behavioral, cognitive, and emotional involvement (Fredricks et al., 2004). Among various theories of student engagement, we find the ICAP theory (Chi, 2009; Chi & Wylie, 2014) to be the most suitable for characterizing student participation in classrooms. The ICAP theory pertains to students' cognitive engagement in learning activities, encompassing Interactive, Constructive, Active, and Passive modes, with a hierarchy of I > C > A > P. Chi et al. (2018) used this theory to design professional development programs for teachers, demonstrating that ICAP is relevant not only to students but also as a guide for teacher behaviors.

Based on the instructional design, the interactions between teachers and students can manifest as either teacher-centered or student-centered. Distinct student responses seem to emerge aligned with each instructional approach. In teacher-centered classrooms, the teacher serves as the transmitter of knowledge, while students are primarily passive listeners. In contrast, student-centered classrooms foster more active student engagement, with teachers playing a role in assisting students' learning (Serin, 2018). To sum up, this taxonomy of instructional approach and the ICAP framework form the foundation for categorizing student behaviors and teacher behaviors in this study.

## 2.2 Epistemic Network Analysis in Education

In previous related research, coding and counting were commonly used, with coding frequency as the basis of analysis. This approach treats behavior captured in a time interval as independent entities, thereby overlooking the connections between them (Swiecki et al., 2020). As a result, the sequential relationships between interactions among participants cannot be discerned. In contrast, Epistemic Network Analysis (ENA) combines qualitative and quantitative analyses, enabling to examine complex processes of human interaction. ENA's theoretical foundation originally lies in the epistemic frame theory (Shaffer, 2017), which views learning as the transformation of an individual's epistemic network manifested in discourse. This network evolves over time during and after the learning process.

Lately, the technique of ENA has been applied in different research domains. The most relevant to our study is the research of collaboration in classroom. For instance, Nachtigall & Sung (2019) conducted a study in an environment characterized by productive failure. In this study, students attempted to solve problems on their own before receiving delayed instruction. Subsequently, ENA was employed to compare the differences in problem-solving approaches between high quality (HQ) solution groups and low quality (LQ) solution groups. The HQ group engaged in debates concerning the problems, while the LQ group focused more on the instructions provided on the worksheet. Peng, Wu, & Hu (2020) conducted an Epistemic Network Analysis (ENA) on the discussion process of pre-service STEM teachers regarding how to design STEM lesson plans in an online context. Their analysis indicated that, despite tutor providing assistance to the low-performing group, the objectives within the lesson plans designed by this group were inconsistent. On the other hand, the high-performing group systematically designed their lesson plans based on the objectives, following a step-by-step approach to intricately incorporate the content relevant to what students were expected to learn. These examples have showcased how this innovative analysis method offers a distinct perspective on the data, surpassing the traditional approach by providing valuable insights.

In most studies, information is collected through methods such as audio recordings, video recordings, or recording interactions via online learning platforms (Elmoazen et al., 2022), followed by additional coding. However, in our research, we employ a classroom observation protocol to collect data in a physical environment. We observe the teaching and learning processes in STEM-PBL courses over the span of six weeks, with each week

focusing on a different theme. Using ENA, we seek to comprehend the overall behaviors of teachers and students across this series of sessions, as well as the variations in behaviors between teachers and students from week to week. Therefore, the following research questions are posed:

RQ1: What is the distribution of individual behavior code frequencies for teachers and students in each week?

RQ2: What does ENA portray in the overall contextual relationship between teacher and student behaviors throughout the six-week course?

RQ3: Does the ENA method identify subtle differences in teacher and student behaviors across the six-week course?

## 3. Methods

## 3.1 Study Context and the Participants

This is a case study focusing on a school in southern Taiwan that adopts a STEM-PBL curriculum emphasizing on Sustainable Development Goals (SDGs) related challenges. Because the hot and humid summers and mosquito issues in the region, the project-based curriculum aims to equip 8th-grade students with relevant knowledge to develop a mosquito prevention product. According to the course plan, the development of the anti-mosquito sprayer was carried out step by step. The observation period lasted for six weeks, with each class lasting 45 minutes. There was a total of 30 students in the class, divided into groups of 6 students each.

## 3.2 Data Collection

## 3.2.1 COPIE-STEM Protocol and Coding Framework

To collect behavioral data, we adopted COPIE-STEM, an observational protocol developed by Hsiao et al. (2022) which is operated on a web-based platform called the "Generalized Observation and Reflection Platform" (GORP; University of California-Davis, n.d.). The protocol was divided into two parts: student behaviors and teacher behaviors. The coding framework for the student behaviors (Table 1) is based on the ICAP theory (Chi, 2009; Chi & Wylie, 2014), while the coding framework for teacher behaviors (Table 2) classified into three categories: teacher-centered, student-centered, or transitional. The raw data, as presented in Figure 1, are collected at 1-minute intervals. A "1" indicates the occurrence of a specific coded behavior, while "0" signifies the absence of that behavior. Finally, a total of 857 student and teacher behavior data were collected in this study, which will be used for subsequent analysis.

Code	Definition	EXAMPLE
Passive	Students simply listen and read or	Simply listening and watching the
behavior (PA)	browse the content of the	teacher.
	textbook.	
Active behavior	Students do some writing on the	Students begin to think about and
(A)	learning sheet or can follow the	briefly discuss marking suitable
	instructions to complete the	locations for anti-mosquito sprayer
	teaching task.	on their learning sheets.
Active	Students participate in group or	Student A: Who is responsible for
discussion	class discussions, such as allotting	giving the presentation on stage?
(AD)	jobs and checking procedures	

Table 1. A Coding Framework for Student Engagement Behavior in the Classroom

	following the instruction of the teacher.	Student B: I can
Constructive behavior (C)	Students go online to search for and analyze information, assemble needed parts, or write programs.	Students think about how to write a program and search for information on the Internet.
Constructive discussion (CD)	While students participate or discus in group, they work and think deeply or integrate information.	Student A: Do we need to add a timed automatic spraying device to the anti-mosquito sprayer?
		Student B: "I think we can. But how often should it be set to spray?"
Interactive behavior (I)	When students participate constructively in group activities, they may discuss more complex learning materials with specific	Student's Question: "Teacher, can I use this method to write the program?"
	group members or the teacher.	Teacher's Response: "Sure! Coming up with your own ideas is impressive."
None (N)	Students doing something unrelated to the classroom	Chatting about non-course related content

Table 2. A Coding Framework for Teacher Behavior in the Classroom

Code	Definition	EXAMPLE
Teacher-center	ed	
Showing (T.S)	Teacher shows slides or products.	Teacher showing finished products or parts
Teaching (T.T)	Teacher explains concepts or principles to the class.	Explain what the five steps of design thinking are
Student-centere	ed	
Answering questions (T.A)	Teacher answers students' questions.	Student's Question: "Teacher, is this assembly correct?"
		Teacher's Response: "You've got it the wrong way, flip it over and it'll be right."
Giving feedback (T.F)	Teacher gives feedback to the class, a group or a student for their works or ideas.	The teacher provides encouragement for the project reported by the student, along with areas that can be improved.
Asking questions (T.Q)	Teacher asks students questions and expects them to respond.	Teachers ask: Where is the suitable place to put anti-mosquito sprayer in the school?
Seeing & Listening (T.SL)	Teacher observes students or listens to their discussion	Teacher observing students assembling parts
Transitional		
Instruct (T.I)	Teacher instructs to illicit certain student behaviors or offer general principles about how to process tasks.	Teacher says, "Now you have five minutes to come up with three solutions for each group.

				Student Behaviors						Teacher Behaviors							
Session	Student	observer	Time	AD	Α	PA	N	CD	С	1	T.A	T.F	T.I	T.Q	T.SL	T.S	T.T
D1	D1.1	Karina	14:28	0	0	0	1	0	0	0	0	0	1	0	0	0	0
D1	D1.1	Karina	14:29	0	0	1	0	0	0	0	0	0	0	0	0	0	1
D1	D1.1	Karina	14:30	0	0	1	0	0	0	0	0	0	0	0	0	0	1
D1	D1.1	Karina	14:31	0	0	1	0	0	0	0	0	0	0	0	0	0	1
D1	D1.1	Karina	14:32	0	0	1	0	0	0	0	0	0	0	0	0	0	1
D1	D1.1	Karina	14:33	0	0	1	0	0	0	0	0	0	0	0	0	0	1
D1	D1.1	Karina	14:34	0	0	1	0	0	0	0	0	0	0	0	0	0	1
D1	D1.1	Karina	14:35	0	0	1	0	0	0	0	0	0	0	0	0	0	1
D1	D1.1	Karina	14:36	0	0	1	0	0	0	0	0	0	0	0	0	0	1
D1	D1.1	Karina	14:37	0	0	1	0	0	0	0	0	0	0	0	0	0	1

Figure 1. Presented through the data table recorded by COPIE-STEM

## 3.2.2 Observer Training

Prior to formal observation, each observer needed to undergo a two-hour training session focused on familiarizing themselves with the coding behaviors of both teachers and students. The inter-rater reliability between observers, measured using Sokal and Michener coefficients (SMC), is 0.92 (Hsiao et al.,2022). A total of three observers were responsible for collecting the student-teacher data for this study. Each observer primarily recorded the behaviors of a target student and the teacher every minute using COPIE-STEM (see Figure 2).



Figure 2. An Example of the Classroom Observation

## 3.3 Data Analysis

This study firstly utilized a coding and counting approach to calculate the proportional distribution of student and teacher behaviors over the course of six weeks. Subsequently, for the second question, we utilized ENA web tool (http://www.epistemicnetwork.org/) to investigate the network models for the data of teachers and students throughout these six weeks of instruction. ENA uses a sliding window to segment content-related behavioral data into sections called "stanza". By calculating the co-occurrence of behavioral data, ENA constructs epistemic networks to comprehend the interactions between these behaviors. In ENA, coded elements serve as nodes within the epistemic network. The thickness of internodal links in these networks represents the frequency of co-occurrence between elements. A thicker link indicates a higher frequency of co-occurrence between elements. For our second research question, we have chosen "student" as the analytical unit. The stanza size is set at 4, and the codes encompass student behaviors and teacher behaviors. This configuration allows us to observe the epistemic network model summarizing the behavioral interaction in the course. As for the third research question, we have selected "week" and "student" as the analytical units with the other settings remaining unchanged. In this context, we focus on observing and comparing the trends and changes in the centroids of the course each week.

## 4. Results

# 4.1 RQ1: What is the distribution of individual behavior code frequencies for teachers and students in each week?

Table 3 presents the frequency distribution of student and teacher behaviors over the six weeks of the course, in which W1 represents the first week and so forth. Passive behavior (PA) constituted the highest proportion almost every week (63%, 40%, 61%, 41%, 38%, 26% for W1 to W6, respectively), except for the last week (W6) where the proportion of Constructive behavior (C) slightly surpassed that of Passive behavior (PA). Based on the results for the first three weeks, we found that most behaviors concentrated on Passive behavior (PA), Active behavior (A), and Active discussion (AD). In contrast, during the last three weeks, Constructive behavior (C) had notably increased compared to the previous weeks. Finally, throughout the six weeks, Interactive behavior (I) consistently remained the lowest proportion among all the behaviors observed across six weeks (0%, 3%, 1%, 3%, 2%, 8% for each respective week).

	١	N1	١	N2	١	N3	١	N4	١	N5	١	N6
	(F)	(%)										
PA	83	63%	59	40%	83	61%	55	41%	62	38%	37	26%
А	26	20%	24	16%	15	11%	7	5%	26	16%	9	6%
AD	7	5%	36	24%	15	11%	28	21%	15	9%	18	13%
С	1	1%	3	2%	8	6%	24	18%	37	23%	38	27%
CD	3	2%	3	2%	3	2%	9	7%	16	10%	15	11%
Ι	0	0%	4	3%	2	1%	4	3%	4	2%	12	8%
Ν	12	9%	19	13%	10	7%	8	6%	4	2%	13	9%
Total	132	100%	148	100%	136	100%	135	100%	164	100%	142	100%

 Table 3. Frequency Distribution of Student Engagement Behavior

Table 4 shows the frequency and proportion of teacher behaviors during the six weeks. In the first and third weeks, the most frequently occurring teacher behavior was Teaching (T.T), accounting for 42% and 38% respectively. The remaining weeks predominantly featured the teacher behavior of Seeing & Listening (T.SL), with proportions of 24%, 44%, 36%, and 35% across the respective weeks. Notably, the behavior of Answering questions (T.A) was nearly negligible, approaching 0% for all six weeks. As frequencies and percentages alone cannot explain complex co-occurrences (Csanadi, 2018), we subsequently employed ENA to gain a deeper understanding of the relationship between student and teacher behaviors within the six weeks of the course.

	W1		W1 W2			W3		W4		W5	W6	
	(F)	(%)	(F)	(%)	(F)	(%)	(F)	(%)	(F)	(%)	(F)	(%)
T.A	0	0%	1	0%	1	0%	0	0%	1	0%	0	0%
T.F	3	2%	54	21%	19	8%	36	15%	15	6%	46	19%
T.Q	7	4%	40	16%	21	9%	31	13%	11	4%	11	4%
T.SL	2	1%	60	24%	53	22%	102	44%	87	36%	87	35%

Table 4. Frequency Distribution of Teacher Behavior

T.I	53	29%	51	20%	50	21%	39	17%	49	20%	39	16%
T.S	39	22%	15	6%	5	2%	6	3%	11	4%	9	4%
T.T	76	42%	31	12%	90	38%	20	9%	71	29%	56	23%
Total	180	100%	252	100%	239	100%	234	100%	245	100%	248	100%

# 4.2 RQ2: What does ENA portray in the overall contextual relationship between teacher and student behaviors throughout the six-week course?

From Figure 3, the epistemic network graph of all the student and teacher behaviors during the six weeks is presented. Singular Value Decomposition (SVD) was utilized to project nodes onto a space with two-orthogonal dimensions (SVD1 and SVD2) by maximizing variance within the data. The positions of nodes in the SVD1 x SVD2 space explain the relationship of the behaviors of interest and the meaing of the dimensions or quadrants (Shaffer et al., 2016).

The first dimension (SVD1) of the ENA space explained 30.8% of the variance within the ENA space, while the second dimension (SVD2) explained 21.6% of the variance. In the first quadrant, there were Constructive behavior (C), Constructive discussion (CD), and Interactive behavior (I), alongside the teacher behavior Seeing & Listening (T.SL). The second quadrant encompassed teacher behaviors Answering questions (T.A), Teaching (T.T), and Instruct (T.I). The third quadrant featured Active behavior (A), Passive behavior (PA), None (N), as well as two teacher behaviors, Showing (T.S) and Asking questions (T.Q). Lastly, the fourth quadrant contained Active discussion (AD) and Giving feedback (T.F).

Overall, the left half of the space predominantly featured teacher-centered behaviors (e.g., T.T and T.S) and passive engagement behaviors (e.g., PA), whereas the right half showcased more active engagement behaviors (e.g., AD, CD, C, etc.) and student-centered teacher behaviors (e.g., T.SL and T.F). We also found that higher level engagement behaviors concentrated on the upper part the space, while lower level engagement behaviors concentrated on the lower part of the space. Continuing, by observing the node sizes, we identified that the more frequently occurring behaviors were T.T, T.I, T.SL, and PA. The connection coefficients between these four nodes were as follows: PA-T.T (0.27), PA-T.I (0.24), PA-T.SL (0.24), T.I-T.SL (0.22), T.I-T.T (0.21), and T.SL-T.T (0.21). Lastly, it's noticeable from Figure 3 that T.A had minimal associations with other behaviors.



Figure 3. Epistemic Network for the Six-Week Course

# 4.3 RQ3: Does the ENA method identify subtle differences in teacher and student behaviors across the six-week course?

In order to further understand the differences between each week's courses, we utilized ENA to display where the centroids of each week are distributed within the overall epistemic network (Figure 4). Each week was represented using a different color with an arrow indicating the temporal development. Notably, there was a gradual shift from the lower left to the upper right.





Figure 5 presents the network graph for each week's course. It's noticeable that the courses in W1, W2, and W3 were more associated with passive and active student behaviors, while W4, W5, and W6 exhibited an increased proportion of constructive and interactive behaviors. The strength of connections between nodes evolved from the initial prominent PA-T.T link to the subsequent more pronounced links of various behaviors and T.SL, partialy confirming the developmental trace that ended at the first quadrant (see Figure 4).



Figure 5. Epistemic Networks of the Each Week Course

#### 5. Discussion & Conclusion

Our study investigates the relationship between student engagement and teacher instructional approaches in a six-week STEM-PBL course. We used coding and counting techniques to analyze student and teacher behaviors, and we applied ENA for in-depth class analysis. Initially, we observed that Passive behaviors(PA) are prevalent in the course, indicating students' primarily passive engagement. Additionally, we identified strong associations between specific behaviors via ENA analysis. For instance, when teachers delivered content and demonstrated materials, students tended to passively listen, reflecting traditional teacher-centered classroom dynamics (Serin, 2018).

Using ENA to track weekly student and teacher behaviors, our study identified a shift from the lower left to the upper right on the graphs. A closer look at individual network graphs revealed an initial focus on student passive engagement behavior and teachercentered instruction, gradually evolving into increased constructive engagement and student-centered approaches. This indicates a growing diversity in classroom dynamics over time, with teachers adjusting their instructional approaches, and students moving to different levels of cognitive engagement accordingly. These findings are consistent with our expectations and parallel the results observed by Akiha et al. (2018) in their study of secondary school STEM courses. The reduction in teacher-centered practices like lecturing and increased emphasis on active learning practices such as teacher guidance and student group collaboration align with their findings as well.

In summary, regardless of whether the instruction is teacher-centered or studentcentered, passive student engagement behavior is surprisingly prevalent. In order to diminish passive student engagement, teachers can find a balance between teachercentered and student-centered instruction while incorporating thoughtfully crafted activities. Implementing impactful instructional techniques, such as actively questioning students about high-level thinking, can result in improved student learning outcomes (Hsiao & Chang, 2023). Furthermore, the insights gained from ENA analysis results can assist educators in evaluating whether their courses are proceeding as planned or expected, thereby enhancing student learning outcomes.

This study also has some limitations. We collected data using the COPIE-STEM classroom observation protocol, which records data every minute, possibly missing some details. Additionally, our research covered only the first six weeks of the STEM-PBL course, which may not apply universally. Future studies should observe more extended programs for a broader perspective. Furthermore, while ENA helped explore teacher-student behavior relationships, we haven't fully explained why certain behaviors correlate strongly. Future investigations may require video analysis or course plan examination to clarify these associations. Overall, this study employed ENA to analyze STEM-PBL course dynamics but offers room for refinement and deeper exploration due to the mentioned limitations.

## 6. References

- Akiha, K., Brigham, E., Couch, B. A., Lewin, J., Stains, M., Stetzer, M. R., ... & Smith, M. K. (2018, January). What types of instructional shifts do students experience? Investigating active learning in science, technology, engineering, and math classes across key transition points from middle school to the university level. In Frontiers in Education (Vol. 2, p. 68). Frontiers Media SA. <u>https://doi.org/10.3389/feduc.2017.00068</u>
- Chi M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73–105. <u>https://doi.org/10.1111/j.1756-8765.2008.01005.x</u>

- Chi, M. T., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., ... & Yaghmourian, D. L. (2018). Translating the ICAP theory of cognitive engagement into practice. *Cognitive science*, 42(6), 1777-1832. <u>https://doi.org/10.1111/cogs.12626</u>
- Chi M. T., Wylie R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. https://doi.org/10.1080/00461520.2014.965823
- Csanadi, A., Eagan, B., Kollar, I., Shaffer, D. W., & Fischer, F. (2018). When coding-and-counting is not enough: Using epistemic network analysis (ENA) to analyze verbal data in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, *13*, 419-438. https://doi.org/10.1007/s11412-018-9292-z
- Elmoazen, R., Saqr, M., Tedre, M., & Hirsto, L. (2022). A systematic literature review of empirical research on epistemic network analysis in education. *IEEE Access*, 10, 17330– 17348. https://doi.org/10.1109/ACCESS.2022.3149812
- El Sayary, A. M. A., Forawi, S. A., & Mansour, N. (2015). STEM education and problem-based learning. *The Routledge international handbook of research on teaching thinking*, 357-369.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, 74(1), 59-109. https://doi.org/10.3102/00346543074001059
- Hsiao, J. C., & Chang, J. S. (2023). Enhancing EFL reading and writing through AI-powered tools: design, implementation, and evaluation of an online course. *Interactive Learning Environments*. <u>https://doi.org/10.1080/10494820.2023.2207187</u>
- Hsiao, J. C., Chen, S. K., Chen, W., & Lin, S. S. J. (2021). Developing a plugged-in class observation protocol in high-school blended STEM classes: Student engagement, teacher behaviors and student-teacher interaction patterns. *Computers & Education, 178*(C), 104403. https://doi.org/10.1016/j.compedu.2021.104403
- Nachtigall, V., Sung, H. (2019). Students' Collaboration Patterns in a Productive Failure Setting: An Epistemic Network Analysis of Contrasting Cases. In: Eagan, B., Misfeldt, M., Siebert-Evenstone, A. (eds) Advances in Quantitative Ethnography. ICQE 2019. Communications in Computer and Information Science, vol 1112. Springer, Cham. <a href="https://doi.org/10.1007/978-3-030-33232-7">https://doi.org/10.1007/978-3-030-33232-7</a> 14
- Peng, X., Wu, B., & Hu, Y. (2020, July). Modeling STEM learning design competence through discourse analysis. In 2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT) (pp. 180-184). IEEE.
- Rojas-Drummond, S., Torreblanca, O., Pedraza, H., Vélez, M., & Guzmán, K. (2013). 'Dialogic scaffolding': Enhancing learning and understanding in collaborative contexts. *Learning, Culture and Social Interaction*, 2(1), 11-21. <u>https://doi.org/10.1016/j.lcsi.2012.12.003</u>
- Saraç, H. (2018). The Effect of Science, Technology, Engineering and Mathematics-STEM Educational Practices on Students' Learning Outcomes: A Meta-Analysis Study. *Turkish Online Journal of Educational Technology-TOJET*, *17*(2), 125-142.
- Serin, H. (2018). A comparison of teacher-centered and student-centered approaches in educational settings. *International Journal of Social Sciences & Educational Studies*, *5*(1), 164-167. https://doi.org/10.23918/ijsses.v5i1p164
- Shaffer, D. W. (2017). Quantitative ethnography. Madison, WI: Cathcart.
- Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A Tutorial on Epistemic Network Analysis: Analyzing the Structure of Connections in Cognitive, Social, and Interaction Data. *Journal of Learning Analytics*, 3(3), 9-45. <u>https://doi.org/10.18608/jla.2016.33.3</u>
- Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., Frank, K., Rupp, A. A., & Mislevy, R. (2009). Epistemic network analysis: A Prototype for 21st-century assessment of learning. *International Journal of Learning and Media*, 1(2), 33–53. <u>https://doi.org/10.1162/ijlm.2009.0013</u>
- Smith M. K., Jones F. H., Gilbert S. L., Wieman C. E. (2013). The Classroom Observation Protocol for Undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices. CBE—Life Sciences Education, 12(4), 618–627. <u>https://doi.org/10.1187/cbe.13-08-0154</u>
- Swiecki, Z., Ruis, A. R., Farrell, C., & Shaffer, D. W. (2020). Assessing individual contributions to collaborative problem solving: A network analysis approach. *Computers in Human Behavior, 104,* 105876. <u>https://doi.org/10.1016/j.chb.2019.01.009</u>

# The Effect of Genres and Reading Orders on Interest, Reading Comprehension, and Process: Evidence from Eye Movement of Multiple-text Reading

Ching-Yu LUAN<sup>a\*</sup>, Zheng-Hong GUAN<sup>b</sup> & Sunny S. J. LIN<sup>c\*</sup> <sup>abc</sup>National Yang Ming Chiao Tung University, Taiwan \*sunnylin.nctu@gmail.com

Abstract: With the rise of online learning, multiple-text reading has become a prevailing trend and involves complex cognitive processes. Previous research often employed controversial socio-scientific issues in the form of expository texts to examine reading processes and comprehension. However, the story is relatively accessible, which can enhance learners' interest in reading. Moreover, the reading order could affect how learners integrate multiple texts. Therefore, in this study, we employed evetracking technology to investigate how different genres and reading orders of conflicting texts influence readers' change of interest, learning performance, and eye movement patterns. The study included 21 participants who were divided into two groups: expository text first and story first. Each participant was provided with two stories and two expository texts to read. Their interests were recorded before and after reading, alongside their eye movements were recorded during the reading. After completing the reading, participants were required to write an essay and take a multiple-choice test to assess their comprehension. The findings revealed that, regardless of reading order, participants' interest heightened after reading. however, the expository-text-first group had higher writing scores than the story-first group. As for the eye movement measures, this study showed that although the first-pass reading time was not affected by reading order and genre, the rereading time was. Specifically, the expository-textfirst group spent more rereading time on the expository text compared to the story-first group, indicating that these readers devote more rereading time to the expository text, focusing on taking its basic concepts and theories, resulting in greater integrated comprehension. Therefore, they exhibit superior performance in deep-level reading comprehension. This study demonstrates that reading order influences readers' reading comprehension, providing useful insights for future scientific topic education and discussions.

**Keywords:** Scientific controversial multiple texts, genre, reading order, eye movement, reading comprehension performance

## 1. Introduction

With the onset of COVID, persons must quarantine themselves at home. It caused athome online learning to become more commonplace. (Fewella, 2023). The process of consuming online information frequently entails navigating through interconnected multiple-text resources (Liu et al., 2008). Consequently, engaging with multiple texts has become a crucial aspect of contemporary life. In this circumstance, the impacts of online learning are mostly influenced by students' self-directed learning (SDL). Students often face difficulties in overcoming challenges and persevering during the learning process (Zhu, 2021). Teachers' direction and instruction are important in students' learning.

Reading multiple texts occurs in a dynamic interplay between several top-down and bottom-up processes (Britt & Rouet, 2012). This intricate process involves the iterative construction of propositions based on textual information, the assessment of reliability and

relevance among the texts, and the subsequent integration of acquired information with preexisting knowledge and personal experiences. This integration culminates in the creation of a coherent mental representation, facilitating a comprehensive understanding of the text (Kintsch, 1988; Richter, 2011).

Prior research has often employed controversial socio-scientific issues written in an expository style as reading materials (Bråten et al., 2018; Stadtler et al., 2020; Yen & Wu, 2017). Recently, the topics of gene editing, nuclear power generation, and ecological conservation have gained significant attention. The texts on this issue often are conflicting, which means that supporting and opposing perspectives exist on this issue. After reading these texts, learners usually need to evaluate the two-side positions and decide to stand for one of the positions. This behavior made readers overwhelmed because readers not only might lack interest in reading this issue (List & Alexander, 2017), but they also might confront difficulty in understanding the expository texts (Clinton et al., 2020). Chambliss & Calfee (1998) pointed out that expository texts are often used to express opinions, state facts, or describe and explain various phenomena, which lead readers to generate more inferences and can be more challenging to read than stories (Clinton et al., 2020). Compared to expository texts, stories that emphasize characters, storylines, and storytelling can help students understand the text's overall meaning. In sum, it is possible to use stories and expository texts simultaneously to evoke readers' interest to read. Given this situation, teachers should carefully consider each student's unique capabilities, reading skills, and prior knowledge while planning the material design and setting the reading order.

Similar research conducted by McCrudden et al. (2022) showed that if the reading order was a principal text first, followed by presenting with pertinent exemplar-based texts, the reading performance could be better than the pertinent exemplar-based texts first. In other words, the reading order would affect how learners process the texts. However, the reading process between exemplar-based text and principal text remains unclear. Moreover, the study did not observe a significant interest difference between groups reading the exemplar-based text first or principal text first. Yet, promoting learners' interest is important in reading multiple texts, which could lead readers to apply more reading strategies to reach deep-level comprehension (List & Alexander, 2017) and need more research.

Previous studies used the think-aloud method to collect reading process data, which might interfere with deep learning results (Cerdán & Vidal-Abarca, 2008). Eye-tracking technology is one of the non-intrusive methodologies and can provide rich and moment-to-moment cognitive processes (Rayner, 1998). Therefore, more and more multiple-text reading research adopt eye-tracking technology to collect reading process data (Salmerón et al., 2018; Stadtler et al., 2020). Given this, the purpose of this study is to use eye-tracking technology to record readers' reading processes when engaging with various genres and reading orders of conflicting socio-scientific multiple-texts. Additionally, we also explore the impact of reading orders on the reading process, comprehension, and changes in interest before and after reading. The research findings will be used as a reference for teacher instruction and material design. Three research questions of this study were as the following:

RQ1: Do different reading orders affect readers' change of interest?

RQ2: Is there a difference in learning performance between different reading orders? RQ3: How do readers' eye movement patterns differ with various reading orders?

## 2. Method

#### 2.1 Participants

21 (16 females and 5 males) postgraduates and undergraduates in northern Taiwan participated in this study. Their ages were between 18 to 30 years old (M = 24, SD = 2.55). They were randomly assigned into two groups: the "expository-text-first group" (N = 11) and the "story-first group" (N = 10). The expository-text-first group involved reading the expository text first and then reading the story, while the story-first group was the reverse.

## 2.2 Materials

#### 2.2.1 Reading materials

All participants needed to read four texts designed by the researchers. The topic was Gene-Edited Babies in Biology. The reading materials were categorized into two types of texts: "expository text" and "story." Two arguments were generated for each type. One was on the opposing side, and the other was on the supporting side. The length of the articles ranged from 351 to 353 words. The presentation order was manipulated into two versions: one with the expository text preceding the story and the other with the story preceding the expository text. This was done using counterbalancing. To maintain the experiment's simplicity, the reading sequence of the participants will be controlled, and they will not be allowed to backtrack during their reading (McCrudden et al., 2022).

#### 2.2.2 Apparatus

An EyeLink 1000 desktop remote eye-tracker system (SR Research Ltd., Canada) with a sampling rate of 1000 Hz and an accuracy of 0.5 degrees recorded participants' eye movement. Two monitors were used. One was for monitoring the eye movement information, and the other was for displaying the stimulus. The participant's head position was fixed using a chinrest placed 60 cm away from the screen. The eye movement measures used in this study were first-pass reading time and rereading time. The first-passing reading time indicated the cumulative time that readers spend on their initial fixation on areas of interest within the text, reflecting the early processing of words and the construction of meaning, such as the extraction of word meanings, decoding, and syntax analysis. The rereading time indicated the cumulative time that readers spent on their subsequent fixations on areas of interest within the text, reflecting the processing difficulties encountered after the first reading or the process of integrating the article with previous information or relevant knowledge (Hyönä et al., 2002; Zhang et al., 2019).

## 2.2.3 Topic interest measure

Topic interest measure was a self-report scale that was developed by researchers to measure readers' interest in genome editing. The scale was given to participants to fill out before and after reading. This scale included six items with a 6-Likert scale (1 = very not agree, 6 = very agree). The coefficient alpha reliabilities for this scale before and after reading were 0.79 and 0.84.

## 2.2.4 Prior knowledge

The prior knowledge test is multiple choice and includes six items selected from the entrance exam to test participants' prior knowledge. These six items were related to the topic of genome editing. Each item scored 10 points. The total score was 60 points.

## 2.2.5 Reading comprehension performance

Reading comprehension performance included two tests: multiple-choice and writing essay tasks. Multiple-choice was designed to evaluate students' memory of the reading material by directly extracting information from the text. On the other hand, writing essays assessed students' ability to integrate information from multiple texts. The participants received the following instruction: "After reading the texts, synthesize and compare different viewpoints presented in the articles, and finally, present and support your stance." Our rubric to score students' written responses was divided into four dimensions: two-sided reasoning, elaboration, integration, and evidence use. Description and examples of the dimensions are detailed in Table 1. We also calculated the scores of all the dimensions as total scores (Lee & List, 2021). Two raters coded all student responses. The inter-rater reliability for each was

higher than 0.8. For the disagreement, a discussion was conducted until two raters reached a consensus.

Dimensions	Description	Example
Two-sides reasoning	Whether readers respond to opposing views, present instances, and make refutations.	"If everyone may select their desired genes, it appears that biological variety may be lost"
Elaboration	Whether readers articulate their own perspectives.	" <u>I think</u> gene editing has a lot of potential in terms of treating human disorders. However, there are still a lot of technological and moral issues to think about"
Integration	Whether readers integrate information across different articles.	"People who favor gene editing believe that it has the potential to change a child's DNA so they can fend off the effects of AIDS. <u>However</u> , it is also evident that unrestricted gene editing may result in problems that cannot be fixed"
Evidence use	Whether readers cite content from the articles.	" <u>As mentioned in the last article</u> , even if one matches their children's genes to exhibit desirable physical attributes and intelligence,"

Table 1. Description and examples of the rubric dimensions

## 2.3 Procedure

This study employed eye-tracking technology to gain insights into the reading processes of the subjects and consisted of three main parts. Firstly, participants were told the experiment procedure and completed the topic interest measure. During the second part, a nine-point calibration was conducted before participants started reading the texts. While reading, participants determined the pace themselves; however, revisiting previous pages was not permitted. After the experiment, participants underwent the writing essay task and the multiple-choice test, followed by an interview to ascertain whether the narrative articles assisted them in constructing a comprehensive understanding of the text during the reading process.

## 2.4 Statistical Method

As for the RQ1, we used mixed-design ANOVA to examine the change of interest (preand post-test) between two groups (Expository-Story and Story-Expository). As for the RQ2, we ran the linear regression to examine whether the differences in the writing essay task (four dimensions and total scores) existed between the two groups after including the number of words as a control variable. We also used the independent t-test to compare the score of the multiple-choice test between the two groups. As for the RQ3, we analyzed the eye-movement measures by using linear mixed model with the reading order and genres as fixed effects while the participants and the sentence as random effects. The eye-movement measures (first-pass reading time and rereading time) were dependent variables. The sentence length was also included in the models as a control variable.

## 3. Result

## 3.1 Pre-test and post-test interest performance of each group

To understand the impact of different reading orders on readers' interest (RQ1). Mixeddesign ANOVA was conducted with pre-post as within variable and reading order as the between variable. As shown in Table 2, readers' interest increased after reading the texts in both groups ( $F_{(1,19)} = 14.75$ , p = .001). Nevertheless, the main effect of reading order and the interaction effect did not reach a significant level (reading order:  $F_{(1,19)} = 0.01$ , p = .92; interaction effect:  $F_{(1,19)} = 0.03$ , p = .85). These findings suggest that readers in both groups exhibited heightened interest after reading the texts, but there was no difference in interest between groups.

•		0 1	
Pre-test	Post-test	Prior	Multiple-choice
interest	interest	knowledge	test
M (SD)	M (SD)	M (SD)	M(SD)
18.45	22.09	41.82	48.18
(4.01)	(6.06)	(13.28)	(9.82)
18.80	22.10	48.00	46.00
(4.16)	(3.21)	(11.35)	(12.65)
	Pre-test interest <i>M (SD)</i> 18.45 (4.01) 18.80 (4.16)	Pre-test         Post-test           interest         interest           M (SD)         M (SD)           18.45         22.09           (4.01)         (6.06)           18.80         22.10           (4.16)         (3.21)	Pre-test         Post-test         Prior           interest         interest         knowledge           M (SD)         M (SD)         M (SD)           18.45         22.09         41.82           (4.01)         (6.06)         (13.28)           18.80         22.10         48.00           (4.16)         (3.21)         (11.35)

Table 2. Results of descriptive statistics for pre-test and post-test interest performance, prior knowledge test and reading comprehension performance of each group

## 3.2 Reading comprehension performance of each group

To comprehend the differences in learning performance between readers who read different text orders (RQ2), linear regression was used for the writing essay tasks with the reading order as the independent variable, and the word count as the control variable. Besides, the t-test was used to compare the multiple-choice score between the two groups. Before the main analysis, the prior knowledge between the two groups was examined. The result of descriptive statistics has shown in Table 2. The result of the t-test showed that there was no significant difference in the multiple-choice scores between the two groups (t(19) = -1.14, p = .26). Therefore, the following analysis didn't consider prior knowledge.

In the case of the writing essay task, a significant difference was found between the two groups. As shown in Table 3, the results of regression demonstrated that students with longer writing lengths achieved higher scores (b = 0.03, SE = 0.01, t = 4.44, p < .001). After controlling for word count, the group that read the expository text first outperformed the group that read the story first (b = 2.46, SE = 0.75, t = 3.30, p = .003).

Furthermore, when comparing the performance of different groups across various dimensions of the writing task, the results showed that the group reading the expository text first performed better in the dimension of "two-sides reasoning" and "integration" than the group reading the story first (two-sides reasoning: b = 0.91, SE = 0.40, t = 2.30, p = .03; integration: b = 0.84, SE = 0.35, t = 2.42, p = .003). However, as shown in Table 4, there were no significant differences between the two groups in the dimensions of "elaboration" and "evidence use" (elaboration: b = 0.28, SE = 0.20, t = 1.41, p = .17; evidence use: b = 0.43, SE = 0.34, t = 1.25, p = 0.23).

Table 3. Results of regression analyse	s for variables	s predicting perfo	rmance on total score,
two-sides reasoning, and integration			

Predictor	Tot	tal sco	res	Two-sides reasoning			lı	Integration			
-	В	SE B	β	 В		SE B	β	В	SE B	β	
Number of words	0.03	0.01	0.69***	0.01		0.00	0.55**	0.01	0.00	0.58**	
Reading order	2.46	0.75	0.51**	0.91		0.40	0.43*	0.84	0.35	0.44*	
*n < 05 **n < 01	***n <	001									

187

Predictor	I	Elaboratio	n		Evidence use				
	В	SE B	β	В	SE B	β			
Number of words	0.00	0.00	0.33	0.01	0.00	0.47*			
Reading order	0.28	0.20	0.31	0.43	0.34	0.26			
*p < .05, **p < .01, **	*p < .001								

Table 4. Results of regression analyses for variables predicting performance on the elaboration and evidence use

## 3.3 Analysis of eye movement data of each group

To comprehend the process of eye-movement patterns between different reading orders (RQ3). The linear mixed model was used in this study. The fixed effects included genre with two levels (story and expository text) and reading order with two levels (story first and expository text first). Besides, the sentence length was entered into models as a control variable. The mean and standard deviation are presented in Table 5, and the statistical results are summarized in Table 6.

Table 5. Mean and standard deviation for eye-movement measure

	Expository f	text first	Story fi	rst	
	Expository text	Story	Expository text	Story	
	M (SD)	M (SD)	M (SD)	M (SD)	
First pass reading time	777.73	703.49	621.08	588.62	
First-pass reading time	(754.22)	(654.26)	(675.19)	(603.09)	
Dereading time	595.80	499.69	326.62	414.78	
Rereating time	(908.57)	(744.71)	(616.83)	(622.20)	

For the first-pass reading time, the sentence length was significant (b = 47.49, SE = 3.73, t = 12.74, p < .001). However, the main effects and interaction effect were not significant (Reading order: b = -135.76, SE = 120.48, t = -1.13, p = .27; Genre: b = -64.11, SE = 38.44, t = -1.67, p = .10; Reading order × Genre: b = 41.79, SE = 46.77, t = 0.89, p = .37). These results indicated that readers spent the equal time constructing the semantic process (such as extracting word meanings, decoding, and syntactic analysis) regardless of genres or reading order they encountered.

For the rereading time, the sentence length was significant (b = 38.13, SE = 3.40, t = 11.24, p < .001). Besides, although the two main effects were not significant (Reading order: b = -177.04, SE = 121.04, t = -1.46, p = 0.16; Genre: b = -12.62, SE = 35.00, t = -0.36, p = 0.72), the interaction effect was significant (b = 184.24, SE = 55.30, t = 3.33, p < .001), as shown in Figure 1.

Table 6. Results of linear mixed model for first-pass reading time and rereading time

	First-pass reading time				Rereading time			
	b	SE	t	р	b	SE	t	р
Sentence length	47.49	3.73	12.74	< .001	38.13	3.40	11.24	< .001
Reading order	-135.76	120.48	-1.13	.27	-177.04	121.04	-1.46	.16
Genre	-64.11	38.44	-1.67	.10	-12.62	35.00	-0.36	.72
Reading Order × Genre	41.79	46.77	0.89	.37	184.24	55.30	3.33	< .001


Figure 1. The interaction effect of reading order and genre on rereading time.

Table 7 shows the result of the simple main effect of rereading time. When reading the expository text, the group reading the expository text first had significantly longer rereading time on the expository test than the group reading story first (b = 269.2, SE = 124, t = 2.17, p = .04); while this effect was not found when reading the story (b = 84.9, SE = 124, t = 0.68, p = .50). Additionally, the group reading the expository text first significantly spent more rereading time on the expository text than on the story (b = 104.7, SE = 43.8, t = 2.39, p = .01). However, the group reading story first slightly spent less rereading time on expository texts compared to story which only reach marginal significant (b = -79.5, SE = 45.4, t = -1.75, p = .08).

<b>Fixed</b> lovel	Compare	Rereading time			
Fixed level	Compare	b	SE	t	р
Expository text	Expository text first - Story first	269.2	124	2.17	.04
Story	Expository text first - Story first	84.9	124	0.68	.50
Expository text first	Expository text - Story	104.7	43.8	2.39	.01
Story first	Expository text - Story	-79.5	45.4	-1.75	.08

Table 7. The simple main effect of the interaction effect on rereading time

### 4. Discussion and Conclusion

The purpose of this study was to investigate the effect of reading order (story first or expository text first) on multiple-text reading comprehension and the process reflected by eye movement measures. Besides, we also examine the effect of reading order on the change of interest. We proposed the following research findings. First, regardless of the reading order, readers' interest increased after reading the texts compared to before reading the texts. This result is inspiring because after reading such a conflicting topic, the readers promote their interest in this topic instead of decreasing it, which implies that readers might change their original attitude after reading the texts and search for more information to read beyond experimental texts. As a result, when creating materials, teachers can take into account the different abilities and prior knowledge of their students. This may increase students' interest in reading and promote their self-directed reading.

Secondly, the comprehension performance of multiple-choice was not different between reading the story first and reading the expository text first. However, as for the writing essay task, the group reading the expository text first performed better than the group reading story first. In this study, the multiple-choice test was regarded as the surface-level representation, whereas the writing essay task represented deeper-level understanding (Chen et al., 2014). In other words, readers in this study had no difficulty in understanding and remembering the basic information the texts provided. However, their deep-level representation was influenced by the reading orders, especially for the "two-sides reasoning" and "integration." These results can be explained by the eye-movement data.

Specifically, the two groups with different reading orders had no significant difference in first-pass reading time. These results showed that readers spend equal time establishing semantic understanding regardless of the reading order, which seems consistent with the impact of the multiple-choice test because readers could remember as much as possible and were not affected by the reading orders. As for the rereading time, readers in the expository first group spent more time on the expository to integrate information than readers in the story first group. This behavior entailed readers to understand the expository text better; therefore, they could make more integrate with the following stories and provide more arguments to justify their positions, which led to better two-sides reasoning scores. However, readers in the reading story first group did not demonstrate this cognitive process which implied that after reading the story, they could not integrate the following expository text better and spent almost equal time on these two genres. This result was consistent with the previous study conducted by McCrudden et al. (2022), who also found that the principal text first promoted readers to have more cross-text integration.

In conclusion, our study yielded several significant insights. As for the interest, we found that readers' interest increased after reading multiple texts, which led the educators to have confidence in using socio-scientific issues as materials to teach judging the information and writing a convincing argumentation. As for the reading process and comprehension, readers might use diverse reading strategies in different reading orders. Although readers had satisfying learning outcomes in surface-level reading comprehension; however, the group that read the expository text first had better deeper-level reading comprehension involving two-side reasoning and integration. The reason for this is that readers spent more time reading the expository text in order to integrate and infer understanding, then they used the story to validate what they had learned, leading to increased deep-level comprehension.

Our study suggests two instructive implications. Firstly, educators can use a teaching method that mixes expository and story genres in computer-assisted instruction, including scientific multiple-text to increase learners' interest and knowledge of the topics. Secondly, teachers can also teach students reading strategies for scientific multiple-text, such as reading the expository part first to understand the issue and then using the story as an illustrative example to achieve better integration and comprehensive reading comprehension.

### References

- Bråten, I., Brante, E. W., & Strømsø, H. I. (2018). What really matters: The role of behavioural engagement in multiple document literacy tasks. *Journal of Research in Reading*, *41*(4), 680-699. <u>https://doi.org/10.1111/1467-9817.12247</u>
- Britt, M. A., & Rouet, J.-F. (2012). Learning with Multiple Documents: Component Skills and Their Acquisition. In J. K. M. Lawson (Ed.), *Enhancing the Quality of Learning: Dispositions, Instruction, and Learning Processes* (pp. 276–314). Cambridge: Cambridge University Press. https://doi.org/doi:10.1017/cbo9781139048224.017
- Cerdán, R., & Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology*, *100*(1), 209-222. https://doi.org/10.1037/0022-0663.100.1.209
- Chen, G., Cheng, W., Chang, T.-W., Zheng, X., & Huang, R. (2014). A comparison of reading comprehension across paper, computer screens, and tablets: Does tablet familiarity matter? *Journal of Computers in Education*, *1*(2), 213-225.

- Clinton, V., Taylor, T., Bajpayee, S., Davison, M. L., Carlson, S. E., & Seipel, B. (2020). Inferential comprehension differences between narrative and expository texts: a systematic review and meta-analysis. *Reading and Writing*, *33*(9), 2223-2248. <u>https://doi.org/10.1007/s11145-020-10044-2</u>
- Fewella, L. N. (2023). Impact of COVID-19 on distance learning practical design courses. International Journal of Technology and Design Education. https://doi.org/10.1007/s10798-023-09806-0
- Hyönä, J., Lorch, R. F., Jr., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology*, *94*(1), 44-55. https://doi.org/10.1037/0022-0663.94.1.44
- Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. New York, NY: Cambridge University Press.
- Lee, H. Y., & List, A. (2021). Examining students' self-efficacy and perceptions of task difficulty in learning from multiple texts. *Learning and Individual Differences*, 90. https://doi.org/10.1016/j.lindif.2021.102052
- List, A., & Alexander, P. A. (2017). Cognitive Affective Engagement Model of Multiple Source Use. *Educational Psychologist*, *52*(3), 182-199. https://doi.org/10.1080/00461520.2017.1329014
- Liu, O. L., Lee, H.-S., Hofstetter, C., & Linn, M. C. (2008). Assessing knowledge integration in science: Construct, measures, and evidence. *Educational Assessment, 13*(1), 33– 55. <u>https://doi.org/10.1080/10627190801968224</u>
- McCrudden, M. T., Kulikowich, J. M., Lyu, B., & Huynh, L. (2022). Promoting integration and learning from multiple complementary texts. *Journal of Educational Psychology*, *114*(8), 1832-1843.
- Rayner, K. (1998). Eye movements in reading and Information processing 20 years of research. *Psychological Bulletin*, *124*(3), 372-422. <u>https://doi.org/10.1037/0033-2909.124.3.372</u>
- Richter, T. (2011). "Cognitive flexibility and epistemic validation in learning from multiple texts," in *Links Between Beliefs and Cognitive Flexibility*, eds J. Elen, E. Stahl, R. Bromme, and G. Clarebout (NewYork, NY: Springer), 125–140.
- Salmerón, L., Gil, L., & Bråten, I. (2018). Using eye-tracking to assess sourcing during multiple document reading: A critical analysis. *frontline Learning Research*, 105-122. https://doi.org/10.14786/flr.v6i3.368
- Stadtler, M., Scharrer, L., & Bromme, R. (2020). How Relevance Affects Understanding of Conflicts Between Multiple Documents: An Eye-Tracking Study. *Reading Research Quarterly*, 55(4), 625-641. <u>https://doi.org/10.1002/rrg.282</u>
- Yen, M.-H., & Wu, Y.-T. (2017). The role of university students' informal reasoning ability and disposition in their engagement and outcomes of online reading regarding a controversial issue: An eye tracking study. *Computers in Human Behavior*, 75, 14-24.
- Zhang, D., Hyönä, J., Cui, L., Zhu, Z., & Li, S. (2019). Effects of task instructions and topic signaling on text processing among adult readers with different reading styles: An eye-tracking study. *Learning and Instruction*, 64. https://doi.org/10.1016/j.learninstruc.2019.101246
- Zhu, M. (2021). Enhancing MOOC learners' skills for self-directed learning. *Distance Education*, 42(3), 441-460. <u>https://doi.org/10.1080/01587919.2021.1956302</u>

# Develop and Analysis of Educational Board Game <The Golden Silk Road> on Cultural Cognition

### Hsin-Ju Wang, Shaun-Wen Chen\*, Ju-Ling Shih

Graduate Institute of Network Learning Technology, National Central University, Taiwan \*seanchen54017@gmail.com

**Abstract:** This study investigated the impact of digital game-based learning on cultural cognition among upper-grade primary school students by developing a tabletop game combining augmented reality (AR) based on the theme of the Silk Road. The game aimed to enhance learning outcomes through multiple media and interdisciplinary integration, promoting multiculturalism, aesthetics education, and technological information. AR technology is used as an assisted learning tool to design game activities that enabled learners to role-play, trade, construct, and experience historical events along the Silk Road, reducing learning pressure while learning in a fun and interactive way. Results showed that <The Golden Silk Road> game activity significantly improved participants' understanding of Silk Road> could arouse students' interest, stimulate their internal motivation to learn and encourage active participation, with positive feedback from participants.

**Keywords:** Digital Game-Based Learning, Cultural Cognition, Augmented Reality, Tabletop Games.

### 1. Introduction

The 21st century emphasizes diverse learning and technology's role in transforming education. Philosopher Dewey promoted "learning by doing," advocating practical, hands-on problem-solving (Williams, 2017). Integrating education into physical board games helps students grow in knowledge, social skills, and emotions. Combining tabletop games with digital tech is a recent trend, providing information, social interaction, competition, and sensory stimulation, bridging gaps in traditional education. Augmented Reality (AR) breaks down barriers, offering a richer learning experience. This study explores "Cultural Cognition" changes in elementary students using the educational board game <The Golden Silk Road> and assesses satisfaction and feedback on incorporating AR into cultural aesthetics education board games.

### 2. Related work

### 2.1 DGBL tabletop game with AR integration

Digital game-based learning (DGBL) combines technology and games with educational learning to assist individuals in achieving ideal learning outcomes. It utilizes a reward system, visual imagery, auditory or interesting elements to attract students and immerse them in the learning process, which is a trend in modern education (Deubel, 2006). Tabletop games, combined with diverse learning fields, not only enhance learning outcomes but also help students to develop a spirit of self-initiation, independent thinking, and mutual growth (Dyson et al., 2016). With the increasing number of cases utilizing technology-assisted game-based learning, smart mobile devices, tablets, or multi-touch screens are commonly used as media. Compared to the traditional desktop computers, smart mobile devices are more lightweight, flexible, and cost-effective (Behnamnia et al., 2022). These new learning

methods deviate from traditional classroom teaching methods. Firstly, students are easily attracted to novel technologies, which arouses their interest. Secondly, the rendering power among peers and the classroom atmosphere can infect everyone's emotions, becoming a driving force. When students immerse themselves in games, their concentration naturally increases, which can improve their absorption of knowledge. These factors are considered in this study.

### 2.2 Silk Road Cultural Cognition

The Silk Road served as a vital transportation route connecting the East and the West, stretching from Xi'an, China to the Mediterranean coast of Europe. The diverse terrain along the route, including deserts, oases, desolate regions, and towering mountains, enriched the multiculturalism of the Silk Road. Early on, missionaries used the Silk Road to spread their faith, while in the 19th century, explorers, geographers, and archaeologists conducted a series of investigations along the Silk Road (Wood, 2002). Human commerce and business transactions were accompanied by the transmission of knowledge, ideas, religions, languages, and customs (Andrea, 2014).

Cognitive psychologist Neisser (1967) defined "cognition" as a series of processes that occur internally, including transformation, deletion, modification, addition, storage, regeneration, and application of sensory input, even in the absence of external stimuli. The result of participating in a way of life that has been passed down by a particular ethnic group or region is species-unique cultural cognition (Tomasello, 2005). Cultural cognition refers to the cognitive behaviors in two forms, external forms such as architecture, daily utensils, artistic creations, entertainment, and internal forms such as social civilization, language, norms, religion, and class, which have been formed by various ethnic groups.

The impact of the Silk Road on historical culture was profound, with trades and resources exchanges serving as mediums. This study allows players to act as the representative of a nation with the goal of promoting national prosperity and experiencing the beauty of the Silk Road.

### 3. Game design

### 3.1 Game Object

Taking into consideration the literacy and reading capabilities of upper-grade elementary students and adjusting for an appropriate level, the visual and interface design of the game <The Golden Silk Road> primarily focuses on graphics, with text serving as a supplementary element to ensure intuitive and easy differentiation. Throughout gameplay, players engage in role-playing, enabling them to witness the historical traces of cultural exchanges, fusion, and coexistence along the Silk Road. <The Golden Silk Road> features four types of game objects: (1) the Silk Road Map, (2) Civilization Map Cards, (3) Tokens, and (4) Mobile Devices.

### 3.1.1 Silk Road Map

Game map of <The Golden Silk Road> (as shown in Figure 1) is modeled after the ancient Silk Road, with Chang'an in the east end and Constantinople in the west end as the main representative cities. Three kinds of geological features are simulated: a) high mountains (brown), which will take two steps pre move; plains (green), which takes one step per move; and deserts (yellow), which takes one step per move and require a camel mount card. The map also lists the major cities (dark red) and small cities (blue) and sacred places (orange) that the Silk Road passed through. The four nations are represented by different colors and are divided into four territorial areas by colored lines: Chang'an area as red outline for Tang Dynasty, Dunhuang area as blue outline for Gupta, Samarkand area as yellow outline for Arab Empire, and Constantinople area as purple outline for Byzantine Empire.



Figure 1. Map of <The Golden Silk Road>

### 3.1.2 Civilization Map Cards

Each nation has its own civilization card (as shown in Figure 2), which describes the country's background, mission details, and abilities, divided into three stages from left to right. After completing the missions as instructed, players can upgrade to the next dynasty and gain special abilities. Missions are divided into three categories: politics, economy, and religion. The second mission of each categories involve exchanging the obtained cultural relics from mission one with the designated nation. Once all three categories are completed, players can use the special abilities listed in the third stage of the civilization card in the next round to help them become the winner of the game.



Figure 2. The Civilization Card for Arab Empire

### 3.1.3 Tokens

There are four types of tokens: a) pawns, b) mounts, c) local treasures, and d) buildings. a) Each nation has a pawn for Silk Road navigation. b) Mount cards are available in major cities, limited to one per nation. c) Each nation has unique regional products, distributed as local treasures at round's end, increasing in value with distance. d) Players can build Waystations, Fortresses, and Temples for their nation each round.

### 3.1.4 Mobile Device

<The Golden Silk Road> has incorporated augmented reality (AR) app technology to add human-computer interaction to game-based learning using the Artivive application. Artivive is

an augmented reality tool designed by a Viennese team in Austria (www.artivive.com), which can display dynamic or static information in a tablet device when scanning key surfaces, explaining which unique artifacts were produced in the local area at the moment of time and space. Players can view ancient artifacts and stories with the mobile device. The visual imagery is segmented into puzzle pieces, requiring players to adjust the tablet's angle to piece together the appearance of the artifacts, thereby increasing the fun and impact of the gameplay. When players move to a city, they can open the Artivive app and point the camera towards the city to display the corresponding information. The scanning reward (as shown in 錯誤! 找不到參照來源。) can also be found in the displayed image, with a total of 32 scanning points included in the entire map.

### 3.2 Game Flow

In <The Golden Silk Road>, four to eight players represent the Tang Dynasty, the Arab Empire, the Byzantine Empire, and the Kingdom of Gupta, guided by a game master. Sessions last 60 to 90 minutes. Players collect coins through AR-aided relic exploration and missions. They choose a nation, receiving nation-specific items. Each turn, players can navigate, trade, build, or raid. Scanning cities with tablets reveals historical events, offering rewards for comprehension, combining learning with gameplay. At round's end, the game master disperses treasure tokens and coins. Participants journey along the Silk Road map, fulfilling tasks at waypoints. The game concludes after two rounds following all players' dynasty advancements. Victory is determined by total coins.

### 4. Research Design

This study involved 50 sixth-grade students (aged 11-12) from a Taiwanese primary school, comprising 24 males and 26 females. Data collection included pre- and post-tests, questionnaires, and interviews. The pre- and post-tests consisted of 10 questions on cultural cognition. Results were analyzed quantitatively and qualitatively. Interviews were coded as country-question number-student (e.g., AR-01-S1 for an Arab Empire student answering question 1). The test assessed Silk Road cultural knowledge with 9 multiple-choice questions and 1 matching question (split into 5 sub-questions). It covered geography, history, cultural integration, and artifact identification to evaluate learning from <The Golden Silk Road> board game. Students completed the pre-test, followed by game rules explanation and gameplay. (as shown in Figure 3).



Figure 3. Game Master leads the students to play <The Golden Silk Road>.

After the conclusion of the game, students are required to complete the post-test and the questionnaire. The questionnaire examines five factors: "Game Motivation", "Visual Interface", "Learning Knowledge", "Game Mechanism Design", and "AR design". Finally, interviews were conducted to support and explained the results of the questionnaires. Research Process is as shown in Figure 4.



*Figure 4.* Research Process of <The Golden Silk Road> Game Activity

### 5. Result

This study investigates the effects of digital game-based learning on cultural cognition, and analyzes the differences in "Cultural Cognition" for elementary school students before and after using the educational board game <The Golden Silk Road>. A satisfaction questionnaire and group interviews were conducted after the tests to understand the students' opinions and suggestions on game design. The effective number of students is 50.

### 5.1 Cultural Cognition

The paired-sample t-test results of pre- and post-tests for cultural cognition were examined. Table 1 shows the overall paired t-test results of the pre-test and post-test scores of cultural cognition for 50 participants. The post-test scores were higher than the pre-test scores in average, and has a significant improvement in culture cognition. During the game process, participants would decide where buildings should be constructed for the most favorable outcome or for completing the mission, thus recognizing the geographical location relationship between different cities. The cultural integration was presented in visual images, and the use of AR technology to read artifact descriptions with artifact imaging in the game process was helpful. Participants acquired historical knowledge in the process of using AR technology and game activities, indicating their familiarity and engagement with board games. They could see artifacts and their details in the game but also experience the Silk Road's methods, and learn about the history of "Cultural Integrations". The participants stated, "Through <The Golden Silk Road>, I can learn more about the culture on the Silk Road and navigate for trading (GU-04-S3)." It proved that participants could learn knowledge through peer cooperation and competition in the game, simulate the situation of the Silk Road by playing roles, and enhance their learning effectiveness and motivation.

	Questions)					
Cultural Cognition	Ν	Means	SD	t	р	
Pre-test	50	29.08	18.121	6 256***	000	
Post-test	50	47.92	19.111	-0.330	.000	
*** 001						

Table 1 Results of Players' Cultural Cognition of Playing <The Golden Silk Road> (QN=Numbers of Questions)

\*\*\*\*p<.001

### 5.2 Overall Satisfactions

According to Table 2, the overall satisfaction of the participants from both classes towards the board game <The Golden Silk Road> was reasonably high. The participants were able to learn cultural knowledge and aesthetic literacy through the game activity. The overall satisfaction of the participants were 3.99 points. This could be due to the fact that the participants are engaged with the game mechanics (M=3.58) and the integration of AR technology (M=3.99) is novelty for six grade elementary students. They also responded positively to the game: "Need to use strategy and have discussions during the game (*TA-06-S2*)" and "Using AR technology can have more interactivity (*TA-06-S2*)". Participants believed that the game had clear mission instructions and that the interface design and configuration were beautiful and attractive, which could help enhance learning motivation and interest in participating in the game (M=4.14). They also believed that the game materials were clearly and reasonably presented and could deepen their understanding of

the Silk Road culture (M=4.11). Participants stated that they would recommend it to their friends, but they thought that the game mechanism and rules were slightly complicated to understand (M=3.58).

Aspects	Numbers of Questions	Means
Game Motivation	3	4.15
Visual Interface	3	4.14
Learning Knowledge	2	4.11
Game Mechanism Design	6	3.58
AR design	6	3.99
Overall Satisfaction	20	3.99

Table 2 Overall Results of Players' Satisfactions

### 6. Conclusion

This study developed a physical board game infused with AR technology to enhance history, geography, and humanities education. The game featured visually appealing elements, including vibrant colors, cultural themes, and regional landscapes, making learning interactive and engaging. "The Golden Silk Road" board game was used to assess the impact on elementary students' "Cultural Cognition" and their satisfaction. The results demonstrated a significant enhancement in cultural cognition, highlighting the effectiveness of well-designed educational games in stimulating learning. Both classes averaged 14 AR interactions, indicating user-friendliness and encouraging exploration of Silk Road culture. The integration of AR with game-based learning immersed students in historical and geographical knowledge. For future game-based learning, simplifying game mechanics and rules explanations for elementary students could enhance their participation and learning experience. This study showcases the potential of combining AR technology with educational games to foster immersive learning.

### Acknowledgements

This study is supported in part by the National Science and Technology Council (previously known as Ministry of Science and Technology) of Taiwan, under MOST 108-2511-H-008 -016 -MY4 and NSTC 112-2410-H-008 -010 -003.

### References

- Andrea, A. J. (2014). The Silk Road in world history: A review essay. *Asian Review of World Histories,* 2(1), 105-127. https://doi.org/10.12773/arwh.2014.2.1.105
- Behnamnia, N., Kamsin, A., Ismail, M. A. B., & Hayati, S. A. (2022). A review of using digital gamebased learning for preschoolers. *Journal of Computers in Education*, 1-34.

Deubel, P. (2006). Game On!. THE Journal, 33(6), 30.

- Dyson, S. B., Chang, Y.-L., Chen, H.-C., Hsiung, H.-Y., Tseng, C.-C., & Chang, J.-H. (2016). The effect of tabletop role-playing games on the creative potential and emotional creativity of Taiwanese college students. *Thinking Skills and Creativity, 19*, 88–96. doi:10.1016/j.tsc.2015.10.004
- Neisser, U. (1976). 1967Cognitive psychology. New York: Appleton-Century-Crofts.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and brain sciences, 28*(5), 675-691. https://doi.org/10.1017/S0140525X05000129
- Williams, M. K. (2017). John Dewey in the 21st Century. *Journal of Inquiry and Action in Education, 9* (1). Retrieved from https://digitalcommons.buffalostate.edu/jiae/vol9/iss1/7
- Wood, F. (2002). The Silk Road: two thousand years in the heart of Asia. Univ of California Press.

### Supporting Peer Help Recommendation Based on Learner-Knowledge Model

### Peixuan JIANG<sup>a\*</sup>, Kensuke TAKII<sup>a</sup>, Changhao LIANG<sup>a</sup>, Rwitajit MAJUMDAR<sup>b</sup> & Hiroaki OGATA<sup>b</sup>

<sup>a</sup>Graduate School of Informatics, Kyoto University, Japan <sup>b</sup>Academic Center for Computing and Media Studies, Kyoto University, Japan \*jiang.peixuan.67n@st.kyoto-u.ac.jp

**Abstract:** With the development of information technology tools and the Internet, computer-supported collaborative learning has become increasingly accessible and promising. Peer help is a popular practice of collaborative learning. In this paper, we propose a way to realize intelligently mediated peer help. We obtain open learning activity data from an integrated learning platform called LEAF for modeling. First, we create a network-based knowledge model. Then, we construct learner models associated with the knowledge model. Based on the knowledge and learner modeling, we propose a method to find problems of a learner based on the order of closeness centrality of knowledge nodes. Also, the system recommends potential peer helpers who can help with these problems. We present a scenario of physics learning at the high school level to explain the practical use of this method which is aimed to enhance learners' initiative during peer help.

**Keywords:** peer help, peer recommendation, learning analytics, knowledge model, learner model

### 1. Introduction

Collaborative learning is a situation in which particular forms of interaction among people are expected to occur, which would trigger learning mechanisms (Dillenbourg, 1999). Peer help is a major practice of collaborative learning. Assigning right peers is the key to the success of peer help, and there are several studies that show that teachers utilize some computer-assisted systems to make their students more satisfied (Urhahne, 2010).

This paper presents an approach utilizing open learning activity data from an integrated learning platform called LEAF (Ogata et al., 2018). This will achieve intelligently mediated peer help in a data-driven approach. A knowledge model is created by automatically extracting knowledge elements from the LEAF system and represents a knowledge network. Based on the knowledge model, a learner model is created from records of individual learning activities from the LEAF system. The learner model combined with the knowledge model can indicate the learner's knowledge proficiency at different stages. This paper uses a scenario of physics learning at the high school level as an example, illustrating the possibility of providing adequate peer help recommendation for learners. The peer help recommendation consists of recommending problems that can be helped by peers, as well as potential peer helpers with complementary knowledge. A learner can choose both a problem and a peer helper from the system.

### 2. Theoretical Background

Three cognitive dimensions were pointed out which appear to be important to the success of collaborative problem-solving and learning are learners' knowledge, information exchange, and roles (Hoppe & Ploetzner, 1999), among which knowledge is the most common consideration when assigning collaborative learning partners (Abnar et al., 2012). Knowledge

network is a tool for graphically presenting a knowledge architecture that allows for a clearer understanding and organization of complex knowledge content. This network can be used to represent knowledge units in the educational area as well, and computers hold the potential to realize this representation. The concept of semantic networks was introduced by Quillian (1968), which represents semantic information such as knowledge. Furthermore, ontology was introduced as a concept for knowledge organization that considers the intrinsic associations of knowledge more than simply describing semantic information (Arroyo et al., 2003). In application areas, Khan Academy builds knowledge networks for subjects such as math and science, which guide open and personalized online learning (Meng et al., 2018).

Learner model is a representation of a real learner, representing the characteristics and level of the learner's knowledge and skills (Chrysafiadi & Virvou, 2013). Open learner modeling allows learners to compare their own models with the class as a whole and with the other learners' individual models (Brusilovsky et al., 2016). In the field of education, using data mining and other techniques to analyze learning records to construct learner portrait models allows education systems to provide more personalized support for learners (Baker & Inventado, 2014). The LEAF system, for example, provides an evidence-based dashboard to analyze and visualize learning logs of e-book reading, quizzes, etc. to demonstrate a student's learning (Majumdar et al., 2019). Experience API or xAPI (2016) is another state-of-the-art open learning record standard. Following the idea of xAPI, this paper adopts a network structure to connect the learner to the learned knowledge.

Peer help is a major practice in collaborative learning and matching an appropriate peer helper is the key to the success of peer help. In intelligently mediated peer help, the individually assessed learner models are used to match better pairs of learners (Hoppe, 1995). Related studies have made attempts to analyze and model the knowledge that learners currently have and identify theoretical correlations between their knowledge in the form of Prologs (Hoppe, 1995) and ontology (Ounnas et al., 2009; Isotani et al., 2009). It is expected that the inner connections between knowledge are considered when meditating peer helpers since these connections provide information of complementary knowledge which can contribute to effective collaborative learning (Isotani et al., 2009). However, existing peer matching methods consider students' overall performance in a certain subject or analyze students' ability to an individual knowledge topic (Liang et al., 2021). Hence, we propose a method that utilizes connections among knowledge units to recommend problems and peer helpers.

### 3. A Method of Peer Help Recommendation

#### 3.1 Technical Framework and Model Construction

Learning and Evidence Analytics Framework (LEAF) presents an integrated learning platform that captures various learning activities and analyzes data to provide learners with the most appropriate support (Ogata et al., 2018). We build knowledge and learner models based on learning log data from the LEAF system. All these knowledge and learner models are built in a graph database application called Neo4j, and we use Python to access the Neo4j database and execute calculations.



Figure 1. Learner Model and Knowledge Model

The knowledge model is a network where knowledge elements are drawn out from the LEAF system by analyzing the learning materials. These learning materials include textbooks, practice problems, etc. learned by the learners. As shown in Figure 1, knowledge elements are represented in the network as knowledge nodes. If there is relevance between two knowledge elements, these two knowledge nodes are connected by edges. Each edge connecting two knowledge nodes possesses an attribute called distance. The distance between the knowledge nodes indicates the degree of relevance. This knowledge model can thus represent complex topological associations between subject knowledge.

The learner model is constructed incorporating the knowledge model. The LEAF system stores learning analytics indicators from learners' learning records in various individual learning activities. The sources of learning records include learners' online textbook readings, homework submissions, test scores, etc. As shown in Figure 1, each learner is represented as a learner node in the network. A learning record of the learner is represented as a log node that connects some knowledge nodes. One log node indicates that the learner has learned some knowledge once. A new log node points to the most recent log node through a directed edge. Each edge connecting log nodes and knowledge nodes has an attribute called proficiency. The proficiency attribute represents a changed value of the learner's proficiency for this knowledge during this learning activity. Each log node is equivalent to a differential snapshot of the learner at different stages of the learning process.

### 3.2 Recommendation Based on Closeness Centrality of Knowledge Node

The recommendation method proposed in this paper creates learner-learner connections beyond the existing knowledge and learner models mentioned above, indicating potential peer helpers for each learner. As each learner model is associated with the same knowledge model, it is possible to find learner-learner connections through the knowledge structure. The proficiency attributes for the same knowledge node by each learner are summed up as the learner's proficiency for this knowledge node. The proficiency is a numerical estimate to represent the quality of the student's specific knowledge (Hoppe & Ploetzner, 1999). If the proficiency of a knowledge node is below a certain level, but above a minimum value, it can be predicted that this learner has learned but does not understand this knowledge very well. So, these knowledge nodes are defined as weak knowledge nodes.

To find out some focused nodes among the weak knowledge nodes associated with a learner, the method of calculating closeness centrality in the weighted network is used. Closeness centrality is a measure of the distance from all other nodes in the network, hence the focused nodes tend to have high closeness centrality scores (Saqr et al., 2020). In the model of this paper, suppose there are a total of N knowledge nodes in the chosen part, each denoted as  $K_n$  ( $1 \le n \le N$ ). Define  $distance(K_n, K_i)$  between node  $K_n$  and node  $K_i$  as the sum of the distance attributes of the edges included in the shortest path of these two nodes. Thus, the closeness centrality of node  $K_i$  is defined as the following formula.

closeness centrality = 
$$\left[\sum_{n=1}^{N} distance(K_n, K_i)\right]^{-1}$$

The system chooses some knowledge nodes with the highest closeness centrality as candidate problems for this learner. After that, the system compares the proficiency of these knowledge nodes of other learners. If some learners have the proficiency of these knowledge nodes greater than a certain level, they are selected as candidate helpers by the system. The cut-off proficiency level, as well as the number of candidate problems and helpers, can be set flexibly based on purpose. When the candidate problems and helpers are ready, this learner can choose them from the system to get help. The following subsection will give a specific example of the method to explain the whole process.

### 3.3 Peer recommendation in physics learning

This subsection presents an example of physics learning at the high school level. The LEAF system extracted knowledge elements from the textbook they are learning to create the knowledge model. After these students have taken several quizzes and submitted some assignments to the LEAF system, the learner model will be created from these learning records. Thus, the models are prepared for recommending suitable problems and helpers for learners.

A portion of the knowledge model associated with Student A, B, C, D, E are shown in Figure 2, Student A has proficiency greater than 0.7 for "velocity", "uniform rectilinear motion" and others, and the system determines that Student A is already familiar with these knowledge topics. However, Student A has proficiency less than 0.5 and greater than 0.1 for "projectile motion", "horizontal projectile motion" and others, and the system identifies these as weak knowledge nodes for Student A. The system then calculates that "horizontal projectile motion" has the highest closeness centrality among these weak knowledge nodes, followed by "free-fall motion". Therefore, the system orderly selects these 2 weak knowledge nodes as the candidate problems that Student A needs help with.



Figure 2. Knowledge Model Associated with Students

The system looks for the proficiency of other students in these 2 knowledge nodes, and finds that Student B, C have high proficiency (greater than 0.7) for "horizontal projectile motion". Besides, Student B, C, D have high proficiency for only "free-fall motion". However, Student E does not have high proficiency for these 2 knowledge nodes. In summary, the system displays the contents of Table 1 in the user interface to Student A, where Student A can select one to get help.

Student	Can Help	Proficiency
В	horizontal projectile motion	0.85
С	horizontal projectile motion	0.76
В	free-fall motion	0.96
С	free-fall motion	0.84
D	free-fall motion	0.73

Table 1. The Options of Help Recommendation

In Figure 3, the "horizontal projectile motion" node has the best closeness centrality, and if Student B and C are familiar with "horizontal projectile motion", theoretically, they could help

explain to Student A about "decomposition of displacement" and "free-fall motion". However, some of them may not be familiar enough with "projectile motion". Finding a peer helper by focused nodes may not guarantee that all the weak knowledge nodes of Student A will be solved. However, at the same time, the helpers may explain some other related knowledge, which also benefits Student A. Such a recommendation method is an open practice of peer help, which means that it does not quantitatively predict exactly what goals should be accomplished, nor does it fully control how learners interact. Even in the case of a misdiagnosis, we can expect human-human communication to provide more flexible and ultimately more successful repair strategies (Suchman, 1987).

Finally, it should be noted that in practical implementations, the scope of knowledge and learners to be considered can be limited accordingly. For example, the scope of knowledge can be limited to knowledge related to a certain topic. Also, the range of learners can be limited to a single class, as in reality, they are more accessible.

### 4. Discussion and Future Work

The recommendation method in this paper is a more general way of analyzing and does not specifically analyze the characteristics of different knowledge. The characteristics of different knowledge can have an impact on the accuracy of recommendation algorithms. Closeness centrality has a lack of applicability to networks with disconnected components (Opsahl et al., 2010). We need to consider different structures for modeling knowledge from different subjects.

Besides, the method in this paper does not fully utilize the complex topological network structure of the knowledge model and the learner model when recommending problems and helpers. Other methods such as clustering may be better to select multiple problems for learners. The method in this paper initially considers creating learner connections through knowledge connections. However, we have simply considered the structural connections of knowledge but not the intrinsic associations that can imply a learning order of knowledge. If we have analyzed the learning order, a directed network can be used to represent it.

Additionally, it is challenging to mediate the interaction between learners in collaborative learning and peer help. The recommendation method in this paper considers the knowledge of the learners and assigns the roles of being helped and helping, which guides the learners' interaction to some extent. Individual tutoring ideally takes place in a "closed loop", but this is practically not feasible in rich peer interactions (Hoppe & Ploetzner, 1999). Thus, in this paper, the intelligent support system is more open that allows a learner to choose helpers as well as problems with high flexibility instead of being passively assigned. In other words, the proposed system assigns roles of peer helpers as recommendations but does not force the interaction to happen, which is aimed at diminishing the dominant role of the system and increasing the initiative of learners.

In future work, we will improve the knowledge model to make it a directed network where the pointers between knowledge nodes will be used to indicate a rational order in which knowledge is learned. Actually, there are variants for directed networks to calculate closeness centrality (Lü et al., 2016). Also, we shall look for other useful algorithms for recommending problems and helpers, making it more theoretically possible for helpers to solve problems. Finally, we will design a complete system with a user interface to implement our method in a real class. We will examine the effectiveness of our method using the peer evaluation system in LEAF (Liang et al., 2022), to figure out the advantage of the current system beyond peer matching with only learner models or random allocation.

### Acknowledgements

This research was supported by the following grants: JSPS KAKENHI 22H03902, NEDO JPNP20006, JPNP18013, and JST JPMJSP2110.

### References

- Abnar, S., Orooji, F. & Taghiyareh, F. (2012). An evolutionary algorithm for forming mixed groups of learners in web based collaborative. IEEE International Conference on Technology Enhanced Education (ICTEE), 1–6.
- Advanced Distributed Learning (2016, August 12). Experience API. https://github.com/adInet/xAPI-Spec/blob/1.0.3/xAPI-About.md#def-activity.
- Arroyo, S., Bruijn, J. D. & Ding, Y. (2003). Ontology Mapping and Aligning. STI Innsbruck.
- Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. Voogt, & G. Knezek, (Eds.), International Handbook of Information Technology in Primary and Secondary Education (pp. 1353–1369). Springer.
- Brusilovsky, P., Somyurek, S., Guerra, J., Hosseini, R., Zadorozhny, V., & Durlach, P. J. (2016). Open social student Modeling for Personalized learning. IEEE Transactions on Emerging Topics in Computing, 4(3), 450–461.
- Chrysafiadi, K. & Virvou, M. (2013). Student modeling approaches: a literature review for the last decade. Expert Systems with Applications, 40(11), 4715–4729.
- Dillenbourg, P. (1999). What do you mean by collaborative learning?. In P. Dillenbourg (Ed.), Collaborative Learning Cognitive and Computational Approaches (pp. 1–19). Oxford: Elsevier.
- Hoppe, H. U. & Rolf Ploetzner (1999). Can Analytic Models Support Learning in Groups?. In P.
  Dillenbourg (Ed.), Collaborative-learning: Cognitive and Computational Approaches (pp. 147– 168). Oxford: Elsevier.
- Hoppe, H. U. (1995). Using multiple student modeling to parameterize group learning. In J. Greer (Ed.), Proceedings of the Seventh World Conference on Artificial Intelligence in Education, (pp. 234–241). Charlottesville, VA: Association for the Advancement of Computing in Education.
- Isotani, S., Inaba, A., Ikeda, M., & Mizoguchi, R. (2009). An ontology engineering approach to the realization of theory-driven group formation. Computer-Supported Collaborative Learning, 4, 445–478.
- Liang, C., Majumdar, R. & Ogata, H. (2021). Learning log-based automatic group formation: system design and classroom implementation study. Research and Practice in Technology Enhanced Learning, 16:14.
- Liang, C., Majumdar, R., Nakamizo, Y., Flanagan, B. & Ogata, H. (2022). Algorithmic group formation and group work evaluation in a learning analytics-enhanced environment: implementation study in a Japanese junior high school. Interactive Learning Environments.
- Lü, L., Chen, D., Ren, X., Zhang, Q., Zhang, Y. & Zhou T. (2016). Vital nodes identification in complex networks. Physics Reports, 650, 1–63.
- Majumdar, R., Akçapınar, A., Akçapınar, G., Ogata, H., & Flanagan, B. (2019). Laview: Learning analytics dashboard towards evidence-based education. In Companion Proceedings of the 9th International Conference on Learning Analytics and Knowledge. Society for Learning Analytics Research (SoLAR).
- Meng, B., Xiao, Y., Tang, T. & Xie, R. (2018). Study and Enlightenment of Khan Academy's Data Platform Function and Data Analysis Framework to University Library's Data based Service Improvement. Library Science Research & Work, 0(9), 41–47.
- Ogata, H., Majumdar, R., Akçapinar, G., Hasnine, M.N. & Flanagan, B. (2018). Beyond learning analytics: Framework for technology-enhanced evidence-based education and learning. Proceedings of the 26th International Conference on Computers in Education, (pp. 493–496).
- Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks, 32(3), 245–251.
- Ounnas, A., Davis, H. C., & Millard, D. E. (2009). A Framework for Semantic Group Formation in Education. Educational Technology & Society, 12(4), 43–55.
- Quillian, M. R. (1968). Semantic networks. Approaches to Knowledge Representation Research Studies, 23(92), 1–50.
- Saqr, M., Nouri, J., Vartiainen, H. & Malmberg, J. (2020). What makes an online problem-based group successful? A learning analytics study using social network analysis. BMC Medical Education, 20:80.
- Suchman, L. A. (1987). Plans and situated actions. The problem of human-machine communication. Cambridge, MA: Cambridge University Press.
- Urhahne, D., Schanze, S., Bell, T., Mansfield, A., & Holmes, J. (2010). Role of the teacher in computer-supported collaborative inquiry learning. International Journal of Science Education, 32(2), 221–243.

# Conceptual Design of WHALE: A Wise Helper Agent for the LEAF Environment

Kento KOIKE<sup>a\*</sup>, Rwitajit MAJUMDAR<sup>a</sup>, H. Ulrich HOPPE<sup>a</sup> & Hiroaki OGATA<sup>a</sup> <sup>a</sup> Academic Center for Computing and Media Studies, Kyoto University, Japan \*koike.kento.7y@kyoto-u.ac.jp

**Abstract:** This approach introduces and elaborates the idea of combining learning path recommendations with the use of pedagogical agents in computer-based learning environments. For this purpose, we conceptualize and propose a pedagogical agent named WHALE (Wise Helper Agent for the LEAF Environment) together with a suite of multiple educational tools enhancing the LEAF learning platform. The main idea is to generate learning path recommendations based on a multi-layered background model comprising an action sequence layer, a resource item layer, and a knowledge graph layer. We expect the acceptability of recommendations to be enhanced by transmitting them through a personalized pedagogical agent. Based on this idea, we describe the design of a pedagogical agent that can recommend what to learn next, what materials to use for that purpose, and what to do with the materials in a multi-facetted educational tool environment.

Keywords: Pedagogical agent, learning path recommendation, knowledge graph

### 1. Introduction

To guide students in their learning in online environments generating recommendations for the learning process itself (i.e., what to learn next) beyond micro-level assistance within given tasks and exercises is important but difficult to implement. Coordinating these activities into a continuous learning process can be a daunting task without assistance (Zimmerman, 2002). Such assistance is usually provided by teachers, but there are challenges: teachers' resources are limited and online learning, where teachers are not usually there, is growing.

Pedagogical agents, digital entities designed to support learning, have the potential to solve this problem. Pedagogical agents are attractive to learners due to their social presence (Haake & Gulz, 2008; Mayer & DaPra, 2012; Woo, 2009), which promotes learners' acceptance of recommendations and other instruction (Castro-Alonso et al., 2021).

However, scaffolding by pedagogical agents has been limited to domain-specific application contexts so far. Usually, pedagogical agents have functionalities to provide pedagogical assistance in the form of scaffolds, such as nudges, explanations, feedback, prompts, support for metacognitive reflection/planning, or error reports/hints (Martha & Santoso, 2019). According to VanLehn (2006), these scaffolds can be classified into two types: guiding problem-solving steps within specific problems (inner-loop scaffolding) and guiding learners to an appropriate next unit through sequential problems or item sets (outer-loop scaffolding).

In order to enable pedagogical agents to guide learning activities, they need to be equipped with capabilities to adaptively sequence multiple learning-related activities as well as to support problem-solving. To complete the learning process, other important activities besides problem-solving, such as reading, taking notes, discussing with peers, self-reflection have to be considered and included. For example, if an educational agent recommends a quiz to a learner with no evidence of prior knowledge, the educational value is limited because the quiz would not be a meaningful indicator for formative assessment.

Therefore, our suggestion aims to address these challenges by developing a pedagogical agent capable of adaptively supporting individual learning paths for expanding learners' knowledge by identifying and recommending adequate learning actions. The

ultimate objective is the development of a pedagogical agent that can guide learners in selecting appropriate activities, considering the array of available options, the learner's habit, and the learner's both state of knowledge and affect. As a part of this objective, in this paper, we proposed a first conceptual design of an agent for guiding the learner's learning path in the LEAF environment (Ogata et al., 2023) which is a suite of multiple educational tools enhancing a learning platform. We design the concept of our agent namely WHALE (Wise Helper Agent for the LEAF Environment) which is an agent for guiding a learner's next action that considers both the learner's habitual learning flow and learning target (topic)-specified general learning flow. We believe that learners will learn what to learn next through our recommendations and be able to adjust their learning paths.

### 2. Background

Several pedagogical agents have been proposed and their effectiveness has been verified. For instance, there are AutoTutor (Graesser et al., 2005), Betty's Brain (Biswas et al., 2005; Leelawong & Biswas, 2008), and SimStudent (Matsuda et al., 2013) as well-known agents. The benefits of introducing pedagogical agents as metaphors, in general, are widely known. In particular, the explicit presence of agents is known to give learners a sense of social presence. Equipping social presence is more effective than only show prompting and feedback, and eliciting engagement (Haake & Gulz, 2008; Mayer & DaPra, 2012; Woo, 2009). Also, it is known that in the recommendation context including education, the existence of an agent affects the acceptance of recommendations (Castro-Alonso et al., 2021).

However, there have been no attempts to assist in the sequencing of learning activities. Accordingly, clarifying what kind of recommendations can be accepted in comprehensive guidance may be an important contribution to the study of pedagogical agents. Learning requires not only resolving impasses in problem-solving or making decisions about specific activities but also deciding what to learn next and how to do it from a more holistic viewpoint. Until now, pedagogical agents have been limited to assisting in the problem-solving process and recommending problems to be solved. Thus, the range of recommendations and support that has been handled by pedagogical agents has been limited.

Individual differences in sequencing learning activities can affect the learner's academic achievement and efficiency (Pintrich, 2000). Such learning sequence have also been conceptualized as learning paths (or learning pathways), i.e., the learner's developmental trajectories towards reaching given learning goals (Cheng et al., 2023; Ortiz-Vilchis & Ramirez-Arellano, 2023; Pintrich, 2000). The selection of adequate learning paths has been based on Vygotsky's notion of Zone of Proximal Development (ZPD) as a theoretical background (Shabana et al., 2022). ZPD theory distinguishes between problems that learners can solve on their own, problems that they can solve with help, and problems that they cannot solve even with help, depending on their learning situation. Furthermore, this theory promotes development by having learners work on the second one with support. If the learner's current knowledge is represented as a subset of a general knowledge graph, the nodes most promising for "knowledge expansion" can be selected from the outer fringe of the student's current knowledge. The next knowledge item should not yet be part of the learner's knowledge, yet should be maximally connected to nodes that have already been learned. This idea has, e.g., been operationalized using the theory of Knowledge Spaces (Falmagne et al., 2013) and recently using a "curriculum graph" in combination with an optimization strategy (Shabana et al., 2022). For instance, methods for giving learning path are proposed with: graph network (Raj & Renumol, 2022); ontology description (Rahayu et al., 2023); bayesian knowledge tracing (Pelánek, 2017); model tracing (Blessing, 1997); and so on. Also, each of these methods depends on some granularity: steps in problem-solving (Blessing, 1997); problem in problem sequence (Falmagne et al., 2013; Raj & Renumol, 2022), and item bag (Pelánek, 2017); problems in a unit(topic) of curriculum (Shabana et al., 2022); but no method of action level.

However, in the context of learning path recommendations, the question is not only which knowledge should be learned next but also how knowledge should be acquired using available resources and actions in the environment. Effectively coordinating the timing and action among various learning activities presents its own set of challenges for learners. To form a learner's learning path, thus, it is important to decide both what to learn and how to learn.

To design the pedagogical agent with recommendations that consider both what to learn and how to learn, a learning environment with multiple options for both topic and learning activities is needed in advance. One such learning environment is the LEAF system, a large-scale educational platform (Ogata et al., 2023). LEAF is a packaged platform that can be extended through LTI connections to an LMS system. An LMS with the LEAF platform comprises multiple modules supporting different learning activities: (1) regular LMS activities (e.g., forum discussions and assignment submissions); (2) learning activities in an e-book reader (e.g., highlighting and memo functions) (Kannan et al., 2022); (3) learning activities in the GOAL subsystem supporting individual learning management across formal and informal spaces (Li et al., 2021); and (4) learning activities in the Al-enabled EXAIT subsystem (e.g., recommending problems to be solved, self-explanation of solved problems, etc.) (Flanagan et al., 2021); (5) learning activities in GLOBE supporting cooperative learning based on automated group formation (Liang et al., 2022). The LEAF system is equipped with multiple learning activities and their support functions, but since there are multiple types of learning activities, it is difficult for the learner to coordinate the activities as in the before discussion.

On the other hand, LEAF's characteristic that all activities in all tools are recorded in the form of xAPI logs can be advantageous for the design of general recommendations. By recording the habitual learning activities of learners, it is possible to extract examples of what worked well in the aggregate and to attempt perturbations from individual habits. Furthermore, the LEAF system has been installed in several schools and is used daily. Therefore, a large amount of data necessary for designing and training the agent that this research aims for has already been accumulated.

# 3. Conceptual Design of WHALE for Learning Action Path Recommendation

We focused on developing a pedagogical agent capable of supporting learners' learning path for expanding learners' knowledge and options for learning actions. Thus, we propose a concept of a pedagogical agent named WHALE that combines the technologies of a pedagogical agent and learning path recommendation in the LEAF environment. To make learning action path recommendations, at the level of knowledge expansion, WHALE stands on ZPD theory. Falmagne et al. (2013) and Shabana et al. (2022), by applying the theory of ZPD to the knowledge map, state that knowledge nodes with links to the learner's current knowledge nodes (called "fringe") are the easiest knowledge to learn. Therefore, WHALE also aims for the learner to learn knowledge in a way that the node is approached step by step from the learner's current knowledge node toward the knowledge node that is set as the destination. However, the remaining challenges are which material and how to learn. To develop a pedagogical agent for addressing these challenges, we raised some working hypotheses as follows: (1) Certain actions on certain material can learn certain knowledge nodes: (2) Between materials and knowledge nodes can make N-to-N relations; (3) Each material limits allowed actions; (4) actions can be discussed independently from materials; (5) There can be an ideal sequence of actions represented as a semi-ordered structure. Based on these working hypotheses, we propose multi-layered learning path recommendations combined from three layers: an action sequence layer, a resource item layer, and a knowledge graph layer (Fig. 1).

The knowledge graph layer is a layer that represents the so-called knowledge graph, a graph that expresses the connections between knowledge and its parent-child relationships.

In several studies, recommendations based on the knowledge graph have been realized (Falmagne et al., 2013; Shabana et al., 2022).

The resource item layer is a layer where items such as textbooks, workbooks, class videos, and teacher's class materials reside. These resource items are the learning contents uploaded in the LEAF environment. There is no connection between these items in this layer alone, and the order relationship in the resource item layer is inferred through the knowledge graph layer. In addition, there is usually an N-to-N relationship between resource items and nodes in the knowledge graph. This is because a textbook contains multiple pieces of knowledge and there are multiple learning materials for a given piece of knowledge.

The action sequence layer is a layer in which a sequence of learned actions is represented as a directed graph. The learning actions are based on the primitive actions that a recommender system can represent. In the LEAF environment, we are considering representing them based on the type of verb represented by xAPI. We also assume that learning actions are sequential: for example, it should make little sense to perform a formative test when no knowledge has been learned. The types of learning actions that can be taken depend on the resource items: simply reading the question booklet or solving the teacher's material is unlikely to be considered.

By implementing a reasoning mechanism based on these three layers, WHALE aims to provide step-by-step recommendations for specific primitive actions to acquire knowledge of goals set by the learner, teacher, and system. In other words, finally, we believe that WHALE can recommend what to learn next, what materials to use for that purpose, and what to do with the materials.



Figure 1. Concept of three layers for inference of recommendation.

As an example, our architecture supports a kind of "means-ends analysis" that spells out through the three layers starting with the identification of "ZPD anchors" and generating the actual suggestions by a mapping from the knowledge layer to the action (sequence) layer. This is the generative part of the process, which can be broken down into the following steps:

- (1) The knowledge graph is static in its underlying structure, but dynamically annotated according to the individual learner's assumed knowledge state (overlay type student model).
- (2) The ZPD / "fringe" analysis selects candidate nodes (knowledge items) for knowledge expansion.
- (3) Based on the selected knowledge item, learning resources supporting this knowledge item are selected (making use of metadata labels such as "Introduction to knowledge X").
- (4) The selected resources are suggested to be accessed using appropriate tools or apps, possibly also including usability support on the concrete action level.

Furthermore, we propose the following strategy to diagnose and update knowledge states based on the results of a certain action of the learner and the connections between each action:

- (1) When any action is performed by the learner, (regardless of whether it is a recommended action or not), it is represented in the model as having gained some degree of experience with that knowledge.
- (2) After that, effective/connective action is diagnosed and recommended based on the action sequence.
- (3) By repetition of (1) and (2), when the "number of experiences" of the knowledge node or when the percentage of correctness to the quiz related to the knowledge exceeds a certain level, the next knowledge node is recommended based on the ZPD fringe.

Our agent equipped with these two mechanisms is provided along with an anthropomorphic agent, and the learner interacts with this agent by selecting recommended actions. For example, the agent can provide these options for the learner's inquiry: "I want to set a goal.", "I do not know what to do next.", "Which knowledge would be good for the next learning?", etc.

### 4. Discussion and Future Work

This paper describes the design of WHALE, a pedagogical agent for learning action path recommendation. However, several issues remain to be solved in order to implement and deploy WHALE. First, it is necessary to introduce and implement the three layers (an action sequence layer, a resource item layer, and a knowledge graph layer) as part of the LEAF environment. For this purpose, we believe that both extraction of the most frequently occurring patterns in each resource item and top-down modeling are necessary for the action sequence. For resource items, from a practical point of view, it is necessary to be able to define semi-automatically what kind of action each resource item is associated with. As for the knowledge graph, we believe that technology to automatically extract the hierarchy of knowledge from the table of contents of textbooks and other sources will be a central issue.

In addition, although we mainly discussed the recommendation mechanism in this paper, we have not yet discussed how to utilize social cues by pedagogical agents. This is also an issue to be addressed in the future.

### Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers JP23H00505, and JP22H03902.

### References

- Biswas, G., Leelawong, K., Schwartz, D., Vye, N., & at Vanderbilt, T. T. A. G. (2005). Learning by teaching: A new agent paradigm for educational software. *Applied Artificial Intelligence*, *19*, 363–392.
- Blessing, S. B. (1997). A programming by demonstration authoring tool for model-tracing tutors. *International Journal of Artificial Intelligence in Education*, *8*(3-4), 233–261.
- Castro-Alonso, J. C., Wong, R. M., Adesope, O. O., & Paas, F. (2021). Effectiveness of multimedia pedagogical agents predicted by diverse theories: A meta-analysis. *Educational Psychology Review*, 33, 989–1015.
- Cheng, S.-C., Cheng, Y.-P., & Huang, Y.-M. (2023). Developing a learning pathway system through web-based mining technology to explore students' learning motivation and performance. *Sustainability*, *15*, 6950.
- Falmagne, J.-C., Albert, D., Doble, C., Eppstein, D., & Hu, X. (2013). *Knowledge spaces: Applications in education*. Springer Science & Business Media.
- Flanagan, B., Takami, K., Takii, K., Dai, Y., Majumdar, R., & Ogata, H. (2021). EXAIT: A symbiotic explanation learning system. *29th International Conference on Computers in Education Conference, ICCE 2021*, 404–409.
- Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education*, *48*(4), 612–618.
- Haake, M., & Gulz, A. (2008). Visual stereotypes and virtual pedagogical agents. *Journal of Educational Technology & Society*, *11*(4), 1–15.
- Kannan, V., Warriem, J. M., Majumdar, R., & Ogata, H. (2022). Learning dialogs orchestrated with BookRoll: Effects on engagement and learning in an undergraduate physics course. *Research and Practice in Technology Enhanced Learning*, *17*(1), 1–22.
- Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: The Betty's Brain system. *International Journal of Artificial Intelligence in Education*, *18*, 181–208.
- Li, H., Majumdar, R., Chen, M.-R. A., & Ogata, H. (2021). Goal-oriented active learning (GOAL) system to promote reading engagement, self-directed learning behavior, and motivation in extensive reading. *Computers & Education*, *171*, 104239.
- Liang, C., Majumdar, R., Nakamizo, Y., Flanagan, B., & Ogata, H. (2022). Algorithmic group formation and group work evaluation in a learning analytics-enhanced environment: Implementation study in a japanese junior high school. *Interactive Learning Environments*, 1–24.
- Martha, A. S. D., & Santoso, H. B. (2019). The design and impact of the pedagogical agent: A systematic literature review. *Journal of Educators Online*, *16*(1), 87–101.
- Matsuda, N., Yarzebinski, E., Keiser, V., Raizada, R., Cohen, W. W., Stylianides, G. J., & Koedinger, K. R. (2013). Cognitive anatomy of tutor learning: Lessons learned with SimStudent. *Journal of Educational Psychology*, 105, 1152.
- Mayer, R. E., & DaPra, C. S. (2012). An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology: Applied*, *18*(3), 239.
- Ogata, H., Majumdar, R., & Flanagan, B. (2023). Learning and evidence analytics framework bridges research and practice for educational data science. *Communications of the ACM*, 66(7), 72–74.
- Ortiz-Vilchis, P., & Ramirez-Arellano, A. (2023). Learning pathways and students performance: A dynamic complex system. *Entropy*, 25, 291.
- Pelánek, R. (2017). Bayesian knowledge tracing, logistic models, and beyond: An overview of learner modeling techniques. *User Modeling and User-Adapted Interaction*, 27, 313–350.
- Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92, 544.
- Rahayu, N. W., Ferdiana, R., & Kusumawardani, S. S. (2023). A systematic review of learning path recommender systems. *Education and Information Technologies*, 28(6), 7437–7460.
- Raj, N. S., & Renumol, V. G. (2022). An improved adaptive learning path recommendation model driven by real-time learning analytics. *Journal of Computers in Education*, 1–28.
- Shabana, K., Lakshminarayanan, C., & Anil, J. K. (2022). CurriculumTutor: An adaptive algorithm for mastering a curriculum. *International Conference on Artificial Intelligence in Education*, 319–331.
- VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education*, *16*(3), 227–265.
- Woo, H. L. (2009). Designing multimedia learning environments using animated pedagogical agents: Factors and issues. *Journal of Computer Assisted Learning*, 25(3), 203–218.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*, 64–70.

# W03

# The Applications of Information and Communication Technologies (ICTs) in Adult and Continuing Education

# Investigating Students' Perceptions of Knowledge-building Environment and Learning Engagement

Wei-Shou CHEN a\*, Min-Hsien LEE b & Jyh-Chong LIANG b

 <sup>a</sup> Graduate Institute of Information and Computer Education, National Taiwan Normal University, Taiwan
 <sup>b</sup> Program of Learning Sciences and Institute for Research Excellence in Learning Sciences, National Taiwan Normal University, Taiwan
 \*max0616@gmail.com

Abstract: Knowledge-Building pedagogy emphasizes engaged learning, with students collaboratively working towards advancing shared ideas within a community. This study explored the relationship between students' perceptions of the knowledge-building learning environment and their learning engagement in Knowledge Forum activities. A total of 259 undergraduate students from courses that used Knowledge Forum participated in this study. The study employed partial least squares structural equation modeling (PLS-SEM) analysis to examine the relationship. The findings revealed that assuming agency positively influenced students' cognitive, emotional, and social engagement, indicating that students who exhibited a sense of autonomy and enjoyment demonstrated higher levels of engagement. Additionally, students' perceptions of fostering community positively predicted cognitive and social engagement, underscoring the importance of students' sense of identification with the community. These results provide potential insights for designing or refining knowledge-building learning environments and activities, highlighting the significance of promoting student agency and fostering a sense of community to facilitate learning engagement.

**Keywords:** Perceptions of learning environment, knowledge-building, learning engagement, PLS-SEM

### 1. Introduction

Knowledge-building is a socio-constructivist approach to learning that places students' ideas at the center and aims to advance community knowledge (Scardamalia & Bereiter, 2014, 2022). In knowledge-building environments, participants are encouraged to actively contribute to developing and enhancing their community's shared ideas and knowledge. Knowledge Forum (KF) is a platform specifically designed to implement knowledge-building pedagogy, providing a digital platform that enables students to engage in knowledge creation and innovation online. Students can use this platform to contribute their ideas, build on each other's contributions, and even elevate their collective thinking to a higher level using the rise above function. In Taiwan, an increasing number of university classrooms are adopting Knowledge Forum to provide students with a knowledge-building learning environment.

In research on technology-enhanced learning environments, understanding students' perceptions of the learning environment is crucial (Chang et al., 2015). By understanding students' perceptions of the learning environment, we can gain insights into the characteristics of the learning environment and further improve it. Understanding students' perspectives becomes even more important in a learning environment like Knowledge Forum, which relies on student communication and collaboration. To assess students' perceptions of knowledge building environment, Lin et al. (2014) categorized the 12 knowledge-building principles by Scardamalia (2002) into three categories: working with ideas, assuming agency, and fostering community, and further developed a scale based on this framework.

In a platform like Knowledge Forum, students' continuous engagement and contributions are also crucial for knowledge-building activity (Chan & van Aalst, 2018). Learning engagement has been a widely discussed topic, with various dimensions of engagement proposed. For example, Fredricks et al. (2004) categorized school engagement into three types: behavioral, emotional, and cognitive. However, behavioral engagement has limited predictive power for higher-level cognitive activities (Sinatra et al., 2015). In a knowledge-building environment that emphasizes students' collaborative creation of knowledge, behavioral engagement is relatively less emphasized. Instead, social engagement becomes a practical form in such a learning environment that deserves attention since it is closely tied to the interaction within the community. A qualitative study by Zhu et al. (2021) found that refining the design of knowledge-building courses in which students participate can positively impact their emotional, cognitive, and social engagement.

In sum, students' perceptions of the learning environment are closely related to their learning engagement (Opdenakker & Minnaert, 2011). However, there is limited quantitative research investigating the relationship between knowledge-building learning environments and learning engagement. By understanding the relationship between students' perceptions of the knowledge-building learning environment and their learning engagement, we can gain more insights into the design of activities on the knowledge forum. To fulfill this objective, we seek to adapt and validate the knowledge-building environment scale (KBES) developed by Lin et al. (2014) with students who have experience using Knowledge Forum by addressing the following questions: What are the structural relationships between the different indicators of KBES and students' cognitive, emotional, and social engagement?

### 2. Methodology

### 2.1 Participants

A total of 259 university students in Taiwan from various courses participated in this survey. All of the courses utilized the Knowledge Forum platform for their class activities. Basically, teachers provide weekly topics such as Socio-Scientific Issue for students to discuss. Students are required to make contributions, which is posting content, and build on, which is replying to others' contributions on the Knowledge Forum platform. Students can participate by accessing the platform through various electronic devices (i.e., PC, Tablet, or mobile phone). Through these discussions, students develop a comprehensive understanding and gain deeper insights into the topics. The teacher or teaching assistants assess students' weekly contributions and provide grades based on their content.

### 2.2 Instruments

After participating in a semester-long course, each student completed two questionnaires to assess their perceptions of Knowledge-Building environment and learning engagement. The details of the questionnaires are described below.

### 2.2.1 Knowledge-building environment scale

Knowledge-building environment scale is a questionnaire developed by Lin et al. (2014) based on the 12 knowledge-building principles. The original questionnaire consists of three constructs, which are working with ideas (e.g., In this course, it is important to embrace divergent ideas.), assuming agency (e.g., In this course, one need to plan and execute one's learning plan.), and fostering community (e.g., In this course, all members have to actively participate in discussion.), with a total of 24 items. The internal consistency (Cronbach's alpha) coefficients for the three constructs are .84, .88, and .91, respectively. Each item uses a 5-point Likert scale from 1 = strongly disagree to 5 = strongly agree for students to select.

### 2.2.2 Learning engagement instrument

This study utilized the science learning engagement instrument developed by Lin (2021) and made some minor modifications to align the context for this study. In the original questionnaire, Lin (2021) categorized five different types of learning engagement. For this study, three types were selected for use: cognitive engagement (e.g., I review the discussion question carefully and make sure I understand the content), emotional engagement (e.g., I

enjoy learning new things related to knowledge building learning), and social engagement (e.g., I can refer to other people's ideas). This scale has excellent reliability and validity, with the composite reliability value ranging from .89 to .92, and the average variance extracted value ranging from .67 to .80. Each item also uses a 5-point Likert scale from 1 = strongly disagree to 5 = strongly agree for students to select.

### 2.3 Data analysis

The partial least squares structural equation modeling (PLS-SEM) technique was employed to investigate the relationship between students' perceptions of knowledge-building environment and their learning engagement. Unlike covariance-based structural equation modeling (CB-SEM), partial least squares structural equation modeling (PLS-SEM) was a variance-based latent variable SEM technique. A confirmatory factor analysis (CFA) was conducted to ensure the reliability and validity of the data. Cronbach's alpha, composite reliability (CR), and average variance extracted (AVE) were obtained to calculate the items' reliability, convergent reliability, construct reliability, and internal consistency (Hair et al., 2011).

This study chooses PLS-SEM due to the limitation of the sample size. PLS-SEM uses a nonparametric bootstrapping method and can be estimated with a smaller sample size (Hair et al., 2011). Additionally, giving the study's exploratory nature, PLS-SEM was chosen to investigate the structural model for students' perceptions of knowledge-building environment and their learning engagement. This study used the SmartPLS 4 software to validate the proposed model.

### 3. Results

### 3.1 The measurement model

The measurement model was obtained by conducting a confirmatory factor analysis (CFA) through PLS-SEM analysis. As shown in Table 1, a total of six factors are in the proposed model. Five items were retained for the "Working with idea" and "Assuming agency" factor, and eight items were retained for the "Fostering community" factor. For engagement, four items were retained for the "Cognitive engagement" and "Emotional engagement" factors, and three items were retained for the "Social engagement" factor.

For items' reliability, most of the factor loadings are over 0.6. Therefore, the factor loadings of these items are acceptable. Besides, Cronbach's alpha values for each variable exceeded the minimum value of .7 (0.80 - 0.93), and the composite reliability (CR) for each variable exceeded the minimum value of .7 (0.87 - 0.94), indicating the data have excellent consistency and construct reliability. Regarding the average variance extracted (AVE) value, each variable exceeded the minimum value of 0.5 (0.62-0.75), indicating the data have a decent convergent validity.

, , , , , , , , , , , , , , , , ,					
Variables and items	Factor	CR	AVE	alpha	Mean
	Loadings			·	(S.D.)
Working with idea (ID)		0.93	0.70	0.89	4.37 (0.55)
ID 1	0.61				
ID 2	0.90				
ID 3	0.89				
ID 4	0.89				
ID 5	0.88				
Assuming agency (AG)		0.91	0.67	0.88	4.22 (0.53)
AG 1	0.84				
AG 2	0.80				
AG 3	0.82				
AG 4	0.81				
AG 5	0.82				
Fostering community (FC)		0.94	0.66	0.93	4.36 (0.53)

Table 1. The item factor loadings, CR, AVE, Cronbach's alpha values and the instrument variable descriptive statistics (n = 259)

FC 1	0.85				
FC 2	0.84				
FC 3	0.87				
FC 4	0.80				
FC 5	0.67				
FC 6	0.84				
FC 7	0.81				
FC 8	0.80				
Cognitive engagement (CE)		0.87	0.62	0.80	4.18 (0.51)
CE 1	0.77				. ,
CE 2	0.79				
CE 3	0.77				
CE 4	0.82				
Emotional engagement (EE)		0.92	0.75	0.89	3.77 (0.66)
EE 1	0.91				
EE 2	0.86				
EE 3	0.84				
EE 4	0.87				
Social engagement (SE)		0.90	0.75	0.83	4.28 (0.54)
SE 1	0.88				
SE 2	0.92				
SE 3	0.80				

Note: CR: composite reliability, AVE: average variance extracted

### 3.2 The structural model

The structural model was examined using a bootstrapping method with 5000 subsamples to establish the significance level for model examination. The results are shown in Figure 1 (only significant paths were drawn). We discuss the relationships by predictors in the following paragraphs.

### 3.2.1 Working with idea

The results indicate that students' perception of "Working with idea" does not predict any form of learning engagement. This might suggest that even if students perceive ideas as important in knowledge-building activities, they might not put much effort into participating. In other words, when it comes to learning engagement, expressing ideas or embracing different perspectives may be less crucial.

### 3.2.2 Assuming agency

The results indicate that students' perception of "Assuming agency" was a positive predictor for cognitive engagement (path coefficient = 0.31, p < .001), emotional engagement (path coefficient = 0.44, p < .001), and social engagement (path coefficient = 0.22, p < .01). In other words, students who perceive planning, reflecting and monitoring their learning process as more important in knowledge-building activities also exhibit significantly higher engagement across the three dimensions.

### 3.2.3 Fostering community

The results indicate that students' perception of "Fostering community" was a significantly positive factor for cognitive engagement (path coefficient = 0.34, p < .001) and social engagement (path coefficient = 0.42, p < .001). Specifically, students who perceive the importance of interacting with and engaging with others in the knowledge building process demonstrate higher cognitive and social engagement.



*Figure 1.* The structural equation model of the relationship among knowledge-building environment and learning engagement. (\*\*p < .01, \*\*\*p < .001)

### 4. Discussion

This study extends the relationship between students' perceptions of the learning environment and their learning engagement into a specific context of the knowledge building environment. Specifically, we examined the relationship between three core principles crucial in knowledge building and three different types of learning engagement. Knowledge building is an approach that differs from traditional teaching methods, which encourage students to produce and develop their ideas. Therefore, working with ideas is a critical concept in a knowledge-building environment. However, the findings of this study suggest that even if students value these ideas and respect, integrating others' ideas, it does not necessarily enhance their learning engagement. This result could be related to students' unfamiliarity with operating the Knowledge Forum platform. In other discussion activities, students could immediately express their ideas into texts, which adds a step and may decrease their level of engagement.

This finding further indicates that assuming agency is crucial in the knowledge-building learning environment. In this study, students with higher assuming agency tended to show positive cognitive, emotional, and social engagement in a knowledge building environment. Moreover, it was noteworthy that assuming agency positively predicts emotional engagement. In general, students may have lower motivation to participate in these knowledge-building activities as they are requested by the teacher. However, students with high levels of assuming agency can enjoy participating in such activities and engage in them autonomously. This is essential for the sustained development of a knowledge-building community.

Last but not least, we found that the perceptions of fostering community positively predict cognitive and social engagement. It is pretty intuitive that students who embrace the concept of community tend to exhibit higher levels of social engagement. As for cognitive engagement, it is possible that students perceive the community as meaningful, and thus they are motivated to refrain from presenting hasty ideas. Consequently, they engage in knowledge-building activities with a more careful and diverse perspective.

The study contributes to a model of factors relating to students' perceptions of the knowledge-building learning environment and their learning engagement. These findings provide insights for designing or refining a knowledge-building learning environment. However, it is worth noting that this study had a relatively small sample size, and the duration of students' engagement in knowledge-building activities was limited to one semester. Future

research could expand the sample size and investigate specific students over extended periods.

### Acknowledgements

This work was financially supported by the "Institute for Research Excellence in Learning Sciences" of the National Taiwan Normal University (NTNU) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan and by the National Science and Technology Council, Taiwan, under the following grant numbers: 109-2628-H-003 -001 -MY3, 110-2511-H-003 - 035, 112-2628-H-003 -003 -MY3.

### References

- Chan, C. K., & van Aalst, J. (2018). Knowledge Building: Theory, Design and Analysis. In Fischer, F., Goldman, S. R., Reimann, P. (Eds.), International Handbook of the Learning Sciences, p. 295-307. New York, NY: Routledge, 2018. https://doi.org/10.4324/9781315617572-29
- Chang, H. Y., Wang, C. Y., Lee, M. H., Wu, H. K., Liang, J. C., Lee, S. W. Y., ... & Tsai, C. C. (2015). A review of features of technology-supported learning environments based on participants' perceptions. *Computers in human behavior*, 53, 223-237. https://doi.org/10.1016/j.chb.2015.06.042
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: indeed a silver bullet. *Journal of Marketing Theory and Practice*, 19, 139–152. https://doi.org/10.2753/MTP1069-6679190202
- Lin, K. Y., Hong, H. Y., & Chai, C. S. (2014). Development and validation of the knowledgebuilding environment scale. *Learning and Individual Differences*, *30*, 124-132. https://doi.org/10.1016/j.lindif.2013.10.018
- Lin, T. J. (2021). Multi-dimensional explorations into the relationships between high school students' science learning self-efficacy and engagement. *International Journal of Science Education*, *43*(8), 1193-1207. https://doi.org/10.1080/09500693.2021.1904523
- Opdenakker, M. C., & Minnaert, A. (2011). Relationship between learning environment characteristics and academic engagement. *Psychological Reports*, *109*(1), 259-284. https://doi.org/10.2466/09.10.11.PR0.109.4.259-284
- Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. *Liberal education in a knowledge society*, 97, 67-98.
- Scardamalia, M., & Bereiter, C. (2014). Knowledge Building and Knowledge Creation: Theory, Pedagogy, and Technology. In R. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (Cambridge Handbooks in Psychology, pp. 397-417). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.025
- Scardamalia, M., & Bereiter, C. (2022). Knowledge Building and Knowledge Creation. In R. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (Cambridge Handbooks in Psychology, pp. 385-405). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108888295.024
- Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist*, *50*(1), 1–13. https://doi.org/10.1080/00461520.2014.1002924
- Zhu, G., Raman, P., Xing, W., & Slotta, J. (2021). Curriculum design for social, cognitive and emotional engagement in Knowledge Building. *International Journal of Educational Technology in Higher Education*, 18, 1-19. https://doi.org/10.1186/s41239-021-00276-9

# Exploring the Framing Effect of Drawing Task Instructions on Science-Major Novice Preservice Teachers' Technology-use Knowledge

Yung-Hsiang HUANG<sup>a</sup>, Yi-Sheng HUANG<sup>a</sup>, Shan-Hui SU<sup>a</sup>, & Chia-Ching LIN<sup>a\*</sup> <sup>a</sup> Graduate Institute of Science Education & Environmental Education, National Kaohsiung Normal University, Taiwan \*cclin.edu@gmail.com

**Abstract:** Digital technology-assisted teaching and learning is a trend in schools at all levels around the world, and beliefs about technology use are one of the predictors to see if, and to what extent, teachers would adopt technology in their classes. In the field of educational research, drawing tasks are one of the potential research tools that have been shown to reflect participants' educational beliefs. We propose that the instructions for a drawing task would frame participants' thinking, as it is difficult for participants to depict images outside the direction of instruction. In the current study, 44 science-majors entering preservice teachers were recruited, and a two-round drawing task along with a random-assigned experimental design were introduced to evaluate this hypothesis. The instruction of the first-round drawing was used to disrupt the framing effect of the instruction of the second-round drawing by activating the participants' target knowledge. The results indicated that the participants depicted more images of technology use in the second-round drawing after their technology knowledge had been activated, but the pedagogical theme was not significant. Methodological, theoretical, and practical implications of the results are discussed.

**Keywords:** digital technology-assisted teaching and learning, beliefs about technology use, a drawing task, framing effect, preservice teachers

### 1. Introduction

Digital technology-assisted teaching and learning is a trend in schools at all levels around the world, as it can not only make teachers' teaching more varied and effective, but students' learning can also be more profound due to the multi-modality characteristic of digital technology (e.g., Ertmer et al., 2015). In addition to research on how meaningful, effective, and efficient digital technology helps teachers' teaching and students' learning, researchers have also established theoretical frameworks to measure teachers' understanding and capabilities, such as TPACK (Mishra & Koehler, 2006). In another line of research, researchers are interested in beliefs about digital technology use and the relationship between technology-use beliefs and instructional effectiveness (Ertmer et al., 2015). Moreover, it has been shown that beliefs held by preservice teachers would influence their learning of how to teach in teacher education programs (Richardson, 2003).

Different terms have been alternatively adopted, such as beliefs (e.g., Ertmer et al., 2012), conceptions (e.g., Yeh et al., 2019), perceptions (e.g., Inaltekin, 2020; Minor et al., 2002), and knowledge (Mishra & Koehler, 2006) to represent the thoughts about technology-assisted teaching and learning. These terms tend to be used alternatively as the needed, but beliefs are used primarily in the current study since it is the most frequently used term to express teachers' thoughts about technology in drawing-based research.

Teachers' beliefs are defined as teachers' mental constructs, propositions, or premises about teaching issues (Richardson, 2003). One of the functions of educational

beliefs is to serve as a filter for teachers to deal with educational issues or information they encounter (Skott, 2015), which in turn would affect how well, deeply, or creatively they adapt technology to their teaching practices (Ertmer et al., 2015). Furthermore, it has been argued that beliefs about technology use are one of the predictors of whether teachers adopt technology in their classes (Hermans et al., 2008; Hew & Brush, 2007), and it can be adjusted by teacher education programs (e.g., Funkhouser & Mouza, 2013).

To assess teachers' beliefs, drawing tasks are a relatively new yet potential method (Chang et al., 2020). Besides, it has been proven that drawing can reflect the drawers' concepts, knowledge, or beliefs about things researchers would like to know (e.g., Elmas et al., 2011). A drawing task can bypass the condition in which subjects find it difficult to express themselves verbally or orally (Schraw & Olafson, 2015). Although there has been an increase in the number of drawing studies on preservice science teachers' beliefs about science teaching (e.g., Alkış Küçükaydın & Gökbulut, 2020; Minogue, 2010; Tatar, 2015) and on preservice teachers' beliefs about technology use (e.g., Funkhouser & Mouza, 2013), research on novice preservice science teachers' beliefs about digital technology integrated into science teaching that uses drawing as a form of assessment is still in its infancy. A recent and relevant study was conducted by Lin (2022). The researcher found that, as revealed by their drawings, the preservice mathematics teacher participants in his study seldom used technology as pedagogical tools. However, the results of Funkhouser and Mouza's (2013) study showed that preservice teachers had beliefs about teacher-centered technology use. The reason why the results of these two studies seem contradictory may have been due to different drawing instructions. Lin's drawing instructions (Lin, 2022) did not ask the participants to depict technology integrated into their teaching directly, i.e., "what are you doing in the mathematics classroom?", but the instructions in some other studies did (e.g., Funkhouser & Mouza, 2013; Yeh et al., 2019). There is a parsimonious explanation why the participants in Lin's study seldom depicted technology in their drawings, which is that they merely did not think of it due to the framing effect (e.g., Kahneman, 2011) derived from the instructions of the drawing task. According to this argument, it is plausible to infer that, in the situation where the drawing instructions do not direct participants to depict technology images, as in Lin's study, participants would not depict technology images in their drawings unless their technology-relevant knowledge is activated. In the current study, therefore, a two-round drawing task were adapted to investigate whether there was a framing effect in a drawing instruction and how to disrupt it.

### 2. Methodology

### 2.1 Participants

A total of 44 (21 female, 23 male) science-major sophomores at a normal university in southern Taiwan, who were enrolled in and had begun educational programs, were recruited in this research. Of the 44 participants, 14, 15 and 15 had majored in chemistry, biotechnology, and physics, respectively. In the coding period, it was found that one drawing in each group was irrelevant to the theme of the drawing task and so was removed from the data set.

### 2.2 Research Design

A two-round drawing task along with a random-assigned experimental design were adapted to investigate the hypothesis described above. The drawing task was conducted in groups at the end of the "Introduction to Education" course in the 2022-2023 academic year. The drawing time was unlimited. There were two groups in our study, the Technology group and the Experience group. Participants were divided into two groups and were asked to complete two drawings. They were randomly assigned to each group, and the assignment was counter-balanced based on participants' gender and departments. The drawing instructions of the two groups in the first-round drawing task were different. The first drawing instruction

for the Technology group played a role in disrupting the framing effect derived from the second-drawing instruction and was: "Please try to imagine the scene where technology or digital technology is integrated into a science course and try to depict it."; in contrast, the first drawing instruction for the Experience group was: "Please try to recall the most impressive science course you attended in your high school and try to depict it.", which was a controlled treatment.

Participants were asked to raise their hand when they finished the first drawing, and the researcher would deliver the second drawing paper to them. The instruction for the second drawing task was the same for the two groups: "Please try to imagine what you will do in a science course if you are a science teacher and try to depict it." The second-drawing instruction played a role of framing effect on both groups because there were no words about technology in the instruction of the second drawing for both groups, and it was expected that participants whose technology knowledge was not activated, i.e., the Experience group, would depict less technology items in their drawings than those in the Technology group. See Figure 1 as an illustration of the research design.



Figure 1. The research design of the current study. Square frames show the instructions in the two drawing situations where there are different instructions in the first-round drawing but the same in the second round.

### 2.3 Data analysis

There were three different types of information we attempted to extract from the drawing data. The first index was how many types of technology items were displayed in the second-round drawing in each group, which was to evaluate if our treatment in the first-round drawing worked. Statistically, to evaluate the difference between the Technology group and the Experience group in terms of how many technology items were depicted in the second-round drawing, the Mann-Whitney *U* Test was utilized.

To echo the definition of technology-use belief (e.g., Ertmer et al., 2015) or technology knowledge from TPACK (Mishra & Koehler, 2006), how many participants depicted images of technology use in the second-round drawing is applied as the second index. The definition of technology use in the current study is that teachers or students were depicted by the participants as using technology or digital technology to teach or learn. For example, a teacher uses an interactive whiteboard to present physics equations in a class, or students wear AR devices to experience the learning content. If a participant expresses such images described above in the second drawing, then the code for this participant would be 1; if not, then the code would be 0 (see Fig. 2 for an example). Accordingly, a chi-square test of independence is performed to examine the relation between groups and numbers of participants depicting technology use in the second-round drawing.



Figure 2. Examples of whether characters in the second drawings used technology in the class. Two drawings by different participants are depicted. In the left-hand drawing, because a teacher is using an interactive whiteboard (IWB) to present a physics equation, the code of technology use in this drawing would be 1. On the contrary, although an IWB is shown in the right-hand drawing, the teacher and students are discussing the steps of experiments without using IWB, and nor is there any information on the IWB. The code of technology use in this drawing would therefore be 0.

In the third index, we examined, at the within-subject level, whether the pedagogical theme displayed in the second-round drawing resembled the first-round drawing by utilizing a chi-square test of independence. The coding rule is as follows. One participant, for example, depicted a teacher-centered pedagogical theme in the first-round drawing and depicted a student-centered pedagogical theme in the second-round drawing, then the code for this participant would be 0, which means non-match of pedagogical belief for this participant; on the other hand, if what a participant depicted in both the first- and second-drawing is, say, a student-centered pedagogical theme, then the code for this participant would be 1, meaning there was a match in pedagogical belief for this participant. The three indices were coded by the two independent coders as well. When inconsistency was found, the two coders with the senior coder would negotiate to reach a consensus.

### 3. Results

# 3.1 Is there a significant difference in the numbers of technology items depicted in the second-round drawings of the Technology group and the Experience group?

The test revealed significant differences in number of technology items depicted by the Technology group (Median = 1, n = 22) and the Experience group (Median = 0, n = 20), U = 293.000, z = 2.048, p = .041, and r = .316, meaning that the effect size of the result was medium. The result indicated that the Technology group depicted more types of technology items in the second-round drawing than the Experience group.

# 3.2 Is there a significant difference in the number of participants in the Technology group and the Experience group who depicted technology-use images in the second-round drawing?

As for technology use, the result of a  $X^2$  test indicated that the relation between groups and numbers of the participants who depicted technology-use images in the second-round drawing was significant, Pearson's  $X^2$  (1, N = 42) = 5.301, p < .05 (p = .021), meaning that more participants in the Technology group depicted technology-use images than those in the Experience group. The cross tabulation is shown in Table 1.

	Technology	No Technology	Total
	use	use	
Tech group	12	10	22
Exp group	4	16	20
Total	16	26	42

Table 1. Cross tabulation of the relation between groups and numbers of technology use displayed in the second-round drawings.

3.3 Is there a significant difference in the number of participants in the Technology group and the Experience group who depicted similar pedagogical themes in both the first-round and second-round drawings?

As for pedagogical beliefs, the result of a  $X^2$  test indicated that the relation between groups and numbers of participants who depicted similarity of pedagogical theme in the first- and second-round drawing was not significant, Pearson's  $X^2$  (1, N = 42) = 0.105, p > .05. In other words, the similarity of pedagogical belief in the first- and second-round drawings did not differ by groups. The cross tabulation is shown in Table 2.

Table 2. Cross tabulation of the relation between groups and similarity of pedagogical theme in the first- and second-round drawings.

	Congruent	Incongruent	Total
Tech group	11	11	22
Exp group	9	11	20
Total	20	22	42

### 4. Discussion and Implications

In the next section, three issue are discussed, the methodological, theoretical, and practical implications.

### 4.1 Methodological Implications

As the 3.1 result shows, when their technology knowledge was activated by the instruction of the first-round drawing, the Technology group would depict more technology items in the second-round drawing than the Experience group whose technology knowledge was not activated, which means that, in the Experience group, the instruction of the second-round drawing would frame the participants' thoughts; in the Technology group, however, the instruction of the first-round drawing help the participants disrupt this framing effect.

Hence, when utilizing a drawing task to investigate preservice teachers' technologyuse beliefs, researchers should make the content of drawing instructions pertinent to research questions and interpret the results of a drawing task with caution since instructions of a drawing task would frame subjects' thoughts. If the content of drawing instruction must be neutral because of the research needed, making positive interpretations is recommended; yet if negative interpretations are necessary, it is suggested to use the tworound drawing task to decrease possible confounding. There is an exemplar in the Technology group (see Fig. 3). The participant did not depict technology-use images in their second-round drawing even when their technology knowledge had been activated in the first-round drawing. In this scenario, we could be more confident in inferring negatively that this participant thinks that they need not use technology or digital technology to assist their classes since we had activated their technology knowledge in the first-round drawing.



Figure 2. Exemplar drawings of the Technology group. These two drawings were produced by the same participant. The participant depicted what s/he thinks about technology or digital technology integrated into a science course in their first-round drawing and depicted what s/he will do in their science courses in their second-round drawing. This participant depicted technology-use images in the first-round drawing but not in the second-round drawing. We can confidently infer that s/he did not think that it is necessary to use technology in the science course.

### 4.2 Theoretical Implications

As noted in the Introduction, Skott (2015) argued that teachers would use their educational beliefs to deal with educational issues or information they encountered, like a knowledge filter. Similarly, Richardson (2003) clarified the definitions of beliefs and knowledge by proposing that beliefs are thought to be subjectively true for the individual, whereas knowledge is not. For example, educational philosophy of pedagogy, like traditional lecturing and student-centered constructivism are usually demonstrated in educational textbooks. For novice preservice student teachers, the pedagogical philosophy is a kind of knowledge, but when preference is involved, if the individual prefers the pedagogy of constructivism to traditional pedagogy, it becomes a belief for them. By applying the belief-as-a-filter theory to explain the result in 3.2, it looked as if the preservice teachers whose technology knowledge was activated had let the activated knowledge flow into their second-round drawing compared to those whose technology knowledge was not activated. It seemed that what digital technology assisted in teaching and learning means to the preservice teachers is a kind of knowledge not beliefs.

On the other hand, it is reasonable to expect the 3.3 result that pedagogical belief displayed in both the first- and second- round drawings would be more consistent in the Experience group, in which the participants' pedagogical memory had been activated than the Technology group, as with the result in 3.2. It was, however, not the case. No more participants in the Experience group depicted similar pedagogical scenes in both the first- and second-round drawing than those in the Technology group. It seemed that the Experience group would not totally accept the pedagogical information or memory activated by the instruction of the first-round drawing. That is, they might depict the images about pedagogy based on their educational beliefs regardless of the activated pedagogical memory. Thus, what images of pedagogy depicted by participants in the Experience group reflected would be their beliefs about pedagogy.

### 4.3 Practical Implications

Taiwan Ministry of Education (2023) has been cultivating the project of digital technology integrated into teaching and learning for about thirty years. The result of the current study echoes the project and would provide a way to promote in practice. In some traditional classrooms, there are less, even no, digital-technology devices for a teacher and students to use, which can be seen as a framing effect of a classroom. Similarly, Ertmer (1999) also argued that lack of access digital technology is an external barrier for teachers to integrate digital technology unto their classroom. As it can be seen in the current study, the participants whose technology-use knowledge were activated would depict more technology-use images than those who were not. It implied that we could and should establish a sound digital-technology-use environment for teachers and students. As noted in Wood (2019), preference is for what people are used to seeing, which also called mere exposure effect. As long as teachers and students are getting familiar to digital technology embedded in their classroom, they would be inclined to utilize it to teach and learn.

### 5. Conclusion and Future Study

To the best of our knowledge, our study is one of the pioneers in utilizing a two-round drawing task along with a random-assigned experiment design to investigate beliefs or knowledge about technology-assisted teaching and learning. One of the advantages of this research design is that it can provide empirical data to refine the theory of educational beliefs, which might still be the subject of philosophical debates (e.g., Richardson, 2003). Work on two-round drawing tasks has, however, just begun; further cultivation is required. For instance, when it comes to beliefs about technology-assisted teaching and learning, and learning, it would be interesting to make a comparison of preservice and in-service teachers to uncover if there are indeed different levels of these beliefs since in-service teachers may build their beliefs about technology-assisted teaching and learning through teaching experience, which preservice teachers do not yet have.

### Acknowledgements

The funding of this study is supported by Natural Science and Technology Council, Taiwan, under grant contract number MOST 111-2410-H-017 -001 -MY3

### References

- Alkış Küçükaydın, M., & Gökbulut, Y. (2020). Beliefs of teacher candidates toward science teaching. Journal of Science Teacher Education, 31(2), 134-150. https://doi.org/10.1080/1046560X.2019.1673603
- Ashton, P. T. (2015). Historical overview and theoretical perspectives of research on teachers' beliefs. In H. Fives & M. G. Gill (Eds.), *International handbook of research on teacher beliefs*, (pp. 31-47). Routledge.
- Chai, C. S. (2010). The relationships between Singaporean pre-service teachers ICT competencies, pedagogical beliefs, and their beliefs on the espoused use of ICT. *Asia-Pacific Education Researcher, 19*, 387-400. <u>http://hdl.handle.net/10497/4791</u>
- Chang, H.-Y., Lin, T.-J., Lee, M.-H., Lee, S. W.-Y., Lin, T.-C., Tan, A.-L., and Tsai, C.-C. (2020). A systematic review of trends and findings in research employing drawing\_assessment in science education. Studies in Science Education, 56(1), 77-110. https://doi.org/10.1080/03057267.2020.1735822
- Elmas, R., Demirdöğen, B., & Geban, Ö. (2011). Pre-service chemistry teachers' images about science teaching in their future classrooms. *Hacettepe University Journal of Education, 40*, 164-175.

- Ertmer, P. A. (1999). Addressing first- and second-order barriers to change: Strategies for technology integration. *Educational Technology Research and Development*, *4*7(4), 47–61. https://doi.org/10.1007/BF02299597
- Ertmer, P. A., Ottenbreit-Leftwich, A., Sadik, O., Sendurur, E., & Sendurur, P. (2012). Teacher beliefs and technology integration practices: A critical relationship. *Computers & Education, 59*, 423-435. <u>https://doi.org/10.1016/j.compedu.2012.02.001</u>
- Ertmer, P. A., Ottenbreit-Leftwich, A. T., & Tondeur, J. (2015). Teachers' beliefs and uses of technology to support 2firstcentury teaching and learning. In H. Fives & M. G. Gill (Eds.), *International handbook of research on teacher beliefs*, (pp. 403-418). Routledge.
- Funkhauser, B.J. & Mauser, C. (2013). Drawing on technology: An investigation of preservice teacher beliefs in the context of an introductory educational technology course, *Computers and Education*, 62, 271-285. <u>https://doi.org/10.1016/j.compedu.2012.11.005</u>.
- Hermans, R., Tondeur, J., van Braak, J., & Valcke, M. (2008). The impact of primary school teachers educational beliefs on classroom use of computers. *Computers & Education, 51*, 1499-1509. https://doi.org/10.1016/j.compedu.2008.02.001
- Hew, K. F., & Brush, T. (2007). Integrating technology into K-12 teaching and learning: Current knowledge gaps and recommendations for future research. *Educational Technology Research* and Development, 55(3), 223–252. <u>https://doi.org/10.1007/s11423-006-9022-5</u>.
- Inaltekin, T. (2020). Examining secondary students' perceptions of the technology-based learning and teaching in science courses. *World Journal on Educational Technology: Current Issues, 12*(2), 071–083. <u>https://doi.org/10.18844/wjet.v12i2.4628</u>
- Kahneman, D. (2012). Thinking, fast and slow. Penguin Group.
- Lin, Y. C. (2022). What are you doing in the mathematics classroom? Using a drawing task to investigate pre-service teachers' beliefs about mathematics teaching and learning. *Education Journal, 50*(1), 31-62.
- Lortie, D. (1975). Schoolteacher: A sociological study. Chicago: University of Chicago Press.
- Minogue, J. (2010). What is the teacher doing? What are the students doing? An application of the draw-a-science-teacher-test. *Journal of Science Teacher Education, 21*(7), 767-781. https://doi.org/10.1007/s10972-009-9170-7
- Minor, L. C., Onwuegbuzie, A. J., Witcher, A. E., & James, T. L. (2002). Preservice teachers' educational beliefs and their perceptions of characteristics of effective teachers. *The Journal of Educational Research*, 96(2), 116-127. <u>http://www.jstor.org/stable/27542420</u>
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers college record, 108*(6), 1017-1054.
- Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula (Ed.), *Handbook of research on teacher education* (2nd ed., pp. 102-119). Macmillan.
- Richardson, V. (2003). Preservice teachers' beliefs. In J. Rath & A. C. McAninch (Eds.), *Advances in Teacher Education series, 6* (pp. 1-22). Information Age Publishing.
- Schraw, G. & Olafson, L. (2015). Assessing teachers' beliefs: Challenges and solutions. In H. Fives & M. G. Gill (Eds.), *International handbook of research on teacher beliefs*, (pp. 13-30). Routledge.
- Skott, J. (2015). The promises, problems, and prospects of research on teachers' beliefs. In H. Fives & M. G. Gill (Eds.), *International handbook of research on teacher beliefs*, (pp. 13-30). Routledge.
- Taiwan Ministry of Education (2023, March 30). Main educational policies: Digital teaching and learning. The History of Ministry of Education. <u>https://history.moe.gov.tw/Policy/Detail/1e540341-4d3b-44aa-bea7-17e55d6f26b5</u>
- Tatar, N. (2015). Pre-service teachers' beliefs about the image of a science teacher and science teaching. *Journal of Baltic Science Education*, *14*(1), 34–44.
- Wood, W. (2019). Good habits, bad habits: The science of making positive changes that stick. Macmillan.
- Yeh, H. Y., Tsai Y. H., Tasi, C. C., & Chang, H. Y. (2019). Investigating Students' Conceptions of Technology-Assisted Science Learning. *Journal of Science Education and Technology*, 28(4), 329-340. <u>https://www.jstor.org/stable/10.2307/48699217</u>
# Exploring the Relationship Between Students' Preferences for Teacher Authority and Learning Approaches: An Example of Student Learning Communication Theory and Computer Technology Contexts

#### Wen-Lung HUANG<sup>a\*</sup>

<sup>a</sup>Department of Communication, Fo Guang University, Taiwan \*wlhuang@mail.fgu.edu.tw

**Abstract:** This research investigated university students majoring in Mass Communication preferences regarding teacher authority within the context of communication theory and computer technology learning and examined how these preferences were associated with learning approaches. Data were collected from 357 Taiwanese university students using two questionnaires. The results identified three distinct clusters in terms of teacher authority: passive learners (Cluster 1), blended learners (Cluster 2), and autonomous learners (Cluster 3). ANOVA results revealed significant differences in learning approaches across the three clusters. Notably, blended learners exhibited higher scores in deep strategy compared to passive and autonomous learners, while both blended learners and autonomous learners demonstrated stronger deep motivation than passive learners. Finally, passive learners displayed the lowest level of learning approaches within the context of communication theory and computer technology learning.

**Keywords:** teacher authority, approaches to learning, communication, computer technology

### 1. Introduction

Most undergraduate students majoring in Mass Communication are required to take both theory and hands-on courses (e.g., computer technology and production technology), and instead find communication theory tedious, difficult to understand, distant and abstract, and of little relevance to their daily lives, leading to limited interest in such courses (Griffin, 2012; West et al., 2010). This may lead students to adopt a passive learning attitude where they use to memorization methods to cope with theory examinations. Therefore, how to implement effective teaching methods (e.g., student-centered, and teacher-centered) to stimulate students' learning motivation and willingness for autonomous learning has always been a concern among educators. However, these teaching methods may vary significantly based on students' characteristics, and even differ according to their academic majors (Akdemir & Özçelik, 2019; Murphy et al., 2020). Therefore, it is worthwhile for mass communication educators to understand the varying levels of preference among mass communication students for both student-centered and teacher-centered pedagogies.

Teacher-centered refers to the approach where the teacher takes responsibility for imparting a significant amount of knowledge to students through a lecture-based mode, and in the teaching environment, the teacher is active while the students are passive. Contrastingly, student-centered instruction creates a learning environment where skills and understanding are actively constructed, with the teacher assuming the role of a guide and facilitator, while

students actively participate in the learning process (Mascolo, 2009; Vavrus et al., 2011; Serin, 2018). In terms of student learning roles, based on previous research shown that students' preference for student-centered pedagogy involves two aspects: autonomy and participation management, while preference for teacher-centered pedagogy comprises two aspects: dependency and teacher control (Lee et al., 2009; Lin et al., 2013).

In the student-centered pedagogy, the two factors of autonomy and participation management mean that learners tend to have autonomy in the learning process and actively participate in making decisions about course content and assessments. Nevertheless, in the teacher-centered pedagogy, the two factors of dependency and teacher control imply that students expect teachers to decide the content knowledge that needs to be taught and to decide on the instructional activities and methods for evaluating learning (Alam, 2023; Lin et al., 2013). Additionally, the student-centered approach is commonly linked to the development of self-regulated and deep learning, as well as the enhancement of critical thinking and seeking understand. On the other hand, the teacher-centered approach is frequently associated with rote memorization and passive learning (Loyens et al., 2008; Alam, 2023).

Based on the above rationale, pedagogy of student-centered versus teacher-centered may have a profound impact on how students learn. Therefore, this study aims to investigate the relationship between students' preference profile of teacher authority and their approach to learning in the context of communication theory and computer technology.

#### 2. Literature Review

#### 2.1 Learner-centered versus teacher-centered learning environment

In the context of mass communication learning, teachers may control not only course content (e.g., theory course) and assessment, but also the learning activities during the technology (e.g., computer technology) production procedures. Particularly in the field of journalism and mass communication education, a great deal of teaching focuses on technology skills curricula (Han & Newell, 2014; Mendenhall, 2017), such as the computer graphics, computer editing techniques and virtual production technologies. In these activities, teachers' authority may play an important role. Therefore, as we emphasize learner-centered and teacher-centered pedagogies in mass communication learning environments, it seems necessary to focus on the issue of teacher authority.

In student-centered learning, students' activities are important indicators in learning process and quality of learning product (Zohrabi et al., 2012). This pedagogical approach integrates flexible learning, experiential learning, and self-directed learning to create a student-centered classroom environment where teachers act as facilitators, considering individual needs and encouraging active participation in the learning process (Acat & Dönmez, 2009). Additionally, in a teacher-centered learning environment, educators assume a significant role in the learning process, emphasizing the use of well-prepared learning materials and objectives. This approach enables teaching students in incremental and manageable steps to achieve specific learning outcomes. It is often considered suitable for large class sizes and facilitates shorter classroom activities. It may lead to less student interaction and limited opportunities for critical thinking and innovation in teaching (Emaliana, 2017).

Hence, considering the above reasons, it is essential to investigate the differences between teacher-centered and student-centered pedagogies and ascertain students' preferences for these approaches. This enables educators to make informed decisions about course delivery, classroom management, enhancing student engagement, improving learning effectiveness, and facilitating students in achieving desired learning outcomes.

#### 2.2 Preferences for teacher authority and approaches to learning

Several studies have revealed that students' learning attitudes are related to their preference for teacher authority (Lee et al., 2009). For example, Cornelius-White (2007) study indicated that students' preference for autonomy and participative management in student-centered

learning is strongly associated with their motivation and engagement in the learning process. On the other hand, in teacher-centered learning, student preference for dependency shows that students positively impact knowledge acquisition and reduce learner pressure and anxiety (Scheja, 2006). However, the teacher-centered approach is often associated with rote learning, memorization, and standardized testing, while student-centered learning prioritizes active student involvement in the learning process (Alam, 2023). Previous studies have revealed that adopting memorization or struggling with fragmented learning knowledge is considered surface approaches to learning in academic activities (Biggs, 1978; Marton & Säljö, 1976). In contrast, seeking to understand the learning content thoroughly and applying critical thinking to reflect on its meaning are driven by internal motivations and involve more sophisticated strategies, which are considered deep approaches to learning (Asikainen & Gijbels, 2017).

However, whether students' preferences profile of student-centered or teacher-centered instruction are related to their approaches to learning has rarely been explored in previous research. As a result, the current study seeks to investigate the potential correlation between students' preference profile of teacher authority and their approaches to learning in the context of communication theory and computer technology learning. With that in mind, this study had two main aims:

- 1. To identify the preference profile of Taiwanese university students for teacher authority within the context of learning communication theory and computer technology.
- 2. To investigate the relationship between preferences for teacher authority and learning approaches in the context of communication theory and computer technology learning among university students in Taiwan.

# 3. Methodology

#### 3.1 Participants

Participants in this study included 357 undergraduate students from five universities in northern, central, and southern Taiwan. All of them majored in the Department of Mass Communication. All participants completed a Chinese version of the teacher authority and approaches to learning questionnaires.

#### 3.2 Instruments

To explore students' perspectives on teacher authority and their approaches to learning communication theory and computer technology, this study employed two questionnaires. First, the Teacher Authority (TA) questionnaire was modified from Lee, Chang, and Tsai's (2009) instrument, which has been shown to be a reliable measure of college students' perceptions of teacher authority in the context of earth science. The TA questionnaire includes four factors that assess teacher authority, namely autonomy, participatory management, dependency, and teacher control. Each factor which has detail description are presented below:

- 1. *Preference for autonomy assesses* the extent to which students value autonomy and the freedom to decide the content knowledge they need to acquire in the context of classrooms.
- 2. *Preference for participatory management* assesses the degree to which students actively engage in decision-making about learning activities and the creation of learning content and grading criteria.
- 3. *Preference for dependency* assesses the extent to which students favor relying on instructors for decision-making and knowledge transmission within the context of communication and computer technology learning.
- 4. *Preference for teacher control* evaluates students' perceptions of the extent to which teachers exert control over content and classroom activities.

The second questionnaire was the Approaches to Learning through Mass Communication (ALMC) questionnaire, modified from Lee et al.'s (2008) Approaches to

Learning through Science questionnaire., which has been shown to be a reliable measure of college students' approaches to learning in the context of science education. It included two dimensions: the deep approach, encompassing deep motives and deep strategies, and the surface approach, consisting of surface motives and surface strategies. Four factors are described below:

- 1. *Deep motive* refers *to* students' intrinsic motivation when participating in the context of communication theory and computer technology learning, characterized by their curiosity and personal interests.
- 2. *Deep strategy* refers *to* students utilizing more meaningful approaches to facilitate their learning communication theory and computer technology, thereby ensuring a thorough understanding of the subject matter.
- 3. *Surface motive* refers *to* students' extrinsic motivation when participating in the context of communication theory and computer technology learning, encompassing goals such as achieving high grades and fulfilling parental expectations.
- 4. Surface strategy refers to students adopting a limited scope of learning objectives and employing rote methods to acquire knowledge in the context of communication theory and computer technology learning.

Two expert mass communication educators were invited to assess the content validity of the revised items in TA and ALMC questionnaire. Each item on the two questionnaire was presented using a five-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree).

### 3.3 Data analysis

Three statistical analyses were employed in this study. Firstly, the study uses exploratory factor analysis to select the items which fit for the scale. Secondly, performing a cluster analysis on all participants' responses to the TA questionnaire. Subsequently, the clusters initially generated by the Ward's method were further examined using the k-means method. Finally, the study employs an analysis of variance (ANOVA) test to differentiate distinct boundaries among diverse teacher authority preference profiles, which are subsequently utilized to examine their impact on mass communication learning approaches.

### 4. Results and Discussion

### 4.1 Exploratory factor analysis of the TA questionnaires

The results by applying the exploratory factor analysis with varimax rotation revealed four factors with a total of 19 items of the students' preference of teacher authority in the context of communication and computer technology learning. These four factors explained variances are 61.29%. The Cronbach's  $\alpha$  coefficients for four factors were 0.79, 0.84, 0.77, 0.87, and 0.88, respectively, and the overall alpha was 0.82, suggesting that these factors have high reliability in assessing students' preference profiles about teacher authority.

#### *4.2 Exploratory factor analysis of the ALMC questionnaires*

The results by applying the exploratory factor analysis with varimax rotation to assess the construct validity of the ALMC questionnaires and extract the existing factors in these research tools. These four factors with a total of 21 items explained variances are 60.56%. The Cronbach's  $\alpha$  coefficients for five factors were 0.87, 0.87, 0.82, and 0.78, respectively and the overall alpha was 0.84, suggesting that these factors are sufficiently reliable for representing learning approaches in the context of communication theory and computer technology learning.

4.3 Identifying the mass communication students' preference profile of teacher

#### authority among Taiwanese university students.

A cluster analysis approach was used to characterize the different preferences of mass communication learners with respect to their perceptions of teacher authority. The results of the cluster analysis show that students' preferences for teacher authority can be categorized into three different clusters, namely Cluster 1 (Passive learners), Cluster 2 (Blended learners) and Cluster 3 (Autonomous learners). Each profile is characterized as follows:

*Passive learners.* The participants in this cluster scored the lowest on the dimensions of autonomy (mean 3.17, standard deviation 0.25) and dependence (mean 3.39, standard deviation 0.41) in teacher authority. This implies that students in this cluster do not possess a strong inclination to engage actively with the content of classroom learning, nor do they exhibit a strong tendency to rely on the knowledge imparted by the teacher. Hence, students within this cluster could potentially be characterized as "passive learners" who may not exhibit a significant concern about the locus of authority.

*Blended learners.* Blended learners' members exhibited the highest average scores in the dimensions of autonomy (mean 4.09, standard deviation 0.40) and dependence (mean 4.25, standard deviation 0.34) within teacher authority. This suggests that students in this cluster tend to prefer engaging in classroom learning content, and they also show a strong preference for relying on the teacher to impart knowledge. Hence, students within this cluster could potentially integrate both learner-centered and teacher-centered orientations, classified as "blended learners".

Autonomous learners. Students in autonomous learners demonstrated a higher inclination towards autonomy (mean 3.90, standard deviation 0.27) and a lower inclination towards dependence (mean 3.50, standard deviation 0.34) in their preference for teacher authority. This shows that this group of students favors a student-centered pedagogical approach, displaying a pronounced inclination to actively participate in classroom content. As a result, they can be classified as "autonomous learners".

# 4.4 The relationship between students' preference profile of teacher authority and approaches to learning

An analysis of variance (ANOVA) was conducted to determine the relationship between students' approaches to learning and different preferences for teacher authority. The results showed that there were significant differences in learning approaches among the three clusters, including deep motivation (F = 39.61, p < 0.000), deep strategies (F = 47.65, p < 0.000) 0.000), surface motivation (F = 18.27, p < 0.000), and surface strategies (F = 26.44, p < 0.000). Subsequent post hoc tests were conducted using the Scheffé method to determine the significant differences among the groups. The results revealed that students in blended learners have statistically higher scores in deep strategy than those in passive learners and autonomous learners. The result implies that students in blended learners utilize deep strategies, such as comprehension and relating ideas, to form a more comprehensive view. Furthermore, students in blended learners and autonomous learners exhibited statistically higher scores in terms of deep motivation compared to those in passive learners. This observation suggests that both clusters of students possess a profound motivation, such as intrinsic interest, driving their engagement in the study of communication theory and computer technology. Besides, students in blended learners statistically scored higher on surface motivation and surface strategy than the other groups.

However, blended learners also exhibited higher scores in both deep approach and surface approach compared to the other groups. That is, students in blended learners tended to have both intrinsic and extrinsic motivations to learn communication theory and computer technology, but simultaneously, they may hold deep and surface strategies to learn communication theory and computer technology. This phenomenon suggests that mass communication students might be directed by instructors to develop critical thinking skills (Morreale et al., 2000), while also being shaped by Taiwan's exam-oriented culture, resulting in a blend of deep and surface learning strategies. Furthermore, the learning patterns of Taiwanese students are molded by the sociocultural context, wherein parental focus on

academic scores remains paramount. This influence has the potential to lead students to adopt surface motivations to engage in mass communication research, especially during exam periods (Huang et al., 2018).

Finally, it is suggested that future research could further utilize higher-order statistical methods for comprehensive inquiry, such as using structural equation modeling to explore the relationship between teacher authority and learning approaches in the context of communication theory and computer technology learning.

#### References

- Acat, B., & Dönmez, İ. (2009). To compare student centred education and teacher centred education in primary science and technology lesson in terms of learning environments. *Procedia-Social and Behavioral Sciences*, *1*(1), 1805-1809.
- Alam, M. A. (2023). From teacher-centered to student-centered learning: the role of constructivism and connectivism in pedagogical transformation. *Journal of education*, *11*(2).
- Akdemir, E., & Özçelik, C. (2019). The Investigation of the Attitudes of Teachers towards Using Student Centered Teaching Methods and Techniques. *Online Submission*, 7(4), 1147-1153.
- Asikainen, H., & Gijbels, D. (2017). Do students develop towards more deep approaches to learning during studies? A systematic review on the development of students' deep and surface approaches to learning in higher education. *Educational Psychology Review*, 29, 205-234.
- Biggs, J. B. (1978). Individual and group differences in study processes. *British journal of educational psychology*, *48*(3), 266-279.
- Biggs, J. (1994). Approaches to learning: Nature and measurement of. *The international encyclopedia of education*, *1*, 319-322.
- Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A metaanalysis. *Review of educational research*, 77(1), 113-143.
- Emaliana, I. (2017). Teacher-centered or student-centered learning approach to promote learning? *Jurnal Sosial Humaniora (JSH)*, *10*(2), 59-70.
- Griffin, E. (2012). A first look at communication theory (8th ed.). New York, NY: McGraw-Hill.
- Han, G., & Newell, J. (2014). Enhancing student learning in knowledge-based courses: Integrating team-based learning in mass communication theory classes. *Journalism & Mass Communication Educator*, 69(2), 180-196.
- Lee, M. H., Chang, C. Y., & Tsai, C. C. (2009). Exploring Taiwanese high school students' perceptions of and preferences for teacher authority in the earth science classroom with relation to their attitudes and achievement. *International Journal of Science Education*, *31*(13), 1811-1830.
- Lee, M. H., Johanson, R. E., & Tsai, C. C. (2008). Exploring Taiwanese high school students' conceptions of and approaches to learning science through a structural equation modeling analysis. *Science education*, *92*(2), 191-220.
- Lin, C.-L., Liang, J.-C., Su, Y.-C., & Tsai, C.-C. (2013). Exploring the relationships between selfefficacy and preference for teacher authority among computer science majors. *Journal of Educational Computing Research*, *49*(2), 189-207.
- Loyens, S. M., Rikers, R. M., & Schmidt, H. G. (2008). Relationships between students' conceptions of constructivist learning and their regulation and processing strategies. *Instructional Science*, *36*, 445-462.
- Huang, W.-L., Liang, J.-C., & Tsai, C.-C. (2018). Exploring the relationship between university students' conceptions of and approaches to learning mass communication in Taiwan. *The Asia-Pacific Education Researcher*, *27*, 43-54.
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—Outcome and process. *British journal of educational psychology*, *46*(1), 4-11.
- Mascolo, M. F. (2009). Beyond student-centered and teacher-centered pedagogy: Teaching and learning as guided participation. *Pedagogy and the human sciences*, *1*(1), 3-27.
- Mendenhall, D. L. (2017). The JMC introductory classroom with and without Miley Cyrus: An experiment in undergraduate media research. *Teaching Journalism & Mass Communication*, 7(2), 32-38.
- Morreale, S. P., Osborn, M. M., & Pearson, J. C. (2000). Why communication is important: A rationale for the centrality of the study of communication. *JACA-ANNANDALE-*, *1*, 1–25.

- Murphy, L., Eduljee, N. B., Croteau, K., & Parkman, S. (2020). Relationship between Personality Type and Preferred Teaching Methods for Undergraduate College Students. *International Journal of Research in Education and Science*, 6(1), 100-109.
- Vavrus, F., Thomas, M., & Bartlett, L. (2011). Ensuring quality by attending to inquiry: Learnercentered pedagogy in sub-Saharan Africa. UNESCO-IICBA Addis Ababa, Éthiopie.
- Scheja, M. (2006). Delayed understanding and staying in phase: Students' perceptions of their study situation. *Higher Education*, *52*, 421-445.
- Serin, H. (2018). A comparison of teacher-centered and student-centered approaches in educational settings. *International Journal of Social Sciences & Educational Studies*, *5*(1), 164-167.
- Stevens, J. (1996). Applied multivariate statistics for the social science (3rd ed.). Mahwah, NJ: Erlbaum.
- West, R. L., Turner, L. H., & Zhao, G. (2010). Introducing communication theory: Analysis and application (Vol. 2). McGraw-Hill New York, NY.
- Zohrabi, M., Torabi, M. A., & Baybourdiani, P. (2012). Teacher-centered and/or student-centered learning: English language in Iran. *English language and literature studies*, *2*(3), 18.

# Learning Analysis infrastructure to support optimal learning based on IRT and Video Enrollment Data

Masahiro SAKAMOTO, Hiroshima City University Graduate School of Information Sciences and Technology, Japan

Yukihiro MATSUBARA, Hiroshima City University Graduate School of Information Sciences and Technology, Japan

Kousuke MOURI, Hiroshima City University Graduate School of Information Sciences and Technology, Japan

di65006@e.hiroshima-cu.ac.jp

Abstract: Currently, local governments in Japan need DX (=Digital Transformation) promotion personnel. However, even today, Japan's civil service system still employs mainly legal and human personnel, resulting in a need for more necessary human resources. Therefore, the demand for DX human resource development for local government officials is rapidly increasing. In response to this, we have been engaged in human resource development using e-learning materials based on Moodle for three years since 2020, targeting approximately 9,500 employees of the Hiroshima City Office, where I work, and through formative evaluation of the materials, we have succeeded in creating materials at a level where all employees who took the course expressed a desire to promote DX. The company succeeded in developing educational materials so that all staff members who took the course were willing to promote DX. However, only about 5% of all staff members completed the course, reaching the limit of the improvement method in formative evaluation from those who completed the course. During the three-year effort, we also identified a wide gap in the IT skills of local officials and that we cannot provide optimal educational materials for 95% of the staff. In this study, we created multiple-choice questions corresponding to the DX literacy standard set by the Ministry of Economy, Trade and Industry (METI) in December 2022, which we regarded as a complete mastery map, and optimized them using IRT. The objective was to develop a model for measuring the DX literacy of local government officials and an e-learning course that provides individualized and optimal learning based on the measurement results. Through these efforts, we would like to establish a method of correlating skill surveys using IRT (=Item Response Theory) with e-learning course attendance data to overcome the limitations of the technique of creating e-learning courses and improving them through formative evaluation and to present a method to overcome the limitations of formative assessment in the ID (=Instructional Design) field. We hope to offer a strategy to overcome the limitations of formative evaluation in ID.

**Keywords:** DX, e-learning, IRT, Local government, DX literacy standard, Video Analysis, Skill Survey

### 1. Introduction

#### 1.1 LMS Infrastructure

An e-learning course for Hiroshima City employees was created in the environment shown in Figure 1, with construction starting in October 2019 and course offerings beginning in March 2020. Furthermore, based on the load data from the two courses initially provided, SSL (=Secure Socket Layer) support and Vimeo (= Vimeo is an online platform for sharing and watching videos. There most significant difference between Vimeo and YouTube is that

creators pay a usage fee. There is no advertising.) plan changes were made. As a result, five classes were delivered to Hiroshima City employees for three years until March 2023, functioning without a single failure as an e-learning infrastructure for 20,000 viewers, 3,000 students, and 800 completed courses; this study allowed us to verify the availability of the LMS information infrastructure construction in terms of required CPU resources, for example.





R stammidsenti 4.4 (4)	· - •	A 18212 # #		0 X
1 C Q B particular contraction and the U		+ -> O Q # undulta outer	e open innen die 'n ook in jaar van op plagt af aan gete 10000 taan de 100	0 0 <b>0 0 0 0 0 0 0</b>
B OSLEDS-ING DER-	· · ··································	<ul> <li>DALE007-107</li> </ul>	en a cara a nater a natera a naterar - estrutt. A tre entre a nova a servician. I par	A - 1374 TR
testing     testing	■ 自治体 消滅	OSCARD-20     OSCARD-20	And	0751457-585

Figure 2. e-Learning Courses.

# 1.2 Results of e-learning Course offerings

Based on about 500 survey responses from participants who completed the course and formative feedback from interviews with about 20 participants, we spent three years improving the course materials to create a system that is "easy for participants to understand." (Table 1).

Table 1. Responses obtained in formative evaluation

	FY2021	FY2022
Fairly inaccessible	8%	2%
A little difficult to approach	25%	6%
It was no different from any other material.	21%	8%
It was relatively easy to approach.	30%	44%
I found it very approachable.	16%	40%

As an example of improvement based on formative evaluation, we worked on improvements such as carefully explaining IT terms in the videos.

In FY2020, as shown in Table 2, videos longer than 14 minutes tended to increase the number of students who left the course. In addition, in the formative evaluation by participants, some said that one class was too long for them to study during their workday, which interfered with their work.

Year	Course Name	Impression	View	Completion	Cmpletion ratio	Video time	Average	Unwatched
	Information security training (Users Section) [Lecture 1]	4025	404	238	58.91%	0:22:31	0:16:59	0:05:32
	Information security training (Users Section) [Lecture 2]	3998	291	177	60.82%	0:22:58	0:18:17	0:04:41
	Information security training (Users Section) [Lecture 3]	3995	227	160	70.48%	0:14:11	0:12:24	0:01:47
	Information security training (Users Section) [Lecture 4]	3997	248	173	69.76%	0:01:51	0:01:34	0:00:17
EV2020	Information security training (Users Section) [Lecture 5]	3998	229	124	54.15%	0:27:53	0:22:46	0:05:07
F12020	Information security training (Managament section) [Lecture 1]	1309	183	110	60.11%	0:23:05	0:17:14	0:05:51
	Information security training (Managament section) [Lecture 2]	1307	144	69	47.92%	0:41:37	0:27:06	0:14:31
	Information security training (Managament section) [Lecture 3]	1282	97	64	65.98%	0:18:23	0:15:03	0:03:20
	Information security training (Managament section) [Lecture 4]	1280	88	49	55.68%	0:39:53	0:29:20	0:10:33
	Information security training (Managament section) [Lecture 5]	1276	81	50	61.73%	0:28:37	0:23:01	0:05:36

Using these results and the formative evaluation as a reference, the course was redesigned in FY2021 so that all videos would be approximately 10 minutes long. As a result, as shown in Table 3, unwatched time was significantly reduced, and the number of students who were likely to leave the course due to the length of the videos was reduced.

Table 3. Video viewing time (FY2021)

Year	Course Name	Impression	View	Completion	Cmpletion ratio	Video time	Average	Unwatched
	2022 teaser	6416	225	166	73.78%	0:00:17	0:00:13	0:00:04
	DX starting from Information Security (General) [Lecture 1]	775	417	280	67.15%	0:10:32	0:08:40	0:01:52
	DX starting from Information Security (General) [Lecture 2]	530	283	209	73.85%	0:11:27	0:10:14	0:01:13
	DX starting from Information Security (General) [Lecture 3]	475	268	194	72.39%	0:12:51	0:11:45	0:01:06
	DX starting from Information Security (General) [Lecture 4]	472	251	191	76.10%	0:10:32	0:09:29	0:01:03
EV2021	DX starting from Information Security (General) [Lecture 5]	435	250	157	62.80%	0:11:17	0:09:39	0:01:38
F12021	DX starting from Information Security (Management) [Lecture 1]	307	189	143	75.66%	0:10:16	0:08:51	0:00:20
	DX starting from Information Security (Management) [Lecture 2]	260	165	111	67.27%	0:09:38	0:08:17	0:01:25
	DX starting from Information Security (Management) [Lecture 3]	215	129	95	73.64%	0:08:33	0:07:37	0:01:21
	DX starting from Information Security (Management) [Lecture 4]	185	120	85	70.83%	0:13:01	0:11:10	0:00:56
	DX starting from Information Security (Management) [Lecture 5]	195	119	89	74.79%	0:09:25	0:08:18	0:01:51
	DX starting from Information Security (Management) [Lecture 6]	205	117	53	45.30%	0:13:03	0:11:36	0:01:07

By the way, the Information Policy Department of the City of Hiroshima provides consultation on DX promotion. As a result of the provision of e-learning courses and improvement of course content, the number of staff consulting on DX promotion has repeatedly doubled since the start of training compared to the previous year in an evaluation using Kirkpatrick's four-level evaluation method. The number of staff starting DX promotion, in other words, reaching Level 3, has also increased (Table 4).

Year	Number	Comment
2018	28	
2019	30	
2020	23	
2021	49	Results for 2020
2022	104	Results for 2021

Table 4. Number of Consultations

However, in this environment, it is not possible to use methods other than formative evaluation, although it is possible to grasp the students' quiz-taking status in each course, which can be obtained through Moodle, the results of questionnaires, and the viewing time of each lecture video provided on Vimeo.

This made it impossible to obtain data to address students who did not cooperate with the formative evaluation or who dropped out of the program, making further improvement unlikely (Table 1).

#### 1.3 Organization of conventional methods and Issues

The program's primary goal is to design an efficient learning experience for the participants based on the data by correlating test results with their attitudes toward the course.

The Hiroshima City employees targeted for the course were widely distributed between the ages of 18 and 65, including some who had graduated from information

technology universities and some who had worked in the private sector, and there were significant differences in their abilities and IT experience.

In the formative evaluation of a previous study, we analyzed the results from the perspective that "older employees are less comfortable with IT." Still, we were able to determine that there was a relationship between pre-training test scores and ease of use rather than age.

However, since the pre-test was merely a questionnaire about the language information domain set based on the course content, the validity of the test itself and the accompanying course evaluation could not be statistically processed.

Several private operators currently offer tools for conducting skills surveys that comply with the DX Literacy Standards. Still, they are "fixed question sets" and require about one hour to take the test. Furthermore, when we inquired with the providers about the content, they said they did not evaluate the validity of the questions. In this regard, by using IRT, we will create a test corresponding to the learning items listed in the DX Literacy Standards and evaluate its validity.

Furthermore, instead of having the respondents answer the entire massive database of questions, the test should be an adaptive test to measure their understanding of the DX Literacy Standards in a 15-minute test that can be administered between work hours. Finally, through these efforts, we hope to quantitatively measure the ability of "DX human resource utilization and development," Which is still in the process of exploring specific methods in local governments and support staff assignment according to their abilities, as well as realize efficient human resource development.

Regarding the attitude toward viewing video materials according to ability, past research has identified the "time of video with the largest number of people leaving" based on Vimeo viewing history data and designed educational materials with this time in mind. Still, we could not analyze which parts of the video were practical or whether there was a relationship between the trend of the video itself and the attitude toward the course. However, we could not analyze from the perspective of which parts of the videos were practical and whether there was a relationship between the videos' trends and the students' attitudes. Therefore, through this study, we would like to gain knowledge on the design of courses according to ability.

### 2. Related Studies

#### 2.1 DX Literacy Standards in Japan

The Digital Skill Standards (Ministry of Economy, Trade and Industry, December 21, 2022) were announced as the roles of personnel who promote DX and the skills they should acquire. It consists of two types of standards: the DX Literacy Standards, which are guidelines for all businesspeople to acquire basic knowledge, skills, and mindsets related to DX, and the DX Promotion Skill Standards, which are guidelines for companies to develop and hire personnel with expertise in promoting DX. However, the DX Literacy Standards do not mention local government officials.

For local government officials, the Ministry of Internal Affairs and Communications (MIC), which has jurisdiction over local governments, has prepared "Initiatives for Staff Development to Promote Municipal DX, June 3, 2022." Still, it remains vague guidelines for initiatives, does not refer to specific learning content such as DX literacy standards, and does not envision the creation of skill standards or the establishment of skill surveys targeting local government officials. Furthermore, it does not envisage the creation of skill standards or the establishment of a skill survey targeting local officials.

Therefore, a database of questions corresponding to the DX Literacy Standards will be created, and IRT will be used to evaluate the validity of the questions. However, since items such as "customer" and "competitive environment" cannot be applied as they are when

focusing on the work-relatedness of local officials, the DX Literacy Standards are redefined in terms of "citizen" and "social cooperation," respectively, and questions are created.



Figure 3 DX Literacy Standards

# 2.2 Creating a Skill Survey Environment Using IRT

Study items are presented in the DX Literacy Standards, but no examples of specific questions are given. Therefore, creating a surveyable question database for each area of DX Literacy is necessary. In Japan, the Information Technology Engineer Examination is a public examination that can be used as a reference for creating questions. The IT Passport Examination is the least difficult of these examinations. It covers so-called IT literacy, but the syllabus for the IT Passport Examination needs to cover the items required by the DX Literacy Standard.

Therefore, it is necessary to create questions that satisfy the items of the DX Literacy Standards while referring to the questions of the IT Passport Examination. The appropriateness of the questions in terms of content and difficulty will be analyzed using IRT, and inappropriate questions will be modified or eliminated. This test edition will be done by asking about 30 prior testers who have cooperated in formative evaluations in previous studies to answer the entire question database.

Once the appropriateness of the problem difficulty level has been confirmed to some extent, a large-scale skill survey will be conducted for all Hiroshima City employees. In this case, adaptive IRT will be employed.

For example, surveying a database of 200 questions by having the examinees answer all of them would require approximately 2 hours, and many examinees would "leave" the test. Therefore, to measure ability in 10 to 15 minutes, the test environment should be constructed as an adaptive test of the 2PL model, taking into consideration that "the question database contains enough questions with a wide variety of difficulty levels" (Yoshiko Tokusada, Takenori Sakumura, Naoki Tabuchi, Hideo Hirose, 2019). Therefore, the test environment is constructed as an adaptive test of the 2PL model.

As a follow-up study to this research, we envision the development of a DX skills information dashboard that measures the distribution of DX skills among employees at the municipal level through a skill survey of local government employees nationwide and supports the planning of DX human resource allocation and training according to the results of the study. Therefore, Microsoft Azure SQL Database was selected as the database, considering the ability to utilize Azure Machine Learning as a data analysis tool (Figure 4).



Figure 4 Azure Machine Learning Model

# 3. Perspectives on Analysis of Video Materials

#### 3.1 Infrastructure to be constructed in this project

Moodle will be used for learning courses. However, the learning infrastructure will be improved from the following perspectives to realize more diverse data analysis. About the Study Course Moodle will continue to be used for learning courses.

Separation of test and course of the study as an extension module of Moodle, a previous study to realize IRT is "Development of Plug-in for Moodle: Adaptive Test Module based on Potential Rank Theory." (Hiroshi AKIYAMA,2012) However, with this method, it is necessary to respond appropriately to upgrades of the

However, with this method, it is necessary to respond appropriately to upgrades of the PHP used in Moodle and to promotions of Moodle itself. Furthermore, since there are plans to develop subsequent applications, such as aggregating the results of the skill survey for each local government, it was decided to create an entirely new system for the IRT test environment.

The video playback status should be able to be analyzed for each student. In addition, Azure Media Services will be used to analyze the videos to explore the relationship between their characteristics and the students' attitudes.

#### 3.2 Focus on the video audience.

Have them take a skill survey and then take a course.

In doing so, we will focus on the relationship between the skills of each item measured in the skill survey and the data on the attendance at the educational materials created for each item, and whether there is a relationship between prior skills in each domain of the DX Literacy Standards and the attitude toward viewing the video materials corresponding to each field. For example, we will identify events such as skipping video playback in areas where prior skills are high and repeating video playback many times in areas where skills are low, as well as viewing attitudes such as the tendency for participants with higher overall skills to play videos faster.

#### 3.3 Focus on the video itself.

Using Azure Video Indexer, the video material itself is analyzed. This function allows us to infer, for example, that specific keywords are repeatedly viewed, in which case it can be concluded that the explanation of those keywords is challenging to understand. Furthermore, if the speaker leaves after a certain period of silence, it can be inferred that there is a silent period during which the learner is bored. In this case, there is a quiet period during which the participants are bored. Table 4 shows the analysis items of Azure Video Indexer, which is planned to be used in this project.

Table 4. Azure Video Indexer Analysis Items

statistics		labels	
Name	Description	Name	Description
CorrespondenceCount	Number of correspondences in the video.	id	The label ID.
WordCount	The number of words per speaker.	name	The label name (for example, 'Computer', 'TV').
SpeakerNumberOfFragments	The amount of fragments the speaker has in a video.	language	The label name language (when translated). BCP-47
SpeakerLongestMonolog	The speaker's longest monolog. If the speaker has silences		A list of time ranges where this label appeared (a label
	inside the monolog it's included. Silence at the beginning	instances	can appear multiple times). Each instance has a
	and the end of the monolog is removed.		confidence field.
SpeakerTalkToListenRatio	The calculation is based on the time spent on the speaker's monolog (without the silence in between) divided by the total time of the video. The time is rounded to the third decimal point		

#### 3.4 Their relevance

By combining video playback data per viewer unit with the analysis results of Azure Video Indexer, we can, for example, extract keywords that are difficult for students with low scores in specific areas in the DX Literacy Standards, relate playback speed to ability before taking the course, and confirm the effectiveness of improving teaching materials based on these data. The data will be used to verify the effectiveness of enhancing teaching materials.

Heer	Level Attitu		Attitudo	KouWord		
User	Topic	1	2	3	Attitude	Reyword
Α	AI				Very Fast,Skip	-
В	AI				Quit	_
С	AI				Slow	Deep learning,Neural network
Α	Network				Fast	subnet mask
В	Network				Normal	subnet mask,IPv6
С	Network				Slow	subnet mask,IPv6,DNS
Α	Security				Very Fast	-
В	Security				Slow	Availability,AES,Public-key cryptography
С	Security				Fast	Availability,AES

Figure 5 Expected Results

#### 4. Future Prospects

#### 4.1 Standardization of DX Literacy Skill Survey in Local Governments

The provision of skill surveys to measure DX literacy standards needs to be improved, partly because it is still early days since the digital skills standards were published.

Furthermore, since this study is the only one that reads the DX literacy standard regarding work relevance for local governments, it is expected to serve as a skill survey standard for local government employees.

#### 4.2 Improvement of teaching materials

Continuously improve teaching materials by analyzing data to understand the responses of "voiceless participants," which could not be ascertained through formative evaluation.

#### 4.3 Creating an Individualized Optimization Environment

First, efforts will be made to improve the completeness of the skill survey and e-learning materials. Indicators include: the results of the IRT analysis of the skill survey question database should be appropriate; the quality of the video materials for each of the DX Literacy Standards should exceed 90% as "easy to catch on"; and the quality of the video materials should exceed 100% as "easy to learn." A 100% pass rate on the post-learning test within three attempts. Achieve these conditions.

The goal is to construct a system that can propose the optimal order of participation for participants to master all DX literacy standards by measuring changes in their willingness to learn when they take the courses in the order of their strengths and weaknesses in the skill survey results, and the difference in the number of days it takes them to complete each course item before starting the next.

	TOPIC	Survey	Course		Survey	Course		Survey
	Social changes	0	Y-1		_	Y-1		_
Why	Changes in customer value	0	Y-2		_	Y-2		-
	Changes in the competitive environment	×	► Y-3		0	Y-3		-
	Data in society	0	T-1		_	T-1		-
	Reading and explaining data	0	T-2		_	T-2		-
	Dealing with data	0	T-3		-	T-3		-
What	Determined by data.	×	► T-4		. 0	T-4		-
VVIIat	AI	×	► T-5		0	T-5		-
	Cloud	×	► T-6		. 0	T-6		-
	Hardware and software	×	► T-7	•	• • • •	T-7		-
	Networks	× –	► T-8		. ×	 T-8	 -	• • •
	Examples of the use of data and digital technologies	0	W-1		-	W-1		-
	Tool use	×	▶ W-2		0	W-2		-
HOW	Security	×	► W-3	•	· 0	W-3		_
	Morale	×	► W-4	•	• • • •	W-4		-
	Compliance	×	► W-5	+ +	• • • •	W-5		-

\_\_\_\_

Discover the best sequence of courses to take.

Figure 6 Optimization through skill surveys

#### Acknowledgments

Hiroshima City Official supports this study.

#### References

- Minoru AKIYAMA. (2012) Development of plug-ins for Moodle -Adaptive testing module based on latent rank theory.
- Mitsuko TOKUSADA., Takenori SAKUMURA., Naoki TABUCHI., Hideo HIROSE. (2015). Considerations on the Optimal Number of Questions in Adaptive IRT Testing.
- Marcella Milana., Sue Webb., John Holford., Richard Waller., Peter Jarvis. (2017) The Palgrave International Handbook on Adult and Lifelong Education and Learning
- Makoto Takahashi. (2021) Introduction to IRT Item Response Theory: How to make a quality test from the basics of statistics.
- Kentaro Kato., Takeshi Yamada., Kazumitsu Kawabata. (2023) Item Response Theory with R
- Yuhiko Mitsunaga., Akiko Nishida. (2022) What are tests for? Considering the Admissions System from Item Response Theory.

Microsoft Corporation. (2023). Analyze video and audio files with Azure Media Services. https://learn.microsoft.com/en-us/azure/media-services/latest/analyze-video-audio-files-concept

- Microsoft Corporation. (2023). Dynamic packaging in Media Services v3. https://learn.microsoft.com/en-us/azure/media-services/latest/encode-dynamic-packaging-concep
- Ministry of Economy, Trade and Industry. (2022). DX literacy standard.
- Ministry of Internal Affairs and Communications. (2022). Initiatives for Staff Development to Promote Municipal DX.
- Hiroaki Ogata., Etsuhiro Eguchi. (2023) Learning Analytics to Transform Learning
- Reza Rad. (2023) Pro Power BI Architecture: Development, Deployment, Sharing, and Security for Microsoft Power BI Solutions
- Louis Davidson. (2023) Practical Graph Structures in SQL Server and Azure SQL: Enabling Deeper Insights Using Highly Connected Data
- Chris Seferlis., Christopher Nellis. (2023) Practical Guide to Azure Cognitive Services: Leverage the power of Azure OpenAI to optimize operations, reduce costs, and deliver cutting-edge AI solutions Maxwell Everstone. (2023) Azure Data Factory: A Comprehensive Guide
- Mervin Weiden. (2023) Mastering SQL Server: Streamlining SQL Server Always on Availability Group Deployment In Azure

# Measuring Understanding in Video-Based Learning

#### Song-Yi LIN<sup>1</sup>, Meilun SHIH<sup>3</sup>, and Hsin-Mu TSAI<sup>\*12</sup>

 <sup>1</sup> Department of Computer Science and Information Engineering <sup>2</sup>Digital Learning Center, Office of Academic Affairs
 <sup>3</sup>Center for Teaching and Learning Development, Office of Academic Affairs National Taiwan University, Taipei, Taiwan 106319 \*hsinmu@ntu.edu.tw

**Abstract:** Measuring students' mental states, such as their understanding during class, helps improve learning efficiency. Automatic approaches implement this idea without interrupting the class by sensing students' reactions through wearable sensors or cameras and applying machine learning models to analyze the data. However, most of the previous works lack adequate annotations of understanding based on students' reactions compared to the number of concepts conveyed during lessons. This paper proposes a scalable framework for efficiently constructing and annotating datasets. Additionally, we have collected a dataset consisting of posture, facial expression, and eye movement features, and benchmarked it for measuring understanding. The results show promising accuracy of 80% even in cases where not all features are available, demonstrating the potential for widespread adoption of the proposed framework.

**Keywords:** Video-based learning, measurement of understanding, machine learning

### 1. Introduction

Formative assessments provide frequent feedback on the learner through close observations (Black & Wiliam, 1998). They can be leveraged to identify the weaknesses of a learner or pinpoint the concepts that require further clarification, bridging the gaps between the teaching and the learning and improving the quality of both. Although effective assessment techniques have been developed after years of research and experimentation (Angelo & Cross, 1993), they often require significant effort and time to implement, such as human labor to make observations and analyses, preventing comprehensive and large-scale studies. In this paper, we aim to develop automation tools to perform one of such assessments - the measurement of understanding of a learner, and to break the constraint in a specific learning environment video-based learning (VBL).

VBL has become an important learning environment as many higher education institutions developed their Massive Open Online Courses (MOOCs) and Open Coursewares (OCWs) in the past decade. Moreover, the pandemic of COVID-19 has also forced many classes in higher education that originally took place in the physical classroom to move online. One common practice is to have the lectures pre-recorded and placed in video streaming platforms, and the learner sits in front of a computing device and watches the video to learn. Compared to the physical classroom (Ahuja et al., 2019; Gao, Shao, Rahaman, & Salim, 2020), in such a VBL setting, the reactions of the learner, e.g., the facial expression (Friesen & Ekman, 1978), the body posture, the gaze (Robal, Zhao, Lofi, & Hauff, 2018a; Zhao, Lofi, & Hauff, 2017), and the control input to the video player, can be easily recorded using the webcam of the device. If these reactions can be analyzed automatically, they can be used to understand the learning process collectively in a large-scale setting or specific to an individual learner.

This work intends to design and implement a system to measure *learners' understanding* based on their reactions. This is different from many past works that focus on other metrics, such as engagement, confusion, and attention level. We believe that resolving confusion, i.e., understanding, leads to engagement and prevents frustrations in cognitive processes (D'Mello, Lehman, Pekrun, & Graesser, 2014). It is therefore useful to directly estimate the learner's level of understanding.

In this paper, we took the following steps to realize this idea. First, we design a scalable framework that allows multiple devices to record a learner's reaction in video-based learning environments. This framework should synchronize all recordings with the lecture video and simplify the annotation process for convenience. Next, we constructed a data set with the proposed framework, which includes subject recruitment. We have decided to collect eye gaze, facial expressions, posture, and the annotation of understanding when watching lecture videos. We established specific guidelines for annotators to follow and to annotate understanding. Finally, different machine learning-based model architectures are used against our dataset to examine their effectiveness in measuring understanding in real-world situations. We tested the models using different feature sets, and determine to what extent these models perform under certain real-world limitations. These findings could demonstrate the potential for extending our system to ubiquitous laptops and highlight the system's effectiveness.

#### 2. Related Work

In order to generate informative feedback from the learners, a variety of targets have been focused on, including their attention, engagement, and various aspects of interest levels. The purpose is to detect undesired behavior to prompt student's self-regulation and reflective learning, to provide beneficial insights for educators to create the right learning climate for students (Gao et al., 2020), and sometimes to predict learners' performance as references for both learners and instructors to enhance learning and teaching efficiency (Ramakrishnan, Zylich, Ottmar, LoCasale-Crouch, & Whitehill, 2021).

However, reaching a consensus on a clear definition for measuring mental states and acquiring ground truth can be challenging, even when the underlying concepts are understood. For example, different methods are applied to detect and measure attention loss. Zhao et al. (2017) asked participants to press a key if they experienced mind-wandering, a form of attention loss, in the past 30 seconds, and they kept ringing a bell to remind them to report feedback. Robal et al. (2018a) implemented several alerting modules to raise the learner's attention when a loss of focus is detected, including pausing the video playback. Kar et al. (2020) use both test scores and human perception scores to measure attention. Each test contains five questions related to the content of the corresponding video, and the attention score measured this way is correlated with the degree of comprehension. Scores from human perception were manually evaluated by other participants, who took turns being observers of their peers.

Although these targets can be subjective and difficult to define, basic features are usually derived from clickstream data, data of the surrounding environment, background audio, eye tracking data, facial expressions, postures, and physiological activities such as the electrical activity of the brain, galvanic skin responses, and heart rates (Monkaresi et al., 2016; Gao et al., 2020). These responses from learners are often collected through cameras, eye trackers, wearable devices, or physically clickable buttons. Some works focus on a single feature, while others combine different features. For instance, Whitehill et al. (2014) draw attention to



Figure 1. System Design

classifying engagement based on facial information in image frames and videos, while Robal et al. (2018b) rely on facial and eye gaze information to track learners' attention. The choice of which features to use usually depends on the specific scenarios of applying the system and the accessibility of the data.

Among these features, our work will mainly focus on the combination of eye movement data, facial expressions, and postures. They are likely available in video-based learning environments nowadays. Most works we refer to were conducted in video-based learning environments, where learners watch videos through mobile devices or computers (Whitehill et al., 2014; Monkaresi et al., 2016; Zhao et al., 2017; Robal et al., 2018a; Kar et al., 2020). In this case, learners are naturally in front of built-in cameras on the device, which minimizes the impact on their reaction. On the other hand, some works that benefit from non-invasive sensing devices and take place in instrumented physical classrooms (Gao et al., 2020; Ramakrishnan et al., 2021). Although the learning environments may not be identical, we value the inspiring ideas and methodologies for constructing a pipeline that automatically collects responses through devices, estimates features of interest based on the recorded data, and analyzes targets representing learning efficiency.

There also exist publicly available datasets, including datasets annotated with affective states such as DAiSEE and Engagement in the Wild, or CLASS-coded datasets like Measures of Effective Teaching (Pianta et al., 2008; Kane et al., 2013). DAiSEE is a dataset created by Gupta et al. (2016) that contains video snippets annotated with levels of affective states as a multi-dimensional score. In these cases, the annotation process is time and labor-consuming, highlighting the value of an efficient annotation tool, which this work aims to develop.

#### 3. System Design

The system developed in this work consists of two parts, as shown in Fig. 1.

The first part involves a toolkit to collect a dataset capturing the reactions of a learner while watching online lecture videos, annotated with the understanding level of the learner. The objective is to create a dataset with fine granularity, multiple features, and accurate annotation. These are essential for creating a machine learning model for estimating the understanding level in the later stage. Most existing publicly available data sets do not satisfy these requirements. For example, some only specify the understanding level for a large section of the video



Figure 2. Interface of the Annotation Tool

or the entire video. With better granularity, the proposed system can pinpoint a video location where the learner starts to fall behind.

The second part is to apply existing machine learning models to the constructed dataset and examine their performance in measuring understanding. We consider this a supervised classification task. The models being compared will include neural-network models and conventional machine learning models, as well as time-series and non-time-series models.

### 3.1 Dataset creation toolkit and data collection

To construct the dataset, we recruited graduate and undergraduate students majoring in computer science as participants to emulate learners taking online courses. We recorded their reactions while they watched lecture videos and asked them to annotate their understanding.

### 3.1.1 Video selections

Four lectures are selected from two different machine learning courses, covering high-level ideas of algorithms and mathematical proofs. The topics covered in these lectures are self-contained to provide participants with sufficient background knowledge, eliminating the need for them to take any preliminary courses in advance. Moreover, the videos are manually split into a number of clips, each covering a complete idea. This is similar to the setting commonly seen in MOOCs. The duration of the video clips ranges between 5 and 12 minutes. The lecture videos use slides in English but the lectures are conducted in Mandarin Chinese, the native language of the participants.

#### 3.1.2 Dataset creation toolkit

We developed two tools to facilitate the data collection process. The first functions as a video player controlled by keyboards, while simultaneously recording reactions. The learner can seek forward or backward by 5 seconds, pause, and resume playback. The playback speed is adjustable, ranging from 0.7x to 2.0x with increments of 0.25x. All keyboard events are logged, representing changes in watching behavior.

In addition, we use the following sensors to collect data at the same time:

1. **Eye tracker.** We use Tobii eye tracker 5 to determine the gaze points at 33 Hz. Each sample represents a coordinate on a plane that corresponds to the screen.

- 2. Webcam. The webcam captures a learner's facial expressions and partial upper body poses while watching lecture video clips. Compared to eye trackers, webcams are more accessible since most laptops have built-in webcams. The recording quality and sampling rate in our setup are at least 1280×720 pixels at 30 frames per second, which most built-in laptop webcams can collect.
- 3. **Microphone.** We collected the audio because the participants might leverage it to recall their understanding during the annotation phase.

The second provides a computer program to facilitate the annotation process. This program displays the video to be annotated and contains a progress bar allowing users to jump onto a specific position. An understanding gauge indicating the label applied to each video frame is aligned with the progress bar. The user can select a section of the video and assign a single label to all frames within the section. Fig. 2 illustrates the graphical user interface of our annotation program. The label is a boolean value, where 0 represents confusion and 1 represents the opposite, indicating fully understood at the moment.

### 3.2 Participants and conducting data collection

A total of 13 students are invited as participants, all of which are male. Out of the 13 students, there are two first-year, three second-year, and one third-year undergraduate students, and four first-year, three second-year master graduate students. The participants have experience taking online courses but have not taken any courses covering the subjects in our selected lecture videos. Before the experiment starts, details of how to calibrate the eye tracker, how to use the programs to record and annotate, etc., are explained.

The participants utilized our tool to watch lecture video clips in a specific order. Each participant watched two out of the four lectures. During each video playback, the participants were asked to remain focused, indicating that they are not allowed to leave their seats or fall asleep. The collection process for each video clip consists of two phases: the recording phase and then the annotating phase. Only after completing both phases will a participant move on to the next clip. We believe the short time span between the two phases and self-reported annotations would result in higher quality and reliability of the labels of our dataset.

The collected dataset includes reactions collected from 13 participants who watched a total of 4 lectures. It contains 748,198 samples with annotations, out of which 224,920 of them are annotated as confused, accounting for approximately 30% of the dataset. The considered reactions include eye movements, facial expressions, body and head poses, and watching behaviors. Each data sample will be represented by a tuple of these elements, annotated with a label indicating the understanding level at that moment. Screen recording is also collected as an informative reference during annotation, although it will not be included in the dataset.

### 3.3 Machine learning model to estimate understanding level

#### 3.3.1 Derived features

The following features are derived from the collected raw data and serve as input to the machine learning model to estimate the understanding level (referred as the estimation model hereafter).

 Eye movement. Past works (Salvucci & Goldberg, 2000) provided us foundation of what parameters could contain useful information for measuring understanding. We use raw gaze points, their velocities, and the accelerations as the features for the estimation model. Note that these can either be obtained from the specialized eye tracker device or estimated from the webcam images.

- 2. Posture. Body poses and head poses are noticeable signs for human observers to perceive the understanding of others. Psychological experiments also support the claim that bodily states affect cognition and cognitive states can be expressed through body language (Glenberg, Havas, Becker, & Rinck, 2005). We rely on OpenPose to extract posture features from webcam recordings. OpenPose is a real-time multi-person keypoint detection library that estimates human faces and bodies frame by frame (Cao, Hidalgo, Simon, Wei, & Sheikh, 2021). It provides 70 key points for the human face and 25 key points for the human body. For the estimation model, the edges between key points are used as features, resulting in 67 edges for the face and 24 edges for the body.
- 3. Facial expression. Here we wish to derive a vector representing the facial expression captured by the webcam images. The main idea is to use pre-trained models for image classification to extract facial features. We use Resnet-50 with pre-trained weights for this task as it is frequently used as a feature extractor for facial expression recognition tasks (Wang, Peng, Yang, Meng, & Qiao, 2020). The end result is a 1000-dimensional vector representing the facial expression.

#### 3.3.2 Model architectures

We selected three commonly used machine learning model architectures for the study: Long Short-Term Memory (LSTM) (Gers, Schmidhuber, & Cummins, 2000), self-attention (Vaswani et al., 2017), and Support Vector Machine (SVM). The main objective is to understand how accurate the models trained with our dataset can estimate the understanding level and compare their accuracy in different settings. LSTM and self-attention models are both neural network-based models, both take a sequence of instances as input and make use of the relations between instances. On the other hand, SVM takes one instance as input and treats consecutive instances as separated data. We adopted the L2-regularized L2-loss classification solver in LIBLINEAR, which is an implementation of linear SVM classifiers (Fan, Chang, Hsieh, Wang, & Lin, 2008). Linear classifiers can perform as well as kernel classfiers when the dataset size is relatively large compared to the dimension of the dataset, which is the case for our dataset. Moreover, SVM is often used as a baseline model for comparison with more advanced architecture.

### 4. Evaluation Results

#### 4.1 Dataset Partitioning

To evaluate the performance, we divide our dataset into a training set and a testing set. The models are trained with data in the training set, and the evaluation results are produced with the data testing set. On the other hand, we employ several partitioning methods according to different scenarios. The first method is partitioning the participants, ensuring that the training set and the testing set do not contain data from the same participant. The second method is partitioning the viewed lectures so that the training set and the testing set do not share data from the same lecture video. The last method combines participant and lecture partitioning, resulting in the training set and the testing set having different participants watching different lecture videos. If not specified, the first method is set as the default partitioning method in the following sections. Fig. 3 illustrated the three methods. Once the partitions are established, we apply a 3-fold cross-validation on the training set to make full use of the limited dataset. The training set is divided into three subsets, and each subset takes turns to validate models trained on the remaining subsets. Finally, we test the performance of all three models, using the testing set, and calculate their arithmetic mean as the final result of accuracy.



Figure 3. Dataset Partition Methods



Figure 4. Overall Performace

Figure 5. Feature Set Selection

# 4.2 Results

Fig. 4 shows the overall accuracy with the error bar showing the standard deviation on the testing set using different partition methods. As expected, all models perform better when training and testing sets share common participants or lecture videos, indicating the presence of detectable patterns in the features. In real-world scenarios, these results imply the usefulness of the proposed system: if the model is trained with data including a particular participant or a particular course, then the model can be leveraged to produce more accurate results when the same participant watches the video clips from other courses, or when a different participant watches the same video clip. When both the participant and the video clip are different, the accuracy reduces by approximately 10%. The results also show that the latter case produces slightly higher accuracy, suggesting similar learning behavior patterns could be captured in the model for different learners watching the same video clip. Results also suggest that a self-attention model should be leveraged due to its superior overall performance and relatively short training time. SVM provides significantly shorter training time, trading off with slightly inferior accuracy.

Next, we aim to determine which feature contributes the most to measuring understanding. Here we test the model with data from different participants. Fig. 5 shows that for SVM and LSTM models, there is no significant accuracy loss when excluding the eye movement features. In most application scenarios, the devices do not have a specialized eye-tracking sensor; in such cases, the system can still achieve good accuracy, regardless of the model architecture applied. Moreover, as shown in Fig. 5, self-attention-based models can achieve 83% accuracy regardless of which features are included in the training set. This implies that the eye movement features contain sufficient information to estimate the understanding level with a strong model. This is useful as in cases where the eye tracking sensor is available, the participant might prefer its use since the webcam images could be considered as invading privacy.

#### 5. Conclusion

In this work, we proposed a scalable framework for measuring a learner's understanding of lecture videos. The framework covers the process of constructing datasets from learners' reactions, training machine learning models on the dataset and measuring understanding. We have also evaluated the feasibility of our framework in real-world situations.

Evaluation results suggest that the best performance was achieved when training the models using data from the same video clips with poses, facial expressions, and eye movement features, achieving 82% accuracy. The case when using data from the same learner but different video clips has a slightly lower accuracy of 78%. This shows that if the model can be trained with data from either the same learner or the same video clip, the model can make use of the behavior it learns from the data to better estimate the understanding level. Moreover, with the strongest self-attention-based model, results also suggest that the information contained in the eye movement, or the pose and facial expression features are both sufficient to achieve approximately 82% accuracy. Using only a webcam (for facial expressions and poses) can achieve large-scale deployment. Using only an eye tracker (for eye movement) could be preferred for privacy-sensitive scenarios. Results show that both of these solutions are feasible with good accuracy. We hope that the implications of these results demonstrate the feasibility of using automated tools to measure understanding in a large-scale and comprehensive manner, and the research community would produce similar open tools to facilitate related future research.

### Acknowledgements

This work is supported in part by National Science and Technology Council of Taiwan under grants 111-2221-E-002-177 and 112-2221-E-002-208.

#### References

- Ahuja, K., Kim, D., Xhakaj, F., Varga, V., Xie, A., Zhang, S., ... Agarwal, Y. (2019, Sep.). Edusense: Practical classroom sensing at scale. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(3), 1–26.
- Angelo, T. A., & Cross, K. P. (1993). Techniques for assessing course-related knowledge and skills. In *Classroom assessment techniques: A handbook for college teachers* (2nd ed., pp. 115–254). San Francisco, CA, United States: Jossey-Bass.
- Black, P., & Wiliam, D. (1998, Jul.). Assessment and classroom learning. Assessment in Education: principles, policy & practice, 5, 7–74.
- Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2021). Openpose: realtime multi-person 2d pose estimation using part affinity fields. *IEEE transactions on pattern analysis and machine intelligence*, 43(1), 172–186.
- D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, 29, 153–170.
- Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., & Lin, C.-J. (2008). Liblinear: A library for large linear classification. *The Journal of machine Learning research*, 9, 1871–1874.

- Friesen, W., & Ekman, P. (1978). Facial action coding system: a technique for the measurement of facial movement. Psychologists Press.
- Gao, N., Shao, W., Rahaman, M. S., & Salim, F. D. (2020, Sep.). n-gage: Predicting in-class emotional, behavioural and cognitive engagement in the wild. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *4*(3), 1–26.
- Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with lstm. *Neural computation*, *12*(10), 2451–2471.
- Glenberg, A. M., Havas, D., Becker, R., & Rinck, M. (2005). Grounding language in bodily states. *Grounding cognition: The role of perception and action in memory, language, and thinking*, 115–128.
- Gupta, A., D'Cunha, A., Awasthi, K., & Balasubramanian, V. (2016). Daisee: Towards user engagement recognition in the wild. *arXiv preprint arXiv:1609.01885*, 1–22.
- Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). Have we identified effective teachers? validating measures of effective teaching using random assignment. In *Research paper. met* project. bill & melinda gates foundation (pp. 1–48).
- Kar, P., Chattopadhyay, S., & Chakraborty, S. (2020). Gestatten: Estimation of user's attention in mobile MOOCs from eye gaze and gaze gesture tracking. *Proceedings of the ACM on Human-Computer Interaction*, 4(EICS), 1–32.
- Monkaresi, H., Bosch, N., Calvo, R. A., & D'Mello, S. K. (2016). Automated detection of engagement using video-based estimation of facial expressions and heart rate. *IEEE Transactions on Affective Computing*, 8(1), 15–28.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system™: Manual k*-3. Paul H Brookes Publishing.
- Ramakrishnan, A., Zylich, B., Ottmar, E., LoCasale-Crouch, J., & Whitehill, J. (2021). Toward automated classroom observation: Multimodal machine learning to estimate class positive climate and negative climate. *IEEE Transactions on Affective Computing*, 1–16.
- Robal, T., Zhao, Y., Lofi, C., & Hauff, C. (2018a). Intellieye: Enhancing MOOC learners' video watching experience through real-time attention tracking. In *Proceedings of the 29th on hypertext and social media* (pp. 106–114).
- Robal, T., Zhao, Y., Lofi, C., & Hauff, C. (2018b). Webcam-based attention tracking in online learning: A feasibility study. In *lui'18: 23rd international conference on intelligent user interfaces* (pp. 189– 197).
- Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on eye tracking research & applications* (pp. 71–78).
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*, 5998–6008.
- Wang, K., Peng, X., Yang, J., Meng, D., & Qiao, Y. (2020). Region attention networks for pose and occlusion robust facial expression recognition. *IEEE Transactions on Image Processing*, 29, 4057– 4069.
- Whitehill, J., Serpell, Z., Lin, Y.-C., Foster, A., & Movellan, J. R. (2014). The faces of engagement: Automatic recognition of student engagement from facial expressions. *IEEE Transactions on Affective Computing*, 5(1), 86–98.
- Zhao, Y., Lofi, C., & Hauff, C. (2017). Scalable mind-wandering detection for MOOCs: A webcam-based approach. In *Data driven approaches in digital education* (pp. 330–344). Springer.

# Research and Design of Digital Learning Resource Management System in Meteorological Adult Training Based on Cloud Computing

#### Jinfang HOU<sup>a\*</sup>

<sup>a</sup>China Meteorological Administration Training Centre, China <sup>\*</sup>happybrocade@126.com

**Abstract:** In this paper, we describe the design of digital learning resource management system based on cloud computing in meteorological adult training. This paper mainly introduces the overall structure and design strategy, function design, structure design and database design of the learning resource management system. It is to provide teachers and students convenient access to resources. The design concept of the system is unified function, cooperative development, centralized sharing, distributed storage. It realizes the integration of all kinds of multimedia data, unifies data construction specifications and classification standards, and provides standard interface management.

Keywords: Learning resource management system, cloud computing, system design

# 1. Introduction

Based on the design concept of unified functions, collaborative development, centralized sharing and distributed storage, Digital Learning Resource Management System based on Cloud Computing in Meteorological Adult Training (LMS-CC-MAT) integrates various multimedia data, unifies data construction norms and classification standards, and provides a standard interface management multimedia data construction environment on the basis of existing teaching platforms and teaching resources. It facilitates the branch and provincial training to actively participate in the construction of resources. It realizes the multimedia data construction environment with convenient data update, perfect system function and good application performance.

# 2. Function Design

The system takes media data digital storage, multimedia teaching resource management, search and recommendation resources as the core. It collects teaching materials from books, Internet, industries, work and life, and establishes a logical relationship with documents, images, audio, video, animation and other types of materials. The core functions of the system include uploading, cataloging, retrieval, storage management; download output, transcoding transmission, copyright management, in addition to user management, system management and other functions.

### 2.1 Upload module

The main function of upload module is to collect compressed or uncompressed digital file materials stored on traditional media and uploaded on the network, generate digital media data objects, and generate corresponding original cataloging information. With the function

of automatic key frame extraction, it provides the necessary basis for fast retrieval. The system can compress the material digitally with broadcast quality and convert it into streaming media, network video and other formats for data exchange.

# 2.2 Cataloging module

All kinds of learning resources and materials must be cataloged into the library. The cataloging system classifies the archived materials in a unified, scientific and standardized way, so as to turn the unordered information into orderly resources. This module provides structured processing and labeling of teaching materials, and supports the realization of shot segmentation, key frame extraction, content feature extraction, automatic extraction of labels and manual labeling labels, etc., which provides the basis for fast and efficient search of materials in comprehensive utilization.

Due to the wide range of online courseware resources, it is necessary to support a powerful cataloging system, which is suitable for large-scale cataloging production and management of pipeline, and is suitable for different cataloging needs of various types of resources such as high-definition audio and video, three-split screen streaming courseware, SCORM courseware, etc.

# 2.3 Query and retrieval module

Learning resources information query is the result of the catalog and data storage. The system supports a variety of compound retrieval methods, and the retrieval method based on media content is added to the traditional retrieval method. It supports to query according to various subject headings, keywords, speakers and other elements to achieve retrieval efficiency and accuracy of retrieval results. Based on the multimedia collection and cataloging of the material, the query terminal can not only use the original retrieval system to query the text and picture information of the material, but also browse the low bit stream material through the network.

Retrieval can be divided into the following categories, including full-text retrieval based on annotated content, retrieval based on visual content, retrieval based on examples, retrieval based on images, etc. The retrieval module should be independent and meet the combination of users' needs. At the client side, the structure of the whole system can be obtained through a retrieval interface. Data search lists can be automatically formed for the retrieved data. The system allows users to submit a download application, and the download application status is visible.

### 2.4 Catalog review module

The cataloging review module is to complete the final review of all cataloged and submitted teaching materials and learning resources from the state of "cataloging to be reviewed". Catalog review should contain different functions such as review, call back, and cancel review. All cataloged information, after being reviewed by professional cataloging and archiving staff, is formally stored in the archived learning resource database for easy retrieval, download and reuse.

### 2.5 Download module

LMS-CC-MAT can save the digital media data, then download or transfer to other systems to achieve browsing, viewing, editing and other needs. It supports fast positioning playback that is accurate to the frame, and supports downloading of program lists. It supports data conversion. It can convert digital files into DVD, streaming media and other formats, and makes intelligent code stream adjustment according to network conditions.

# 2.6 Transcoding module

The system fully supports the interconversion between various video and audio formats, and is committed to applying to mainstream video and audio formats to ensure the efficiency of format conversion, the success rate of conversion and the quality of transcoding. The system supports all kinds of video and audio equipment and collecting, editing and broadcasting systems for meteorological distance education, aiming to improve the efficiency and quality of video and audio production, management and broadcasting. This module can quickly transcode video and audio files of various formats, transcode courseware resources of various formats. It is compatible with all domestic non-editing software, courseware editing software, with the most flexible scalability and good usability and it is convenient for users to edit and use data.

### 2.7 Statistics and analysis module

Statistics and analysis module can timely understand more system information, including visitors, visit rate, teaching resource utilization rate and so on. It can understand the changing trend of teaching resources in real time, understand the scale of effective coverage of users and other important information. For example, through the statistical analysis of the target user characteristics of the use of teaching resources, it can provide an important basis for the collection and construction of teaching resources.

# 2.8 Statistics and analysis module

This module mainly includes parameter setting, log management, backup management, user management, permission management and security management. It has a three-level account management mode of "person-role-permission". It supports logging and query functions and supports automatic backup and recovery management of databases. The system sets secret levels for teaching resources, from low to high, and requires users with corresponding permissions to access operations to ensure the confidentiality of teaching resources. In addition, according to different users for storage space, departments, permissions, processes and other aspects, the system will carry out a comprehensive personalized settings and management.

# 2.9 Statistics and analysis module

The system provides a user account for each operator, and the user account is unique for each operator. Only after the user name and password have been verified, you can use the system. The system has detailed log records for each operation of each user and network administrator. The system administrator can add and delete personnel, departments and columns; establish the login password of the personnel, the setting of personnel permissions, system log management and so on.

# 3. Structure Design

LMS-CC-MAT is developed by B/S model. Based on the distributed network environment, this system constructs a platform to share multimedia resources and various materials for digital teaching and learning. This structure can improve the security of the system, effectively reduce the cost of system development and maintenance, and adapt to the changing needs of the business. It can effectively improve the concurrent processing ability of the system.

No specific software is required to be installed on the client, and the client can access the server simply by using an Internet browser. With a simple click of the mouse, you can access information in databases such as text, images, audio, video and animation by browsing the web. The system is relatively simple to maintain. There's almost no development work to do on the client side. However, most of the work is focused on the intermediate transaction logic layer and the back-end database layer. The three-layer B/S mode prevents users from waiting for the processing results when they query a large amount of data and cannot process other tasks, and also solves the time delay caused by the system processing complex logic. The advantages of using the three-layer structure include: the programmer can care about a certain layer of the overall system structure; the interdependence between layers is reduced; it is conducive to the standardization of modules, but also is conducive to the reuse of logic at each layer.



Figure 1. Three-layer application system architecture diagram.

# 4. Overall Architecture and Design Strategy

The overall architecture includes application layer, data logic layer and database layer.

### 4.1 Application layer

The application layer includes teaching resource management subsystem, teaching resource browsing and retrieval subsystem, teaching resource statistical analysis subsystem.

### 4.2 Data logic layer

The data logic layer establishes the information flow between each system and database, and the external data interface includes the teaching resource management system interface, the training file management system interface, and the Single Sign On (SSO) interface.

#### 4.3 Database layer

The following four logical databases will be built in the database layer: basic information database, courseware database, teaching case database and tutoring database.

LMS-CC-MAT adopts B/S model, based on open source J2EE architecture design, integration of Web2.0, Spring, Strusts, Hibernate, Ajax and other frameworks, to achieve advanced and complete enterprise application environment. The design concept of SOA is applied to realize data exchange through interfaces. The SSO mechanism is used to implement authentication and trust access. Based on WebLogic application middleware and Oracle RAC enterprise-class database cluster, the application system is stable. Link Load Balance and P2P technologies are used to ensure efficient network utilization. CDN technology is used to distribute and share data.

A network storage System based on the Internet Small Computer System Interface (ISCSI) technology provides massive storage. Digital Watermark technology is used to embed hidden marks in digital multimedia data by means of signal processing to achieve copyright protection of multimedia data. Using SVN (SubVersion) technology, the version management of multimedia data is realized.

In the program implementation, using the advantages of B/S structure design, we use JSP+Ajax+XML+Xhtml+css+div technology and Oracle database to achieve the data platform.

#### 5. Conclusion

In this study, we propose a cloud-based digital learning resource management system design model for online training, that allows learners and teachers to explore a variety of real open learning resource data. Online learning resource management system allows learners to explore the content they want to learn through other means than the content designed by the researcher. This actually shows that shows the learners' commitment to the exploration of the topic, and online guided query mode can increase students' learning motivation.

#### Acknowledgements

This research is funded by China Meteorological Administration Training Centre research project "Research of Digital Learning Resource Management System in Meteorological Adult Training Based on Cloud Computing" (2020CMATCZD06). We would like to thank all the people who prepared and revised previous versions of this article.

#### References

- Boulter, C. J., & Buckley, B. C. (2000). Constructing a typology of models for science education. *In Developing models in science education* (pp. 41-57). Springer, Dordrecht.
- Buck, L. B., Bretz, S. L., & Towns, M. H. (2008). Characterizing the level of inquiry in the undergraduate laboratory. *Journal of college science teaching*, 38(1), 52-58.
- Duit, R., Roth, W. M., Komorek, M., & Wilbers, J. (2001). Fostering conceptual change by analogies—between Scylla and Charybdis. *Learning and Instruction*, 11(4-5), 283-303.

Hakkarainen, K. (2003). Progressive inquiry in a computer-supported biology class. *Journal of research in science teaching*, 40(10), 1072-1088.

- Haury, D. L. (1993). Teaching science through inquiry. *Striving for excellence: The national education goals*, 2, 71-77.
- Looi, C. K. (1998). Interactive learning environments for promoting inquiry learning. *Journal of Educational Technology Systems*, 27(1), 3-22.

National Research Council. (1996). National science education standards. National Academies Press.

- National Research Council. (2000). Inquiry and the national science education standards: A guide for teaching and learning. National Academies Press.
- Pedaste, M., & Sarapuu, T. (2006). Developing an effective support system for inquiry learning in a web-based environment. *Journal of computer assisted learning*, 22(1), 47-62.
- Robottom, I. (2012). Socio-scientific issues in education: Innovative practices and contending epistemologies. *Research in science education*, 42(1), 95-107.

# W04

The 7th International Workshop on Information and Communication Technology for Disaster and Safety Education (ICTDSE)

# Collaborate, Design, and Generate Cybercrime Script Tabletop Exercises for Cybersecurity Education

#### Joshua DWIGHT<sup>a\*</sup>

<sup>a</sup>School of Science, Engineering, & Technology, Royal Melbourne Institute of Technology, Vietnam \*Joshua.dwight@rmit.edu.vn

**Abstract:** Cyber incidents are frequent, urgent, and sophisticated, and there is a global shortage of cybersecurity professionals. The demand for cybersecurity graduates is high, and a key priority for universities is improving successful graduate outcomes, employability, and work readiness. This research paper proposes a novel teaching approach to develop the skills and experience of higher education students for cybersecurity roles. The approach synthesizes problem-based learning, work-integrated learning, tabletop exercises, and crime script analysis. The paper presents a theoretical model based on the National Institute of Standards and Technology (NIST) Test, Training, and Exercise (TT&E) methodology and suggests using Generative AI for the development of draft documentation. The Cybercrime Script Tabletop Exercise can be integrated into a cybersecurity curriculum or as a stand-alone workshop. Future research can be conducted to compare, interview, and observe student outcomes such as skill and experience acquisition, work-life awareness, and levels of industry professional involvement.

**Keywords:** problem-based learning, work-integrated learning, tabletop exercise, crime script analysis, cybersecurity education, generative Al

### 1. Introduction

Cyber incidents are becoming more frequent, urgent, and sophisticated (Brilingaité et al., 2017; Zhang et al., 2021; Dwight, 2023), and there is a global shortage of 3.5 million skilled cybersecurity professionals (Angafor et al., 2020; Towhidi & Pridmore, 2023; Lunn et al., 2021). An international Fortinet research report specified that 68% of organizations face a shortage of cybersecurity skills (2023). Cybersecurity training is more important than ever (Chowdhury & Gkioulos, 2022). However, organizations do not have the necessary skill set to defend against sophisticated and continually evolving threats and cybercrimes (Angafor et al., 2020), and security training programs are often poorly developed and received (Reeves et al., 2021). To meet the labor demand, preparing higher education students for the cybersecurity workforce is a crucial priority for many universities, business organizations, and governments.

RQ: How can higher education academics develop and design tabletop exercises for students to develop skills and experience for cybersecurity roles?

This research paper aims to develop a novel teaching approach to develop the skills and experience of higher education students for cybersecurity roles. This paper will start with a literature review providing background and the gaps between cybersecurity education, learning and design paradigms, tabletop exercises, and crime script analysis. Then, this paper presents a theoretical model based on the National Institute of Standards and Technology (NIST) Test, Training, and Exercise (TT&E) methodology. Next, the theoretical model showcased the integrative model and utilized Generative AI for the development of sample

tabletop exercise documentation. Lastly, the paper wraps up with a discussion of the significant themes, limitations, future research, and implications.

# 2. Literature Review

### 2.1 Cybersecurity Education

Cyber incidents are prevalent (Brilingaité et al., 2017; Zhang et al., 2021), and in the industry, there is a global shortage of skilled cybersecurity professionals (Angafor et al., 2020; Towhidi & Pridmore, 2023). Cybersecurity training and education are more important than ever (Chowdhury & Gkioulos, 2022), as is the demand for cybersecurity graduates (McGettrick, 2013). An educated workforce is essential for developing trustworthy systems, but pedagogical teaching issues exist (Schneider, 2013). The education sector is often sluggish to transform and adapt to technologies (Lim et al., 2023). Improving successful graduate outcomes, employability, and work readiness has been a critical priority for many universities (Aprile & Knight, 2020; Kay et al., 2019). Developing cybersecurity education and curriculum can use both problem-solving and skill development types of active learning and design paradigms.

# 2.2 Problem-Based Learning (PBL)

Problem-based learning (PBL) is an educational approach whereby a problem is the starting point of the learning process (De Graaf & Kolmos, 2003). PBL typically involves eight to ten students working on a scenario with a facilitator to guide the students in a multi-step process (Bate et al., 2013). PBL helps students develop knowledge and skills to solve real problems (Hallinger, 2020). PBL has been successfully used in a variety of higher education disciplines for over 30 years (Clancey, 2020). The benefits of PBL include better industry performance, efficiency, competency, and interaction between students and faculty (Bate et al., 2013). PBL is the most comprehensive documented active learning method used in education today, although dominated by the medical education field (Hallinger, 2020). The emergence of the active learning school of thought has led to increased adoption and blending of PBL with other learning methods (Hallinger, 2020). However, there is very little literature blending PBL and WIL. This study proposes blending PBL with WIL to meet the demands of developing a cybersecurity industry workforce.

# 2.3 Work-Integrated Learning (WIL)

Work-integrated learning (WIL) is gaining popularity, where significant labor and skill shortages exist (Smith & Worsfold, 2015). WIL is an educational approach incorporating authentic industry practice where students learn and develop by working, networking, and engaging with industry professionals (Kay et al., 2019; Bilsland et al., 2019). The fundamental tenet of WIL is integrating university study and professional practice (Smith & Worsfold, 2015). Work-integrated learning traditionally entails short-term job placements or internships (Kay et al., 2019; Smith & Worsfold, 2015; Bilsland et al., 2019). However, emerging and innovative WIL practices such as co-designed programs (hackathons), active engagement in industry associations, event-based partnerships, use of technology (simulations), regional and global projects, practicums, supervised practice and greater flexibility in duration (brief placements or micro-placements) are gaining traction (Kay et al., 2019; Smith & Worsfold, 2015). Aprile and Knight (2020) note students experienced benefits such as obtaining real-world application of skills, learning through role models, and professional awareness of work life. WIL curriculum design is intended to integrate and apply theory with workplace practices (Smith & Worsfold, 2015). Tabletop exercises could be a helpful design avenue for a range of PBL and WIL opportunities.

### 2.4 Tabletop Exercises

When offering cybersecurity educational opportunities, tabletop exercises should be considered due to their ability to increase participants' engagement and success in skill acquisition (Chowdhury & Gkioulos, 2022). Tabletop exercises (TTX) are staged events where participants meet in an open forum to discuss actions for a response to a specific real-life scenario or incident (Brilingaite et al., 2017; Angafor et al., 2020). TTXs are based on communication and information knowledge sharing that may involve multiple roles within government, business, and other organizations (Brilingaite et al., 2017).

Tabletop exercises can include technical and non-technical activities (Brilingaitė et al., 2017; Angafor et al., 2020). These activities can consist of planning, discussion, and improvement of cyber plans and procedures. Reviewing and improving the skills and abilities of cyber incident response teams. Practicing and improving incident management coordination, prioritization, escalation, communication, and reporting strategies. Reviewing and practicing operational response capabilities such as tools and applications. Practicing and building awareness with different roles and responsibilities during an incident (Angafor et al., 2020). Tabletop exercises can provide several benefits. TTXs offer the acquisition of practical experience, nurture both technical and soft skills, provide opportunities for reflective practice, and are cost-effective to develop and implement (Angafor et al., 2020). Tabletop exercises can be designed in different ways.

Chowdhury and Gkioulos (2022) findings concluded that the involvement of industry professionals should include ease of implementation, time and resource constraints, remote accessibility, and shorter duration. Lunn et al. (2021) performed a study integrating tabletop exercises into the cybersecurity curriculum. The findings specified a significant increase in students' confidence in exploiting software vulnerabilities and implementing network security protocols. Chowdhury and Gkioulos (2022) conducted several exercises with different participants from industry and higher education students. Lunn et al. (2021) and (Chowdhury & Gkioulos, 2022) alluded to work-integrated learning but did not focus on the WIL theory and its connection to cybersecurity education and tabletop exercises. This paper proposes using crime script analysis as the basis for the tabletop design.

### 2.5 Crime Script Analysis

Introduced by Cornish (1994), crime scripts are models that describe predictable and sequential criminal actions, locations, and roles (Dehghanniri & Borrion, 2021). Crime script analysis (CSA) is an investigation profiling method that breaks down the actions of a criminal into stages to understand the behaviors, feelings, and decisions associated with the offense (Bada & Nurse, 2021). Cornish (1994) established seven stages that included preparation, entry, pre-conditions, instrumentals, doing, post-conditions, and exit. Later academic works have created similar truncated versions of the initial four stages, such as preparation, pre-activity, activity, and post-activity, that incorporate the other elements (Bada & Nurse, 2021; Bodker et al., 2022; Leppänen et al., 2020). Then the CSA model identifies areas of disruption by appropriate stakeholders (Cornish, 1994). This model provides a great outline to identify cybercriminal behaviors and attacks and allows cybersecurity professionals and students to recognize where and when cybersecurity measures can be used to mitigate cyber offenses.

Clancey's teaching pedagogy was influenced by problem-based learning and the integration of CSA into that PBL teaching pedagogy (Clancey, 2020). Clancey noted this pedagogy encouraged deeper thinking about motivations and root causes beyond crime risk factors and opportunities for criminal offenses (2020). See Table 1 for a summary of the paradigms and benefits. However, there was minimal literature applying and connecting these models.

Each of these models exhibits elements of active learning, although not expressed explicitly in the academic literature. Active learning is the activities designed to engage students in their learning with discussion, problem-solving, and or learning from each other (Nguyen et al., 2021).

Paradigm	Summary	Learning Benefits
Problem-based learning	Industry problem scenario	Industry preparation
(PBL)	Multi-step/procedural	Skill and knowledge
	Active learning	competency
Work-integrated learning	Integration of university study	Practical skill acquisition
(WIL)	and industry practice	and application
	Active learning	Work-life awareness
		Industry preparation
Tabletop exercise (TTX)	Industry problem scenario	Practical skill acquisition
	Multi-step/procedural	Reflective practice
	Active learning	Cost-effective
Crime script analysis	Industry problem scenario	Profiling cybercriminals
(CSA)	Multi-step/procedural	Skill and knowledge
	Active learning	competency

Table 1. Summary of paradigms and benefits

# 2.6 Literature Gaps

Even though active learning has seen positive results, STEM courses have low adoption rates and still predominately focus on lecturing (Nguyen et al., 2021). So it is imperative to try out different active learning techniques to nurture active learning. There are two literature gaps between PBL, WIL, tabletop exercises, and CSA. First, crime script analysis is used primarily as an investigative, analytical method and has not been applied as a mode of learning. Thus, it has not been applied in the cyber security context with PBL, WIL, or tabletop exercises. Secondly, there has been very little application between WIL and tabletop exercises. Lunn et al. (2021) and Chowdhury and Gkioulos (2022) alluded to using industry professionals for different aspects of tabletop exercise development but did not explicitly identify and apply the WIL approach. This paper aims to create a novel approach integrating these areas to better prepare and develop university students for professional cybersecurity roles to fill the skills gap.

# 3. Methodology

The National Institute of Standards and Technology (NIST) Test, Training, and Exercise (TT&E) methodology was used to help design a tabletop exercise integrated with PBL, WIL, and CSA for this paper. The NIST tabletop exercise methodology consists of four phases. The phases are design, development, conduct, and evaluation (Grance et al., 2006). The design phase establishes the topics, scope, objectives, participants, and logistics of the exercise. The development phase is the formation of all documentation and the establishment of tools necessary to conduct the exercise. These can be materials such as manuals, evaluation criteria, simulations, and gamification software. The conduct phase performs the exercise based on the design and development. The evaluation phase captures the lessons learned, discussions, and reflections from the exercise (Grance et al., 2006).

# 4. Cybercrime Script Tabletop Theoretical Framework

#### 4.1 Design Phase

The design phase establishes the topic, scope, objectives, and participants of the exercise (Grance et al., 2006). The topic proposed for this approach is cybercrime incidents and mitigation response using CSA. Problem-based learning (PBL) can be applied design of the tabletop exercise. PBL starts with a problem-solving scenario (De Graaf & Kolmos, 2003). For this paper, cybercrimes are the problem and topic for the tabletop exercise. The scope of the TTX exercise will focus predominately on the tactical level roles and responsibilities. The time scope is limited to 1-3 hours. The delivery scope should be virtual or hybrid to accommodate the industry professional. The objective of the TTX is to identify and analyze cybercrimes based on the CSA. This paper proposes groups of 8-10 student participants per team with facilitators from both academia and industry. Each participant will be provided with a real-world role and responsibility based on input from an industry professional.

Work-integrated learning (WIL) can be applied through the academic lecturer and industry professional working together to design the specific cybercrime, threat, and or attack as the topic, scope, and objective based on a real-world scenario. Some examples are account takeover, credit card fraud, credit card testing, gift card fraud, subscription fraud, triangulation fraud, social engineering, credential stuffing, romance scams, 1st person misuse, identity theft, denial-of-service, phishing, ransomware, advanced persistent threats (Dwight, 2023; Han et al., 2023). The industry professional can also act as a facilitator of the exercise. See Table 2 for the cybercrime tabletop exercise design outline.

Description	Model Synthesis
Cybercrime incident response and	PBL, WIL, CSA
mitigation	
Participant level: Tactical level	WIL
Duration: 1-3 hours	
Delivery: Virtual/blended	
Identify and analyze cybercriminal actions	PBL, WIL, CSA
and behaviors	
Identify and analyze cybersecurity	
mitigations and strategies	
Create a cybercrime script document	
based on the exercise	
Discuss the results of the exercise	
Groups of 8-10 Students	WIL
Two facilitators (one academic and one	
industry professional)	
	Description Cybercrime incident response and mitigation Participant level: Tactical level Duration: 1-3 hours Delivery: Virtual/blended Identify and analyze cybercriminal actions and behaviors Identify and analyze cybersecurity mitigations and strategies Create a cybercrime script document based on the exercise Discuss the results of the exercise Groups of 8-10 Students Two facilitators (one academic and one industry professional)

#### Table 2. Cybercrime Tabletop Design Outline

### 4.2 Development Phase

The development phase identifies all the documentation and tools necessary to conduct the exercise. This includes a briefing, facilitator guide, participant guide, and lessons learned report (Grance et al., 2006). The briefing consists of the agenda and logistics information. The facilitator guide consists of the purpose, scope, objectives, scenario, and list of questions. The participant guide consists of the shortened version of the facilitator guide without the list of questions. The lessons learned report contains evaluation criteria and reflective practice (Grance et al., 2006).

There are numerous tools to support the development and delivery of the exercise. The development of the documentation can use Generative AI. Generative AI presents an

opportunity to innovate and transform learning as higher education has been sluggish and underwhelming (Lim et al., 2023). Generative AI tools allow new frontiers in the way we learn, interact, and work with each other (Lim et al., 2023). Educational learning opportunities should use Generative AI to encourage transformation in the higher education realm. For this paper, we propose using Generative AI for the development of draft documentation and then revising it with input from the academic and industry partner.

For the delivery of the tabletop exercise, this paper proposes utilizing video conferencing software and a web-based interface (Zhou et al., 2015) to host the activity virtually or in a blended learning environment. Additionally, web-based interfaces can have meaningful effects on the scenario presentation and delivery (Zhou et al., 2015). This helps afford flexibility and cost-effectiveness for industry participation and academics (Chowdhury & Gkioulos, 2022).

# 4.3 Conducting Phase

The conduct phase is where the designed and developed exercise is executed with the relevant learners, processes, and systems. The facilitators start with the briefing and provide learners with the participant guide. The facilitator will procedurally go through the cybercrime script analysis stages of preparation, pre-activity, activity, and post-activity. At each stage, the learners will discuss and choose appropriate cybersecurity mitigation strategies. Each group of learners will fill out the CSA procedures and mitigations. Then a comparison of the results with other groups should occur in the evaluation phase.

# 4.4 Evaluation Phase

The evaluation phase captures the lessons learned, discussions, and reflections from the exercise. The learners will present their cybercrime scripts to each other and the facilitators. Then discuss the practicalities of the security measures with the facilitators. The industry professional can provide context about their experiences with the specific cyber-attack. See Figure 1 for the theoretical integrative approach. See Figure 2 for a sample draft briefing for a ransomware scenario, sample draft facilitator guide, participant guide, and after action guides created with the Generative AI technology Notion AI.


#### Figure 1. Cybercrime Script Tabletop Exercise Approach.

#### Crime Script: Ransomware Attack on an Ecommerce Company

onday morning, the IT department of an ecommerce company received an odd message on their computers. The message read "Your files have been encrypted. Pay \$10 million dollars in bitcoin to retrieve them." The IT department was confused and assumed it was a prank. But soon they realized that their entire system had been compr mised by a rans

The attackers had infiltrated the company's network through a phishing email that was sent to an unsuspecting employee. Once the employee clicked on the link in the email, the ransomware virus was downloaded onto the company's network. The virus quickly spread throughout the system, encrypting all of the company's files and rendering their data inaccessible

The attackers demanded a ransom of \$10 million dollars in bitcoin in exchange for the decryption key needed to unlock the company's files. The company was forced to shut down their website and all online sales until the issue could be resolved. The IT department worked tirelessly to try and recover the data, but all their efforts were in vain.

The company had no choice but to pay the ransom, as they were losing millions of dollars every day their systems were down. The payment was made and the attackers provided the decryption key. The company was able to retrieve their data, but not without significant financial and reputational damag

The incident was a wake-up call to the company and other ecommerce businesses, reminding them of the importance of cybersecurity measures and the need to stay vigilant against potential threats. The company took steps to improve their security protocols, including conducting regular security audits, educating employees on how to recognize phishing emails, and implementing multi-factor authentication.

Despite their efforts, the company's reputation took a hit and they lost customers due to the prolonged outage. The incident also attracted negative media attention and raised concerns among stakeholders about the company's ability to protect customer data.

In the aftermath of the attack, the company worked to repair its reputation and regain the trust of its customers. They offered compensation to affected customers and implemented additional security measures such as encryption and data backups. The company also engaged with cybersecurity experts to conduct a thorough investigation of the incident and identify ways to prevent future attacks.

While the company was able to recover from the attack, the incident highlighted the devastating impact that ransomware attacks can have on businesses. It also underscored the need for companies to prioritize cybersecurity and take proactive steps to protect their data and customers

#### TTX Facilitator Guide

This facilitator guide is based on the National Institute of Standards and Technology (NIST) Special Publication 800-84, "Guide to Test, Training, and Exercise Programs for IT Plans and Capabilities." The purpose of this guide is to provide facilitators with the necessary information to conduct a tabletop exercise (TTX) for ransomware incidents.

#### Introduction

The objective of a TTX is to test the organization's response plan for a ransomware incident. The TTX is designed to simulate a vare attack and evaluate the organization's response to the incident. It is important to conduct a TTX on a regular basis to identify gaps in the response plan and to improve the overall effectiveness of the plan.

#### Preparing for the TTX

ducting a TTX, the facilitator should identify the scope of the exercise, define the objectives, and identify the participants The facilitator should also create a scenario that simulates a ransomware incident, including the initial infection vector, the impact on the organization, and the response actions required.

#### Conducting the TTX

During the TTX, the facilitator should guide the participants through the scenario, providing them with information as needed. Th facilitator should also ensure that the participants adhere to the response plan and identify any gaps in the plan. The TTX should be conducted in a realistic and controlled manner, simulating the stress and pressure of a real ransomware incide

#### Evaluating the TTX

After the TTX, the facilitator should conduct a debriefing session with the participants. The debriefing session should identify any issues or gaps in the response plan and discuss ways to improve the plan. The facilitator should also document any lessons learned and update the response plan accordingly.

#### ... Conclusion

Conducting a TTX is an important part of an organization's overall ransomware incident response plan. It allows the organization to identify gaps in the plan and improve the overall effectiveness of the plan. This facilitator guide provides the necessary information to conduct a TTX for ransomware incidents based on NIST Special Publication 800-84.

#### TTX Participant Guide

ted to provide instructions to participants of the Tabletop Exercise (TTX) on how to respond to a ware attack. The guide is based on the guidelines provided in the National Institute of Standards and Technology (NIST) Special Publication 800-84, "Guide to Test, Training, and Exercise Programs for IT Plans and Capabilities".

#### Introduction

Ransomware has become a significant threat to organizations as it can cause significant disruptions to business operations and result in financial losses. It is essential for organizations to be prepared to respond to a ransomware attack effectively. A TTX is an excellent opportunity to test the organization's incident response capabilities and identify areas for improve

#### Objectives

The objectives of this TTX are as follows

- 1. Identify and analyze cybercriminal actions and behavior
- 2. Identify and analyze cybersecurity mitigations and strategies
- 3. Create a cybercrime script document based on the ex
- 4. Discuss and reflect on the results of the exercise

#### Scenario

The TTX will simulate a ransomware attack on the organization's network. The attack will start with an email phishing campaign, which will result in the installation of ransomware on several endpoints. The ransomware will then spread throughout the network encrypting files and demanding a ransom payment in exchange for the decryption key.

#### Roles and Responsibilities

The following roles and responsibilities have been assigned to the participants

#### Incident Response Team

The incident response team is responsible for identifying and containing the ransomware attack. The team will coordinate with other eams to ensure that the attack is contained and the impact on the organization's operations is m

#### TTX After Action guide

This document serves as an after-action guide for ransomware attacks based on the guidelines provided in NIST Special Publication 800-84. The purpose of this guide is to assist organizations in responding to ransomware attacks and to provide rec improving their security post

#### Summary of the Attack

Provide a brief summary of the ransomware attack, including the date and time of the attack, the systems and data affected, and the mpact on business operations.]

#### Initial Response

[Describe the organization's initial response to the attack, including who was notified, the actions taken to contain the attack, and any challenges encountered during the response.]

#### Recovery and Restoration

[Detail the steps taken to recover and restore the affected systems and data, including any external resources used, such as cybersecurity consultants or law enforcement.]

#### Analysis and Assessment

de an cybercrime script analysis of the attack, including the type of ransomware used, the attack vector, and any vulne exploited. Also, describe the impact of the attack on the organization's business operations, reputation, and financial status. Finally assess the effectiveness of the organization's security controls.

#### Lessons Learned

[Identify the lessons learned from the attack, including any improvements that can be made to the organization's security posture, incident response plan, or employee training programs.

#### Recommendations

(Provide recommendations for other organizations based on the lessons learned from this attack. These recommendations may include improving security controls, updating incident response plans, or increasing employee training.]

#### Figure 2. Sample Briefing, Facilitator, Participant, After Action Guides.

Note. Sample draft text generated using NotionAI. (2023, May). Prompt: "Write a crime script about ransomware at an ecommerce company." "Write a ransomware facilitator guide based on

https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-84.pdf". Prompt: "Write a ransomware participant guide based on https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-84.pdf". Prompt: "Write a ransomware after action report based on https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-84.pdf". Then the draft text was adapted by the researcher.

#### 5. Discussion

The aim of this research paper is to develop a novel teaching approach to develop the skills and experience of higher education students for cybersecurity roles. The gaps in the literature showcased active learning methods but did not map these methods to address the cybersecurity skills gap. The novel approach presented in this paper synthesizes PBL, WIL,

TTX, and cybercrime script analysis to deliver a cost-effective collaborative skill development model.

#### 5.1 Cybersecurity Education

Cybercrimes, threats, and attacks such as social engineering, fraud, and identity theft are sophisticated and continually evolving (Dwight, 2023; Han et al., 2023). Educators will need to stay abreast of the current threats to adapt the cybercrime script tabletop exercises for learners. This allows for many opportunities to continually update the problem and cybercrime scripts for up-to-date learning for higher education students. Educators should network and discuss current cyber threat trends with industry professionals. Industry professionals can help in any phase of the cybercrime script TTX approach including development, design, delivery, and evaluation.

The Cybercrime Script TTX can be integrated into a cybersecurity curriculum or as a standalone class session or workshop. However, this exercise should not be the sole active learning activity for an educator and student. This activity should complement other active learning activities such as internships, hackathons, labs, case studies, and others (Kay et al., 2019; Smith & Worsfold, 2015; Bilsland et al., 2019).

Using Generative AI can help with cyber education development. However, the technology is still in its infancy, and many people are wary about adopting the technology (Lim et al., 2023). This brief research used the technology to create rough drafts of the tabletop exercises. The ability of the Generative AI was not comprehensive, and academics will need to review and adapt the output. Additionally, gamification (Angafor et al., 2020), mixed reality (Shaytura et al, 2021), and other technologies can be integrated, used, and tested in any part of the TTX design, development, delivery, and evaluation process.

## 5.2 Industry Collaboration (WIL)

The level of involvement of industry professionals can be time-consuming and costly. This paper suggests minimizing the effort on the industry professional while providing an advantageous opportunity for the learners to potentially gain skills and interact with an industry professional in a real-world scenario. Chowdhury and Gkioulos (2022) Specified industry professionals preferred easy, cost-effective, and short-term activities. The level of involvement will ultimately depend on the time and resources available to industry professionals and academics. The more exposure the students have to the industry, the better they will be prepared for the real world.

#### 5.3 Limitations

There may be some limitations to the breadth of the paper. The extent to which the study's findings may be generalized and applied to other situations must be left to the reader's judgment. The scope of this paper was limited to designing a small, cost-effective event-based method that incorporated PBL, WIL, TTX, and CSA. All the technology and process nuances and variances associated with these paradigms may not be reviewed, evaluated, and discussed.

## 6. Contributions and Future Research

This cybercrime script tabletop exercise contributes to the body of knowledge with a novel approach to developing cybersecurity university students. This research paper provides a way to develop the skills and experience of higher education students for cybersecurity roles through the integration of active learning methods with The National Institute of Standards and Technology (NIST) Test, Training, and Exercise (TT&E) methodology (Grance et al., 2006).

Additionally, students can experience a real-world training method while interacting with an industry professional.

Second, this approach provides a way for industry professionals to get more involved with the development of university students. This model is especially useful to industry professionals who have limited time and resources to devote to students but want to contribute to student development and recruitment.

This paper serves as a starting point to apply the cybercrime script tabletop exercises in cybersecurity teaching practice. The next phase of this research can apply quantitative and qualitative methods to compare, interview, and observe student outcomes such as skill and experience acquisition, work-life awareness, and levels of industry professional involvement. Future studies can evaluate the differences between exercises with and without an industry facilitator, perceptions of students, and perceptions of industry professionals. Other research can investigate variations of this method, such as students going onsite to observe and participate in industry tabletop exercises.

#### 7. Conclusion

Cyber incidents continue to grow in sophistication and urgency, and there is a shortage of capable cybersecurity professionals. This paper contributes to the knowledge base by providing a theoretical teaching approach to integrate active learning methods of problembased learning (PBL), work-integrated learning (WIL), tabletop exercises (TTX), and crime script analysis (CSA) methods into The National Institute of Standards and Technology (NIST) Test, Training, and Exercise (TT&E) Event methodology to help bridge the cybersecurity industry skills gap. Future research can be conducted to compare, interview, and observe student outcomes such as skill and experience acquisition, work-life awareness, and levels of industry professional involvement.

#### Acknowledgements

The author would like to thank his family, the ICTDSE2023, ICCE2023, and the Royal Melbourne Institute of Technology for their support.

#### References

- Angafor, G. N., Yevseyeva, I., & He, Y. (2020, October). Bridging the cyber security skills gap: Using tabletop exercises to solve the CSSG crisis. In *Joint International Conference on Serious Games* (pp. 117-131). Cham: Springer International Publishing.
- Aprile, K. T., & Knight, B. A. (2020). The WIL to learn: Students' perspectives on the impact of workintegrated learning placements on their professional readiness. *Higher Education Research & Development*, 39(5), 869-882.
- Bada, M., & Nurse, J. R. (2021, June). Profiling the cybercriminal: a systematic review of research. In 2021 international conference on cyber situational awareness, data analytics and assessment (CyberSA) (pp. 1-8). IEEE.
- Bate, E., Hommes, J., Duvivier, R., & Taylor, D. C. (2013). Problem-based learning (PBL): Getting the most out of your students–Their roles and responsibilities: AMEE Guide No. 84. *Medical teacher*.
- Bilsland, C., Carter, L., & Wood, L. N. (2019). Work integrated learning internships in transnational education: Alumni perspectives from Vietnam. *Education+ Training*, *61*(3), 359-373.
- Bodker, A., Connolly, P., Sing, O., Hutchins, B., Townsley, M., & Drew, J. (2022). Card-not-present fraud: using crime scripts to inform crime prevention initiatives. *Security Journal*, 1-19.
- Brilingaite, A., Bukauskas, L., Krinickij, V., & Kutka, E. (2017, October). Environment for cybersecurity tabletop exercises. In *ECGBL 2017 11th European Conference on Game-Based Learning* (pp. 47-55). Academic Conferences and publishing limited.
- Clancey, G. (2020). Teaching crime prevention and community safety. *Scholarship of Teaching and Learning in Criminology*, 59-85.
- Chowdhury, N., & Gkioulos, V. (2022, September). A Framework for Developing Tabletop

Cybersecurity Exercises. In *European Symposium on Research in Computer Security* (pp. 116-133). Cham: Springer International Publishing.

- Cornish, D. B. (1994). The procedural analysis of offending and its relevance for situational prevention. *Crime prevention studies*, *3*(1), 151-196.
- De Graaf, E., & Kolmos, A. (2003). Characteristics of problem-based learning. *International journal of engineering education*, *19*(5), 657-662.
- Dehghanniri, H., & Borrion, H. (2021). Crime scripting: A systematic review. *European Journal of Criminology*, *18*(4), 504-525.
- Dwight, J. (2023). Ecommerce Fraud Incident Response: A Grounded Theory Study. *Interdisciplinary Journal of Information, Knowledge, and Management,* 18, 173-202.
- Fortinet. (2023). 2023 Cybersecurity Skills Gap Report. Retrieved from <u>https://www.fortinet.com/content/dam/fortinet/assets/reports/2023-cybersecurity-skills-gap-report.pdf</u>
- Grance, T., Nolan, T., Burke, K., Dudley, R., White, G., & Good, T. (2006). Guide to test, training, and exercise programs for IT plans and capabilities.
- Hallinger, P. (2020). Mapping continuity and change in the intellectual structure of the knowledge base on problem-based learning, 1974–2019: A systematic review. *British Educational Research Journal*, *46*(6), 1423-1444.
- Han, K., Choi, J. H., Choi, Y., Lee, G. M., & Whinston, A. B. (2023). Security defense against long-term and stealthy cyberattacks. *Decision Support Systems*, *166*, 113912.
- Kay, J., Ferns, S., Russell, L., Smith, J., & Winchester-Seeto, T. (2019). The Emerging Future: Innovative Models of Work-Integrated Learning. *International Journal of Work-Integrated Learning*, 20(4), 401-413.
- Leppänen, A., Toiviainen, T., & Kankaanranta, T. (2020). From a vulnerability search to a criminal case: script analysis of an SQL injection attack. *International Journal of Cyber Criminology*, *14*(1), 63-80.
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, *21*(2), 100790.
- Lunn, S., Ross, M., & Liu, J. (2021, December). Cybersecurity Integration: Deploying Critical Infrastructure Security and Resilience Topics into the Undergraduate Curriculum. In 2021 International Conference on Computational Science and Computational Intelligence (CSCI) (pp. 866-871). IEEE.
- McGettrick, A. (2013). Toward effective cybersecurity education. *IEEE Security & Privacy*, *11*(6), 66-68.
- Nguyen, K. A., Borrego, M., Finelli, C. J., DeMonbrun, M., Crockett, C., Tharayil, S., ... & Rosenberg, R. (2021). Instructor strategies to aid implementation of active learning: a systematic literature review. *International Journal of STEM Education*, *8*, 1-18.
- Reeves, A., Calic, D., & Delfabbro, P. (2021). "Get a red-hot poker and open up my eyes, it's so boring" 1: Employee perceptions of cybersecurity training. *Computers & security*, *106*, 102281.
- Schneider, F. B. (2013). Cybersecurity education in universities. *IEEE Security & Privacy*, *11*(4), 3-4. Shaytura, S., Olenev, L., Nedelkin, A., Ordov, K., Minitaeva, A., & Guzhina, G. (2021, November).
- Mixed reality in education and science. In 2021 3rd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA) (pp. 667-673). IEEE.
- Smith, C., & Worsfold, K. (2015). Unpacking the learning–work nexus: priming'as lever for high-quality learning outcomes in work-integrated learning curricula. *Studies in Higher Education*, 40(1), 22-42.
- Towhidi, G., & Pridmore, J. (2023). Aligning Cybersecurity in Higher Education with Industry Needs. *Journal of Information Systems Education*, *34*(1), 70-83.
- Zhang, Z., He, W., Li, W., & Abdous, M. H. (2021). Cybersecurity awareness training programs: a cost–benefit analysis framework. *Industrial Management & Data Systems*, *121*(3), 613-636.
- Zhou, B., Sun, G., Zhang, X., Xu, J., Lai, J., Du, X., ... & Sakurada, Y. (2015). Development of webbased tabletop emergency earthquake exercise system. *Journal of Disaster Research*, 10(2), 217-224.

# Local Safety Map for a Cornerstone of Local Disaster Prevention: A Case Study of a Historical Local Town in Japan

Yasuhisa OKAZAKI \*, Hiroshi WAKUYA, Yukuo HAYASHIDA & Nobuo MISHIMA \*Faculty of Science and Engineering, Saga University, Japan \*okaz@cc.saga-u.ac.jp

Abstract: This paper examines the requirements for creating a local safe map, which is the core of regional disaster prevention, through a case study in a historical local city in Japan. Amid the rising awareness of disaster prevention, efforts are being made in various places to voluntarily create disaster prevention maps through town walks. It is expected to improve residents' awareness of disaster prevention, improve local disaster prevention capabilities, and revitalize local communities. What are the requirements for a local safety map creation to continue to play a central role in regional disaster prevention rather than being a temporary event? Based on our past experiences, we review our research activity what is necessary to make such map-making the core of regional disaster prevention. The following is findings through our experience in the map-making activities. Sustainable development is possible by making it a regular activity in the district rather than a one-off event. Sharing of subjective information in cooperation with the disaster prevention division leads to improve reliability of information by consensus with the members and revitalization of activities through workshops. it is also necessary to select media that provide information from the user's perspective.

**Keywords:** local disaster prevention, safety map, voluntary disaster prevention, historical local town, community participation

## 1. Introduction

In response to the occurrence of natural disasters such as earthquakes and heavy rains, hazard maps that expresses disaster risks are attracting attention(ICTDSE2021 in ICCE2021, Japan Tourism Agency,2022). Hazard maps are used for disaster preparedness and evacuation behavior. In addition to these hazard maps, voluntary disaster prevention maps are also actively being worked on (Mitsuhara, H., 2018; Yoshino, T., et al., 2017). By walking around one's own town and researching it, information on possible hazards in the area is gathered, and the information is used to improve each person's awareness of disaster prevention and evacuation behavior (Enokida, S., et al., 2018). In particular, in areas where historical townscapes have been preserved (Japan Guide.com, 2012), it is difficult to respond to disasters through road improvements and rebuilding. Therefore, it is desirable to highlight and share the dangers that local residents feel in their daily lives from the perspective of local residents (Mishima, N., et.al, 2014).

We have selected Hizen-Hamashuku, Kashima City, Saga Prefecture, which still retains the townscape of the Edo period (Agency for Cultural Affairs, 2017; Saga Trip Genius, 2014), as a model area, and have been working with local voluntary disaster prevention organizations to support the creation of local safety maps using ICT (Okazaki, Y. et al., 2016; 2017; 2018; 2019; 2020;2021). There are two national Important Preservation Districts for Groups of Traditional Buildings here. The government continues to implement measures based on past disasters, such as installing sprinklers in thatched houses, installing fire hydrants in areas where wooden houses are densely built, and repairing the river that flows

through the center of the district. In addition, active disaster prevention activities have been carried out through collaboration among residents through voluntary disaster prevention organizations.

In this study, we examine what is necessary for a local safety map to continue to play a central role in regional disaster prevention. From the perspective of (1)Revitalization of voluntary disaster prevention organizations through efforts to create local safety maps, (2)the relationship with the official hazard map and (3)Information provision media from the perspective of residents, we examine the mechanism for continuing to work as an activity that contributes to the safety and security of the region, not limited to natural disasters, and present future issues.

## 2. History of our research activities

Our activities over the past seven years can be divided into four phases (Table 1).

Phase 1	2015-2017	Map creation					
Collecting information by walking around the town + Summarizing as a map							
Local danger informat	ion sharing (Limited to t	those involved)					
(Okazaki, Y. et al., 20	16; 2017; 2018)						
Phase 2	Phase 2 2018-2021 Map improvement						
Review of information	presented through wor	kshops and meeting					
Update information, re	consider display metho	ods and media					
Review of information by local residents							
(Okazaki, Y. et al., 2019; 2020;2021)							
Phase 3 2022 Expansion of target area							
Expand area and update information							
Expansion and improvement of information							
Phase 4 2023 Expansion of target people (plan)							
Update and review information for distribution to all households							
Local danger information sharing (Everyone in the community)							

Table 1. History of our community-based disaster prevention activities using ICT

In Phase 1, we developed a hazard map creation support system with the aim of collecting specific local information and centrally managing and visualizing it. We realized information gathering by voluntary disaster prevention officers walking around town, centralized management of information using a server, and information visualization.

In Phase 2, we verified the collected information and improved the local safety map. We examined the type of information to be presented, the presentation method, and the media to be presented to improve the visibility and reliability of the local safety map.

In Phase 3, the target area was expanded to the entire Hizenhama district. Participation from districts that had not participated in the past was also made, and further information was added and updated to enhance the map. As a result, we were able to create a foundation for a relief map for the entire district.

In Phase 4, we aim to distribute the map to all households in the district, and we are proceeding with work toward completion of the map as a training session for voluntary disaster prevention. Figure 1 shows the latest version of the map.

Through the ongoing implementation of these community activities, we analyze the impact and challenges of creating the local safety map on community activities.

## 3. Continuous map making activities for safe community

## 3.1 Revitalization of local community activities

Since before 2015, when this activity began, firefighting drills and meetings of voluntary disaster prevention officers have been carried out by local voluntary disaster prevention organizations. At the meeting, risk information was exchanged between the officers of each district, but it was an oral exchange of opinions, and it was not enough to confirm the scattered danger information of areas.



Figure 1. The local safety map(January 2023 edition).

Therefore, first of all, efforts were made to collect and visualize the danger information of these areas. We developed a system using an iPad, and were able to collect information efficiently through city walks by district voluntary disaster prevention officers (Okazaki, Y. et al., 2016). By using an information collection system, it is possible to collect data, including photographs, in a predetermined format immediately on the spot, and by saving that data on a server, centralized management of data is possible (Okazaki, Y. et al., 2017).

In order to verify the collected data, a workshop was held as a training session for voluntary disaster prevention. It was an opportunity to reconsider the whole area by aggregating the scattered danger information of the areas. Participating in the workshop gave the participants new realizations, and it was a good opportunity for them to update their own local danger information (Okazaki, Y. et al., 2018).

## 3.2 Relationship with the disaster prevention division and hazard maps

The local safety map is an extension of the Voluntary Disaster Prevention Map. The information is based on the subjectivity of the residents, and may not have sufficient scientific basis. There is also the possibility of overlooking invisible dangers. From this point of view, it is important to cooperate with the disaster prevention division in government agencies.

In our activities, city officials in charge of disaster prevention participate in town walks and information review workshops, so that information on map creation as a voluntary disaster prevention activity of the district is conveyed to the administrative side. In addition, we sent the created map to the disaster prevention division of the city which is in charge of creating the hazard map in the area, and receive feedback on evaluation comments, which we use to make improvements. In addition, in March 2021, we conducted an interview survey with the staff of the disaster prevention division and evaluated the contribution to the safety and security of local residents by various hazard maps provided by the city and this local safety map (Okazaki, Y. et al., 2021).

## 3.3 Paper media or digital devices

The choice of media for providing information should be carefully considered, taking into account the circumstances of use and the characteristics of each media. We have developed and put into practice a local safety map creation support system using an iPad. Its effectiveness was demonstrated through information gathering through town walks and information sharing and verification through workshops by members of voluntary disaster prevention organizations (Okazaki, Y. et al., 2019).

It was shown that it is possible to operate with a tablet that is not usually used, and we thought that we could proceed with the creation of a local safety map using a digital map. However, there was an opinion that conventional paper maps would be preferable considering the use of maps in addition to resistance to putting local information, especially information related to crime prevention, online. Therefore, we have explored two directions: digital maps and paper maps (Okazaki, Y. et al., 2020). In terms of data management of digital maps and paper maps, we created the map using an electronic map Zi20 by ZENRIN CO., LTD. A problem was found in the size and detail of the map.

Assuming use on the Internet, we also tried porting to the map service "Locamo Share" by Locamo AI. However, integration with paper maps was difficult, and there were two independent maps, which left challenges in data management. In response to the request from a voluntary disaster prevention organization to distribute a paper map to households so that it can be used widely, the paper map will be prioritized first (Okazaki, Y. et al., 2021).

## 4. Discussion

In this section, we discuss the revitalization of the activities mentioned in 3.1 and the relationship with the disaster prevention division mentioned in 3.2 from the viewpoint of information aggregation and sharing of subjective information, as well as the information provision media.

## 4.1 Effects and Issues of Map Creation Based on Subjective Information

We have created a community-based voluntary disaster prevention map that complements government hazard maps by aggregating the subjective information of local residents. In addition, we have held workshops using the map and have been working on the safety and security of the district. Surrounding the map on which the specific information that they had collected led to more lively discussions than ever before, and information about the district was widely shared. It is thought that the activities were activated by visualizing information, sharing information, and working together on the map.

On the other hand, problems caused by subjective information became clear. Information from walking around town is subjective information, and lacks objective scientific grounds. Cooperation with the government is necessary to ensure the quality of information. This cooperation is significant for the administration as it provides an opportunity to learn about the awareness of residents. In addition, as circumstances change, it becomes necessary to update the map. Besides that, requests for map corrections from residents continued. The factors are classified into subjective gaps and individual fluctuations. In the former, since the subjective view differs from person to person, the point and the range of dangerousness may differ from person to person. In the latter, even those who have participated in the activity so far are requested to correct existing information and add or delete new information in the middle of the activity. This is thought to be due to the fact that individual perceptions have changed as a result of repeated examination of information.

In any case, instead of reflecting the opinion of one individual on the map as it is, we have been discussing in groups and incorporating information that has been agreed upon. Continuing these activities will promote the sharing of information that has reached a consensus within the community, and we believe that once a certain amount of information is shared, there will be fewer changes in information due to differences in subjectivity. At this stage, we believe that by distributing the map to the entire district, we can play a role as a relief map that contributes to the safety and security of the entire district.

#### 4.2 Map providing media

We have been making a local safety map by local residents using the safe map making support system. The map creation support system centrally manages information on a server, and has the advantage of being able to visualize information on terminals, which actually contributed to workshops using iPads. Compared to the usual gatherings and oral information exchanges, active exchanges of opinions took place while looking at the iPad screens. On the other hand, problems due to digitization have also become apparent.

The advantage of digital maps is that information is easy to manage, and it is possible to search and select the information to be presented, and it is possible to present a variety of information as needed. However, it was pointed out that the lack of a bird's eye view of the whole and the usability issue that it cannot be viewed unless it is started on the device are disadvantages. These disadvantages have a large impact due to the nature of the local safety map. Also, the workshop was conducted using both an iPad and a paper map, but the paper map was easier to implement. This is thought to be due to the direct operation feeling of pasting sticky notes and the bird's-eye view, which gives a real feeling of working.

Also, the quality of the paper map is an element that needs attention. The map offered by ZENRIN CO., LTD. describes the shape of each house. Also, the trails are accurately marked. When compared with other free maps, some commented that there was a clear difference in terms of the clarity of visuals. It can be said that accurate description of the details of familiar areas is an important factor in gaining the trust of local residents.

Our work demonstrates that the premise that it should be digital may not work. In order to solve the issues of bird's-eye view and visibility, a new fusion of paper and digital is required. Electronic paper exists as a means to solve the problem of bird's-eye view digitally, but at present it is expensive and it is not realistic to post it casually like paper.

#### 5. Summary and future works

In this study, we considered the requirements for a local safety map to continue to play a central role in local disaster prevention based on our experience of research activities over the past eight years. The target is a district where historic townscapes are being preserved. In addition to solving the problem of insufficient information sharing due to sporadic verbal communication before our research, by creating a map, we found that collaborative work through the map activates local disaster prevention activities. The issue of reliability of subjective information obtained from town walks can be resolved through continuous study through workshops and cooperation with administrative agencies. We have created a local safety map that summarizes the information collected by the local people. Trough workshops using the map, we clarified the variation in danger information that relied on individual subjectivity, and by repeating active discussions, we were able to contribute to share safe information with a certain level of reliability in the district. In the process, we also showed that it is necessary to consider the method of providing information by assuming the actual use situation.

Future issues include verification of the validity of the information posted on the map and visualization of the information update process. Also, the current map is from an adult's point of view. Children may have a different point of view than adults. Reflecting the opinions of such children is also an issue for the future. Furthermore, It is also necessary to reconsider research methods. Evaluation of our research up until now has been by activity observations, questionnaires and interviews. It is difficult to conduct quantitative verification based on experiments and statistical data in social case studies. Reexamination of the qualitative research approach, which is known as one of the social science methodologies, is also our future issue.

#### Acknowledgements

This study is supported by JSPS KAKENHI Grant Number 19H02315. We would like to thank all who understood and cooperated on our on-site field work and our interview survey.

#### References

Agency for Cultural Affairs. (2017). Preservation Districts for Groups of Traditional Buildings August 1, 2021, Retrieved April 26, 2023, from

http://www.bunka.go.jp/english/policy/cultural\_properties/introduction/historic\_buildings/

- Enokida, S., Fukushima, T., Yoshino, T., Sugimoto, K., Egusa, N. (2018). Proposal of an Integrated Support System for Disaster-preparedness Map Making Through Town-walk Type of Gathering Information. *IPSJ Journal*, 59(3), 992-1004.
- ICTDSE2021 in ICCE2021. (2021). Retrieved April 26, 2023, from https://sites.google.com/view/ictdse/ Japan Guide.com. (2012), Historic Sites. Retrieved April 26, 2023, from
- http://www.japan-guide.com/e/e2422.html
- Japan Tourism Agency. (2022). Safety tips for travelers, Retrieved April 26, 2023, from https://www.jnto.go.jp/safety-tips/eng/
- Mishima, N., Miyamoto, N., Taguchi, Y., Kitagawa, K. (2014). Analysis of current two-way evacuation routes based on residents' perceptions in a historic preservation area, *International Journal of Disaster Risk Reduction*. 8, 10-19.
- Mitsuhara, H. (2018). Special Issue: Educational Systems for Safe and Secure Society Practical/Support Systems for Programming and Information Technology Education, *Transactions of Japanese Society for Information and Systems in Education*. 35(2), 64-93.
- Okazaki, Y., Mori, S., Wakuya, H., Mishima, N., Hayashida, Y. & Min, B-W. (2016). Development of a Sustainable Community-based Hazard Map Creation Support System for Traditional Towns with Local Heritage, *International Journal of Contents*. 12(2), 58-65.
- Okazaki, Y., Kozaki, S., Matsuo, S., Wakuya, H., Mishima, N., Hayashida, Y. & Min, B-W. (2017). Prototyping of Community-based Hazard Mapping Support System for Traditional Towns with Local Heritage. *25th International Conference on Computers in Education Workshop Proceedings*, 171-179.
- Okazaki, Y., Kozaki, S., Matsuo, S., Wakuya, H., Mishima, N., Hayashida, Y., & Min. B-W. (2018). Review of Hazard Information Collected for Local Disaster Prevention by Residents in a Historical Local Town. *26th International Conference on Computers in Education Workshop Proceedings*, 1-5.
- Okazaki, Y., Matsuo, S., Wakuya, N., Hayashida, Y., & H., Mishima, (2019). Practical Evaluation of ICT-based Self-made Regional Safety Map through Residents' Workshop in a Historical Local Town. 27th International Conference on Computers in Education Proceedings, Vol. II, 209-216.
- Okazaki, Y., Taniguchi, T., Wakuya, N., Hayashida, Y., & H., Mishima, (2020). Prototype of Paper Map for Practical Use of Regional Safety Map "Hamādo-map" and Its Questionnaire Survey. 28th International Conference on Computers in Education Proceedings, Vol. II, 208-214.
- Okazaki, Y., Megro, T., Wakuya, N., Hayashida, Y., & H., Mishima, (2021).Qualitative Evaluation of Information Display in a Regional Safety Map "Hamādo-map", *Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education*, Volume II, 564-569.

Saga Trip Genius. (2014). Hizenhamashuku Area. Retrieved April 26, 2023, from http://www.saga-tripgenius.com/tourism\_search/hizenhamashuku-area.html

Yoshino, T., Hamamura, Y., Fukushima, T., Egusa, N. (2017). Making of Disaster-prevention Maps by Local Residents Using the AkariMap Evacuation Support System. *IPSJ Journal*, 58(1),215-22.

# Non-Player Characters for Evacuation Training in Metaverse: Preliminary Experiment

Kaito OE<sup>a\*</sup>, Yusaku ICHINO<sup>a</sup>, Hiroyuki MITSUHARA<sup>b</sup> & Masami SHISHIBORI<sup>b</sup> <sup>a</sup> Graduate School of Sciences and Technology for Innovation, Tokushima University, Japan <sup>b</sup> Graduate School of Technology, Industrial and Social Sciences, Tokushima University,

Japan

\* c612235054@tokushima-u.ac.jp

**Abstract:** Evacuation training is indispensable in terms of surviving disasters. Recently, virtual reality technologies have been used to facilitate evacuation training. The authors previously proposed the Evacuation Training in Metaverse (ETM) system. This system is currently being evaluated in a prototyping phase, where many users can spend time in a three-dimensional virtual university campus. In the ETM system, an earthquake occurs suddenly, and the users must determine various actions, e.g., whether to evacuate and where to evacuate. However, the ETM system does not always have a sufficient number of active users. Thus, the ETM system can generate non-player characters (NPC) based on the evacuation logs of previous participants. This paper overviews the ETM system and describes the method used to generate NPCs. In addition, the results of a preliminary comparative experiment are discussed.

Keywords: Non-player character, evacuation training, virtual reality, metaverse

## 1. Introduction

Natural and manmade disasters can destroy lives, and people are particularly vulnerable to unpredictable disasters, e.g., earthquakes and terrorist attacks. As a result, learning how to survive disasters is becoming increasingly important. To survive disasters, we must evacuate to a safe place quickly; however, speedy evacuation is difficult when facing unanticipated disaster situations. For example, if a recommended evacuation route is impassable, we must determine another route. In addition, when encountering an injured person, we must decide whether to attempt a rescue. Traditional evacuation training does not necessarily simulate such situations. In other words, traditional evacuation training is conducted according to a simple scenario that encourages participants to remember recommended evacuation routes. In addition, we tend to avoid evacuation training because we frequently feel that disasters will not affect us directly. Thus, a new evacuation training paradigm is required to remove such feelings and motivate participation.

A prospective approach is to introduce virtual reality (VR) into evacuation training. VR technologies can express disaster situations realistically using sophisticated three-dimensional (3D) computer graphics and have attracted attention due to their technological novelty. VR-based evacuation training (VRET) can provide participants with simulated evacuation experiences according to various scenarios in a safe environment, and many VRET systems have been proposed (Gagliardi et al., 2023). If VRET is conducted in a virtual world resembling the real world, participants are expected to feel that a disaster is a direct and imminent problem. For example, a VRET system provides simulated earthquake evacuation in a 3D virtual world that is faithfully modeled after a real hospital (Feng et al., 2020). By focusing on fire evacuation, a VRET system can realistically reproduce structure fires by simulating and visualizing spreading fire and smoke, as well evacuees (Lorusso et al., 2022). Another VRET system was designed to train underground rail station employees to

guide passengers to exits without encountering fire, while also introducing fire and passenger agent simulations (Wang et al., 2023).

Recently, VR has been popularized by the advent of reasonable consumer VR headsets, e.g., the Meta Quest and HTC VIVE devices. In addition, the metaverse has attracted increasing attention as a social VR paradigm that enables multiple users to communicate and interact in a virtual world. The metaverse can be applied to various fields, including education (Tlili et al., 2022). For example, a university campus has been reproduced as a metaverse to provide lectures via a presentation tool and interactions among avatars (Nagao, 2023). In addition, a multiuser collaborative VR system allows participants to learn how to survive active shooter events (Sharma, 2020).

The authors previously proposed Evacuation Training in Metaverse (ETM) and are prototyping an ETM system that focuses on earthquake evacuation (Mitsuhara & Shishibori, 2022; Oe et al., 2022). The ETM begins when an earthquake occurs suddenly in the metaverse. The ETM system is prototyped using a the Unity game engine that enables cross-platform software development. The target VR headset is the Meta Quest 2 device. The current ETM system, which includes a virtual world that resembles the Tokushima University campus, allows at most 200 users (i.e., avatars) to access the virtual world synchronously. Evacuation processes with many users can generate or emphasize difficult situations. In other words, including many participants can make evacuation training more realistic and lead to higher training effect. For example, if many users rush into a narrow path, the evacuation will be delayed due to congestion. Such simulated evacuation experiences can facilitate effective learning, e.g., "I choose a wider route calmly." However, many users are not always present in the metaverse, and evacuation training with a smaller number of users may not be realistic or effective.

From this perspective, we introduce non-player characters (NPC) into the ETM system to simulate the presence of many users in the metaverse system. Note that many VRET systems have introduced NPCs. For example, Feng et al. (2022) developed a VRET system for earthquake evacuation in a hospital environment that included doctor and visitor NPCs. They found that the participants tended to follow a doctor NPC without making independent decisions. Liu et al. (2023) examined the effects of video-based emergency evacuation training by observing whether participants followed NPCs in a VR emergency simulator. In VRET systems, NPCs are preprogrammed based on evacuation behavior models (e.g., Dijkstra's algorithm). We believe that such models do not necessarily represent realistic evacuation behaviors, e.g., panic and upset. Thus, we propose a method that generates NPCs from participants' evacuation logs.

The remainder of this paper is organized as follows. Section 2 overviews the ETM system. Section 3 presents the method used to generate NPCs. Section 4 reports on a preliminary comparative experiment conducted to examine real participants and NPCs. Section 5 concludes the paper, including suggestions for future work.

## 2. ETM System

The ETM system has the fundamental metaverse functions, e.g., a 3D virtual world, synchronous multiuser access, and voice communication.

#### 2.1 Requirements

In addition to the fundamental functions, the ETM system satisfies the following requirements for earthquake evacuation training.

**High-fidelity 3D virtual world:** The ETM system's virtual world, i.e., the 3D university campus, should be identical to the real-world university campus. Thus, the 3D virtual world was modeled in reference to actual building blueprints. In the high-fidelity 3D virtual world, students can take a class or converse with other users as if they were in the real-world

classroom. In other words, the 3D virtual world should be familiar and comfortable. Figure 1 shows screenshots of the 3D virtual world.



Figure 1. 3D virtual world (left: panoramic view; right: a classroom).

**Avatars:** Metaverse users may want to use favored avatars (e.g., an animal or a robot) to represent their identities. In addition, the users may expect that avatars have superhuman abilities, e.g., flying, teleportation, or immortality. However, the ETM system focuses on evacuation training; thus, the avatars should resemble users in appearance (e.g., the face and clothes) and abilities (e.g., movement speed and arm strength). For example, a user whose leg is injured should be represented in a squat down position with reduced movement speed. Such avatars can make evacuation training more realistic and effective. Figure 2 shows screenshots of different avatars in the ETM system. Currently, the system uses 3D avatars created by the Ready Player Me avatar modeling service (<u>https://readyplayer.me/</u>).



Figure 2. Avatars (left: standing; right: squatting down).

**Sudden earthquake:** The ETM system should simulate the sudden occurrence of an earthquake event. In Japan, several seconds before a large magnitude earthquake occurs, an early warning system signals the occurrence via smartphone, television, radio, etc. Thus, the ETM system functions according to the warning. Several seconds later after the warning signal, the virtual world begins to shake. In other words, the ETM system suddenly shifts from normal conditions to emergency conditions in the virtual world, and this sudden shift to earthquake conditions facilitates realistic and effective evacuation training.

**Reflection:** Evacuation training frequently ends when the participants reach a designated safe location. To enhance the effect of training, the participants should also reflect on their evacuation (e.g., the evacuation route and speed) after reaching the designated safe location. With sufficient reflection, the participants are expected to obtain ideas or their own rules to realize more successful evacuation, e.g., "We should move more speedily on a wider route." Thus, the ETM system should provide a function to support this reflection process.

## 2.2 Modes

The ETM system has three modes, i.e., the normal time, emergency time, and reflection time modes. In the normal time mode, the users can act freely in the 3D virtual campus. When a sudden earthquake event occurs, the mode switches to the emergency time model, where the users can also act freely. In other words, the users can self-determine whether to evacuate. Here, some users may begin an evacuation process, and the other users may choose to not evacuate. Users who begin the evacuation process are considered the participants in the ETM. After a predetermined amount of time passes in the emergency time mode, the mode switches back to normal time, i.e., the evacuation training finishes. In the normal time mode, the participants can freely enter the reflection time mode. Figure 3 illustrates the transitions between these three modes.



*Figure 3.* Mode transition.

#### 2.2.1 Normal Time Mode

Regardless of the time mode, the ETM system must accommodate multiple simultaneous users that can move and communicate in the virtual world. To achieve this, the ETM system utilizes network frameworks provided by Photon. Photon Fusion can synchronize multiple users in a virtual world (a conceptual room for synchronization created by a host computer) via a server on the Photon Cloud Network. In addition, Photon Voice supports voice chat in the virtual world.

#### 2.2.2 Emergency Time Mode

When a large magnitude earthquake occurs in Japan, the ETM system receives information about the earthquake from the Japan Meteorological Agency and generates an earthquake in the virtual world by shaking the 3D university campus environment, and the emergency time mode is initiated. This shaking influences the 3D objects in the virtual world; however, this increases the rendering load (especially with the Meta Quest 2 device). To reduce the rendering load and shake-induced VR sickness, the ETM system can omit rendering the shakes. Disaster situations occur in the virtual world based on a scenario that describes the type (e.g., fire and smoke) and location (x, y, and z coordinates) of the disaster situations.

In the emergency time mode, the participants move to a safe place, e.g., a building exit. Note that the ETM system does not present any message to indicate evacuation locations or routes. If the user has a microphone, the participants can communicate via voice; thus, they may evacuate cooperatively. When recognizing non-evacuating users, the participants may underestimate the necessity of evacuation and stop evacuating. This situation is caused by the cognitive biases (e.g., normalcy bias and conformity bias) observed in real-world evacuation situations.

The ETM system evaluates the success or failure of evacuation by determining whether each participant reached a safe location within the predetermined time. In the emergency time mode, the ETM system records all participants' evacuation logs.

#### 2.2.3 Reflection Time Mode

The participants can begin reflecting on their evacuation process by selecting an evacuation log listed on a menu. In the reflection time mode, the participants' logs are visualized such that

they can observe their evacuation behaviors objectively from a first-person or third-person perspective. Figure 4 shows screenshots of the reflection time mode.



Figure 4. Reflection time mode (left: first-person perspective, right: third-person perspective).

## 3. NPC Generation

The ETM system can heighten realism and training effect by including many users (avatars) in the virtual world. However, many users are not always present the virtual world; thus, the ETM system must be able to increase the number of users in the virtual world as required by generating NPCs based on evacuation logs.

## 3.1 NPC Generation based on Evacuation Logs

In video games, NPCs frequently perform intelligent actions, and artificial intelligence (AI) technology is actively used to provide such realistic behavior. In crowd evacuation simulation systems, NPCs are referred to as agents and are frequently preprogrammed based on behavior models. If preprogrammed NPCs are introduced into the ETM system, participants can be induced to fail in evacuation. For example, if NPCs are preprogrammed to rush to a short but narrow path, the participants may follow these NPCs and become stuck in a crowded path. The ETM system attempts to provide participants with simulated evacuation experiences and promote effective reflection, especially in terms of evacuation failures. When introducing preprogrammed NPCs that induce failure, the ETM system can provide a better training effect.

The NPCs generated based on evacuation logs are expected to represent realist evacuation behavior and influence the participants' evacuation behavior. For example, past participants who were panicked or upset may be represented as NPCs moving back and forth, which could induce emotional instability in the participant. An advantage of this NPCs generation method is that various beneficial situations can be represented without requiring complex behavior models. However, a disadvantage is that the ETM system cannot necessarily induce participants to fail in evacuation because NPCs cannot be controlled. In this study, to avoid the need for behavior models, we generate NPCs based on participant evacuation logs.

The ETM system enables participants to conversate via voice chat; however, the system does not currently record voice conversations, which means that the NPCs can only represent visual aspects. In real evacuation situations, conversation among evacuees can influence their evacuation behavior. For example, if an evacuee loudly proclaims an evacuation route, other evacuees may follow that route even though safer evacuation routes may be available. We consider that NPCs in the ETM system should focus on visual and auditory aspects.

## 3.2 Recording evacuation logs

In the emergency time mode, the ETM system collects participants' evacuation logs (for both users and NPC participants), i.e., each avatar's location and rotation (relative angle) in 3D coordinates, and it transfers the collected logs to the server every 0.1 s. On the server, a log recorder module stores the received logs as text files for each scenario.

## 3.3 Reproducing evacuation logs

In the emergency time mode, the ETM system generates NPCs based on the data in the evacuation logs. Here, every 0.1 s, a log reproducer module on the server fetches the evacuation logs of the current scenario to influence the behavior of the corresponding NPCs. The ETM system renders NPCs with hand and leg animations when moving, which is limited within the viewing field of each avatar to reduce rendering load for the Meta Quest 2 device. Note that the rendered NPCs are not semitransparent characters (e.g., ghost characters in video games). Currently, NPCs are rendered as a few fixed avatars that do not resemble real users. Figure 5 shows the mechanism of generating NPCs with screenshots of NPCs.



Figure 5. NPC generation mechanism.

## 4. Preliminary Experiment

A preliminary experiment was conducted to investigate and compare the influence of NPCs in the ETM system.

#### 4.1 Experimental settings

#### 4.1.1 Participants

In this experiment, 12 participants (Tokushima University students who typically spend time at the university campus) were expected to move without disorientation inside the 3D virtual campus. The participants were divided into two groups. In each group, three participants wore an immersive head-mounted display device (i.e., a Meta Quest 2), and the other participants (who were prone to VR sickness) used a desktop PC.

**Group A (N = 6):** Six participants evacuated simultaneously by operating their avatars in the virtual space. In other words, Group A was a multiuser setting.

**Group B (***N* **= 6):** Each participant performed the evacuated using their avatar with five NPCs generated from Group A's evacuation logs (i.e., the evacuation route and body orientation). In other words, Group B was a single-user setting.

#### 4.1.2 Procedure

The experimental procedure was identical for both groups, except for the presence or absence of NPCs. In the emergency time mode, the participants' evacuation logs and voices were recorded. The experimental procedure is summarized as follows.

- I. The participants spawned outside a six-floor lecture building.
- II. According to the experimenter's request, the participants moved to the fifth floor of the building. This step allowed the participants to practice operating their avatars.

- III. When all participants (including the spawned NPCs) entered a lecture room on the fifth floor, an earthquake early warning was issued. Several seconds later, the corresponding earthquake occurred, and the emergency time mode began.
- IV. Immediately after the shaking stopped, a fire alarm went off.
- V. The participants were allowed to behave freely. Here, some participants may evacuate to safe places (i.e., building exits), and others may not evacuate. In the building, the participants encountered different earthquake-caused situations (e.g., scattered chairs, fire, smoke, and an injured person). Figure 6 shows the layout of the prepared disaster situations in the building.
- VI. A successful evacuation was realized if the participants reached a safe spot within five minutes from Step III; otherwise, the participants failed to evacuate.
- VII. Finally, the participants answered a questionnaire comprising five-point Likert scale questions and free descriptions.



Figure 6. Building layout and disaster situations.

## 4.2 Results

#### 4.2.1 Evacuation Behavior

All participants successfully evacuated to exits on the first floor. In other words, the participants avoided going toward the top floor, and they detoured around fire and smoke. In addition, all participants encountered an injured person but did not identify any way to rescue the person from their location. However, this is a current limitation of the ETM system, which does not include interactive functions to perform rescue activities.

For Group A, participants A1–A3 communicated verbally to promote caution (e.g., "Watch out") and discussed whether to rescue the injured person (e.g., "Should we rescue the injured?"). In contrast, participants A4 and A5 did not communicate. Here, participant A6 evacuated independently without communication, except for indicating "I have to crouch down," which was communicated when the earthquake occurred.

For Group B, nearly participants evacuated silently except for words spoken when surprised (e.g., "Wow!"). When encountering the injured person, participant B1 spoke to themselves, "I cannot rescue the injured alone. I do save my own life." Note that no participant in Group B followed the NPCs reproduced from Group A's evacuation logs.

#### 4.2.2 Questionnaire

The questionnaire was administered to identify and describe the participants' emotions during the emergency time mode. Table 1 shows the mean values and standard deviations for each group's answers. For all emotions, the mean values of Group A were higher than those of Group B; however, no significant differences were observed between the two groups in the results of a t-test. In addition, only the participants in Group B were asked about the realism of the NPCs, and all participants provided neutral responses. In terms of VR sickness, we

calculated the mean values for the HMD and PC devices, which were 3.83 and 1.17, respectively. However, no significant difference was observed in the results of a t-test.

Emotion	Group A ( $N = 6$ )	Group B ( $N = 6$ )	HMD $(N = 6)$	PC ( <i>N</i> = 6)
Anxiety	4.17 (SD = 0.74)	3.50 (SD = 1.37)	N/A	N/A
Fear	4.00 (SD = 0.63)	3.83 (SD = 1.46)	N/A	N/A
Surprise	4.50 (SD = 0.83)	3.67 (SD = 1.50)	N/A	N/A
NPCs' Reality	N/A	3.0 (SD = 0.0)	N/A	N/A
VR Sickness	N/A	N/A	3.83 (SD = 1.60)	1.17 (SD = 0.41)

Table 1. Mean values of questionnaire answers (five-point Likert scale)

Question example: *Did you feel anxiety during evacuation?* 

Options: 5: Definitely yes; 4: Yes; 3: Neutral; 2: No; 1: Definitely no.

The participants in Group B provided reasons for the scores assigned in terms of the realism of the NPCs.

- Participants B1 and B6: It was realistic that NPCs were going back and forth or directing to stairs.
- Participants B1, B5, and B6: It was unrealistic that NPCs were evacuating silently.
- Participants B3 and B4: It was unrealistic that some NPCs did not start evacuating though a fire alarm went off.

The participants in both groups were asked to freely describe the lessons learned during the evacuation training. The participants identified the following lessons.

- Participants A1, A3, A6, and B1: I should remember evacuation routes.
- Participant A2: I must make some plans of evacuation routes in advance and cultivate my judgement. Because a disaster may make impassable routes.
- Participant A4: *I must calm down and behave when evacuating.*
- Participant A6: I had a sense of ease if I can evacuate with someone else.
- Participant B2: I would like to acquire knowledge of disaster management.
- Participant B3: I should think beforehand what to do when a disaster suddenly occurs.
- Participant B4: I could demonstrate evacuation since the disaster occurred in the building I usually come.
- Participant B6: In a real fire event, I will have no time to stop to consider evacuation routes since the fire can quickly spread.
- Participant B6: To prepare for real fire events, I would like to train myself to take good decisions. For example, should I call for help toward the outside when I have no idea about evacuation route? What if the fire occurs at night?

## 4.3 Considerations

The participants in Group A (i.e., multiple avatars operated by real users) felt stronger emotions than those in Group B (i.e., a single avatar operated by a real user). If evacuation realism is shown as anxiety, fear, and surprise, the avatars should not be represented by NPCs. In other words, all avatars should be operated by real users. For Group A, we found that the participant conversations may have heightened emotions more than the specific disaster situations (e.g., fire and smoke) because decision making can become complicated or difficult through conversations. For example, when discovering an injured person, the participants may encounter conflicts in terms of opposing opinions, i.e., whether or not to perform a rescue.

The mean value of the NPC realism was 3.0; however, negative reasons were relatively dominant. Participants B1 and B6 showed both positive and negative reasons. In terms of the positive reason, these participants indicated that the NPCs' movements may have contributed to increasing realism, which can be considered an advantage of the NPCs generated using the participants' evacuation logs. In addition, none of the participants in Group B followed the NPCs, and participants B3 and B4 described the corresponding negative reason. The

participants did not follow the NPCs because the NPCs were moving back and forth inside the lecture room, which delayed the evacuation. We consider that the participants in Group B felt that this behavior was unrealistic; thus, the participants were not motivated to follow the NPCs. However, in a real evacuation situation, some people may not start evacuating immediately due to cognitive biases even though they have perceived an early earthquake warning and/or fire alarm. This indicates that non-evacuating NPCs can be realistic occasionally. Thus, silent NPCs reduced realism, as indicated by participants B1, B5, and B6. We also believe that the NPCs should reproduce the participants' recorded voices to improve realism, and more evacuating NPCs should be included.

Focusing on the participants' behavior in terms of the injured person, participants evacuating individually did not attempt a rescue. The participants may have recognized that they could not have rescued the person by themselves. Note that two or three participants in Group A did attempt to rescue the injured person. The participants may have felt that the rescue was possible because they were able to discuss and cooperate. The observed situations in the experiment are expected to occur in real evacuations; thus, we consider that the ETM system should implement interactive functions to facilitate rescue activities in order to further improve realism.

Expected results were observed regarding VR sickness. Immersive HMDs can provide high levels of immersion; however, measures must be taken to mitigate the effects of VR sickness. Thus, the ETM system should be compatible with three types of devices, i.e., HMDs, PCs, and smartphones.

In each group, five participants described the lessons learned during the experiment. These lessons can be considered as the effects of the simulated evacuation experience using the ETM system. If the participants reflected on their evacuation using the ETM system's reflection function, the effects would be strengthened even further.

#### 4.4 Limitations

In this study, the number of participants was insufficient to clarify the influences of the NPCs in the ETM system. The number of generated NPCs was insufficient to influence participants' evacuation behavior and emotions. In addition, the appearance of avatars and NPCs lacked sufficient visual realism. These limitations may have been reflected in the experimental results.

## 5. Conclusion

This paper has described the ETM system with a focus on generating NPCs based on evacuation logs. The generated NPCs are expected to influence the participants' evacuation behaviors and emotions. However, a preliminary experiment revealed that the generated NPCs did not influence evacuation behaviors and emotions sufficiently compared with avatars operated by human participants. This unexpected result may have been caused by the small number of participants in the experiment. Thus, a larger-scale experiment with many participants should be conducted.

Note that the ETM system is still in the prototyping stage and must be improved from various perspectives. For example, interactive functions that simulate rescue activities should be implemented because this is a likely occurrence during real evacuations. In terms of the NPCs, the following important tasks should be addressed.

- As the number of evacuation logs increases, more NPCs must be rendered in consideration of the overall rendering load. Thus, the ETM system should determine the maximum number of rendered NPCs or implement effective conditions to select which NPCs should be rendered.
- The NPCs should be able to communicate with participants because conversation content can potentially influence evacuation behavior. It has been shown recently that generative AI technologies can realize dynamic conversations between NPCs and

players in video games, and such technology can be implemented in the ETM system to make evacuation training situations more realistic.

NPCs may need to be preprogrammed to intentionally create difficult situations where
participants cannot make decisions quickly. For example, if human participants
observe NPCs taking actions, e.g., jumping across a large crack in the floor, the
participants may be motivated to perform similar actions. We believe that the ETM
system can introduce NPCs as good examples of what not to do to facilitate effective
training.

The authors plan to complete the development of the ETM system with these considerations in mind.

#### Acknowledgements

We would like to thank Itsuki Tanioka for his great effort in this study. This work was supported by JSPS KAKENHI Grant Numbers JP18H01054 and JP23H01004.

#### References

- Cui, L., Zhu, C., Hare, R., & Tang, Y. (2023). MetaEdu: A new framework for future education. *Discover Artificial Intelligence*, *3*(1), 10. <u>https://doi.org/10.1007/s44163-023-00053-9</u>
- Feng, Z., González, V. A., Amor, R., Spearpoint, M., Thomas, J., Sacks, R., Lovreglio, R., & Cabrera-Guerrero, G. (2020). An immersive virtual reality serious game to enhance earthquake behavioral responses and post-earthquake evacuation preparedness in buildings. *Advanced Engineering Informatics*, 45, 101118. <u>https://doi.org/10.1016/j.aei.2020.101118</u>
- Feng, Z., González, V. A., Spearpoint, M., Thomas, J., Trotter, M., Cabrera-Guerrero, G., & Paes, D. (2022). A sequence analysis of behaviors in immersive virtual reality for indoor earthquake and post-earthquake evacuation. *International Journal of Disaster Risk Reduction*, 75, 102978. <u>https://doi.org/10.1016/j.ijdrr.2022.102978</u>
- Gagliardi, E., Bernardini, G., Quagliarini, E., Schumacher, M., & Calvaresi, D. (2023). Characterization and future perspectives of Virtual Reality Evacuation Drills for safe built environments: A Systematic Literature Review. *Safety Science*, *163*, 106141. <u>https://doi.org/10.1016/j.ssci.2023.106141</u>
- Liu, R., Zhu, R., Becerik-Gerber, B., Lucas, G. M., & Southers, E. G. (2023). Be prepared: How training and emergency type affect evacuation behaviour. *Journal of Computer Assisted Learning*, *39*(5), 1493–1509. <u>https://doi.org/10.1111/jcal.12812</u>
- Lorusso, P., De Iuliis, M., Marasco, S., Domaneschi, M., Cimellaro, G. P., & Villa, V. (2022). Fire emergency evacuation from a school building using an evolutionary virtual reality platform. *Buildings*, *12*(2), 223. <u>https://doi.org/10.3390/buildings12020223</u>
- Mitsuhara, H., & Shishibori, M. (2022). Toward evacuation training in metaverse. In *Methodologies and use cases on extended reality for training and education* (pp. 97–124). IGI Global. https://doi.org/10.4018/978-1-6684-3398-0.ch005
- Nagao, K. (2023). Virtual reality campuses as new educational metaverses. *IEICE Transactions on Information and Systems*, *E106.D*(2), 93–100. <u>https://doi.org/10.1587/transinf.2022ETI0001</u>
- Oe, K., Tanioka, I., Mitsuhara, H., & Shishibori, M. (2022). Prototype system of evacuation training in metaverse. In *Proceedings of the 30th International Conference on Computers in Education*, 2 (pp. 122–130).
- Sharma, S. (2020). Improving emergency response training and decision making using a collaborative virtual reality environment for building evacuation. In *Proceedings of the HCI International 2020–Late Breaking Papers: Virtual and Augmented Reality* (pp. 213–224). Springer. https://doi.org/10.1007/978-3-030-59990-4\_17
- Tlili, A., Huang, R., Shehata, B., Liu, D., Zhao, J., Metwally, A. H. S., Wang, H., Denden, M., Bozkurt, A., Lee, L. H., Beyoglu, D., Altinay, F., Sharma, R. C., Altinay, Z., Li, Z., Liu, J., Ahmad, F., Hu, Y., Salha, S., . . . Burgos, D. (2022). Is Metaverse in education a blessing or a curse: A combined content and bibliometric analysis. *Smart Learning Environments*, 9(1), 24. <u>https://doi.org/10.1186/s40561-022-00205-x</u>
- Wang, Z., Mao, Z., Li, Y., Yu, L., & Zou, L. (2023). VR-based fire evacuation in underground rail station considering staff's behaviors: Model, system development and experiment. *Virtual Reality*, 27(2), 1145–1155. <u>https://doi.org/10.1007/s10055-022-00718-3</u>

# Investigating How Technology May Negatively Affect the Academic Performance and Sleep Quality of Students

Cedric Miguel CHAN<sup>a</sup>, Josh SARTE<sup>b</sup>, Allen Peter SZE<sup>c</sup>, Jat Cedric TALAMPAS<sup>d,</sup> Judith AZCARRAGA<sup>e\*</sup>

<sup>a-e</sup> De La Salle University, Philippines \*judith.azcarraga@dlsu.edu.ph

**Abstract:** Students have been utilizing IT services such as social media, video games, and streaming services for entertainment, communication and even for coping mechanism to stress and academic workload. However, improper and heavy usage of these services can lead to bad habits and practices that can negatively affect their physical well-being particularly the quality of sleep and academic performance. It is important to investigate how to protect the students from the negative impact of excessive use of technology. This study aims to investigate how high usage rate of IT services such as social media, video games, and streaming services, can negatively affect the academic performance and sleep quality of high school students. Thirty (30) STEM high school students participated in this study over 3 different periods of academic workload. The results of the respondents and even beneficial since they are used as coping mechanism when faced with stress and heavy workload. However, most believe that such services have negative impact on their quality of sleep.

**Keywords:** Information technology, Sleep Quality, Academic Performance, Academic Workload

#### 1. Introduction

A person's sleep plays a crucial role in their overall lifestyle, with the average individual requiring 8 hours of sleep to maintain good health. Poor sleep quality can have a detrimental impact on physical health, emotions, and cognitive thinking of students. Sleep disorders, like sleep deprivation and excessive daytime sleepiness, are common among college students, impairing their attention span, cognitive abilities, and academic performance (Gaultney, 2010). Sleep and stress are closely related, with stress influencing sleep quality, and vice versa, affecting academic performance (Han et al., 2012). Poor sleep quality may lead to chronic fatigue, depression, stress, and lower quality of life (Fuller et al., 2017), impacting their ability to learn and overall health. Addressing sleep-related issues in students is vital for promoting their well-being and academic success.

Sleep deprivation can lead to emotional and behavioral issues that may affect a student's academic performance (Suni & Truong, 2023), making it crucial to analyze the factors contributing to such problems among students. The existence of rapidly developing IT platforms has driven adolescents to use these services for educational, entertainment, and social purposes, leading to both positive changes, such as improved visual reasoning and creativity, and potential negative effects on sleep quality (Rose, N. et al., 2022).

Technology has continuously evolved to meet our needs and provide entertainment. IT services, including streaming platforms, social media, and video games, have greatly enhanced our efficiency and entertainment. However, the improper use of technology remains a common issue, particularly among adolescents. During puberty, teenagers prioritize socialization, making friends, and expressing themselves, leading to extensive use of IT services that can disrupt their sleep patterns. This excessive usage can result in poor sleep quality or, in more severe cases, sleep disorders. The study of Al Kazhali et.al.(2023) suggests that social media use may be linked to poor sleep quality. Electronic social media use and its associations with insomnia, daytime sleepiness, mood, and sleep duration in adults is investigated in the work of Bhat (2018). The use of electronic social media in bed is found to be associated with sleep and mood dysfunction in adults (Bhat, 2018). According to Charmaraman et al. (2021), there is a significant relationship between the quantity of social technology use (e.g., checking social media, problematic internet behaviors, mobile use), content viewed (e.g., emotional or violent videos, risky behaviors), and social context (e.g., bedtime behaviors, starting social media at an early age) and later bedtimes and fewer hours of sleep on school nights.

While technology has undoubtedly improved our lives, it is important to investigate how to protect the students from the negative impact of prolonged use of technology for entertainment purposes. This study aims to analyze the impact of technology such as social media, video games, and streaming services on high school students' academic performance and sleep quality.

## 2. Related Works

The widespread use of technology, such as smartphones, social media, gaming, and streaming services, has influenced students' lifestyles and has the potential to negatively affect their sleep quality. The study of Fuller (2017), suggests that children who use technology around bedtime had 30 minutes to an hour less sleep than those who did not, with a substantial influence on sleep quality. Increased tablet and mobile phone screen time before bed would have a negative influence on sleep quality and attentiveness in children, resulting in morning tiredness and a lower likelihood of eating breakfast (Fuller, 2017). Additionally, a study by Olson et al. (2014) found that a significant number of adolescents brought their phones to bed, continued texting past their bedtime, and were frequently awakened by text messages during the night. These findings suggest a decline in sleep quality and quantity among adolescents who engage in electronic device use before bedtime.

Sleep is essential for memory consolidation and cognitive performance. Okano et al. (2019) investigated the association between sleep and academic performance. The study underlined the negative consequences of sleep deprivation on attention, cognition, and academic performance. Sleep deprivation has repeatedly been linked to increased fatigue, drowsiness, and poorer cognitive performance in controlled trials. Furthermore, the study found that poor sleep quality and shorter sleep duration are linked to worse levels of concentration and attention in the classroom.

## 3. Methodology

Thirty (30) STEM high school students (Grade 12) where there are 18 female, 11 male and 1 gender neutral, participated in this study. The respondents answered a survey through a Google Form. Consent forms were sent to the participants before they responded to the survey. The questions in the survey are based from the Pittsburgh Sleep Quality Index (PSQI) and The Cleveland Adolescent Sleepiness Questionnaire. Included in the questions are the IT platforms the students typically use such as social media, video games, and streaming services. Open-ended qualitative questions where the participants are asked about their thoughts on the effect of IT usage on their academic performance and sleep, are also included in the survey.

Data collection were performed in three (3) different periods of 2 school terms – end of the first term (PH), start of the second term (PL) and middle of the second term (PMH). Since PH covers the last 2 weeks of the term, most students have heavy workload due to submission of final requirements and preparation for the final exams. PL is described to be

the period with light workload since it covers the third to fourth week of the term where there are less academic requirements. Students have moderate to heavy workload during PMH which covers the fifth to sixth week of the term.

The students are divided into three(3) groups according to period. PH and PL periods were participated by eleven (11) students each while PHL was participated by 8 students.

## 4. Results and Discussions

#### 4.1 Social Media Average Usage on Different Periods

In comparing the results from each period's average daily social media usage, PH group had a mode of 4-6 hours with only a few outliers being above it. The % of responses below 4-6 hours had a combined total of 45.5%. The expected outcome was that with a heavier schoolwork, students would have a higher average daily usage of social media probably in order to communicate to each other about their group tasks. However, another counter-argument could be made that most students are busy studying individually for their final exams. These may also be due to a small sample size.

The responses from PL group on the average daily social media usage shows a lot more hours compared to the PH group. The mode remains the same being 4-6 hours. However, the % of responses below 4-6 hours had a combined total of 36.4. This means that students choose to use social media during their free time and that social media is used primarily for entertainment.

Lastly, and most interestingly, the PMH group had the highest average social media usage with the mode being 6-8 hours from 50% of the respondents. Meanwhile, 4-6 hours and 8+ hours had 25% each. This result is interesting because midterms should be a time where most students should be studying instead of communicating with others, or at least this is the "normal" assumption. One reason for this could be the small sample size.

Students with heavy and light workload tend to have the same hours of social media usage i.e. 4-6 hours. However, those with moderate to heavy workload tend to use social media longer (6 or more hours) than the other groups (4-6 hours).

#### 4.2 Gaming and Streaming Services Usage to Accumulated Grade Average and Sleep Quality

Table 1 shows that 63.3% of students (19 out of 30) spend only 0-2 hours of gaming per day and of which 5 (26.3%) had poor sleep quality i.e. 1-2, 8 (42.1%) had sleep quality of 3 and 6 (31.6%) had good sleep quality i.e. 4-5. Moreover, there is no correlation at all with a value of 0.04 between sleep quality and number of hours of gaming per day.

		Sle					
		1	2	3	4	5	Total
Average number of hours for gaming use per day	0-2 hrs	1	4	8	6	0	19
	2-4 hrs	2	0	2	1	1	6
	more than 4 hrs	0	0	3	1	0	4

Table 1. Frequency of respondents based on sleep quality and hours of gaming use

Table 2 shows that 70% of students spend only 0-4 hours of streaming per day and of which 5 (4.8%) had poor sleep quality i.e. 1-2, 8 (38.1%) had sleep quality of 3 and 8 (38.1%) had good sleep quality i.e. 4-5. Moreover, there is a negative correlation of -0.41 between the number of hours used of video streaming and the quality of sleep.

		Sleep Quality (1-worst, 5 -best)					
		1	2	3	4	5	Total
Average number of hours for streaming use per day	0-2 hrs	0	1	3	4	1	9
	2-4 hrs	2	2	5	3	0	12
	more than 4 hrs	1	1	6	1	0	9

One participant, who reported to have spent 8+ hours on streaming services and another 8+ hours on social media per day, had only 4 hours of sleep at night with a very low sleep quality of 1. However, this student performs well in school with an accumulated grade average of 3.3 where 4.0 is the highest average.

Other respondents who reported to have 4+ hours on both gaming and streaming services showed an average of 2.9, showing that healthy sleep patterns do help with maintaining a higher grade average.

Meanwhile, another participant, reported spending 0-2 hours on both gaming and streaming services. The student spend on an average of 6-7 hours of sleep and an overall sleep quality of 4. The grade average also reported to be high (3.8). As a matter of fact, respondents who answered 0-2 hours on both gaming and streaming services, had an average of 3.18 which is on the higher end of the group in terms of academic performance.

#### 4.3 Sleep Quality and Social Media Use

Table 3 show that 70 % of the respondents or 21 out of 30, use social media ranging for more than 6 hours on the average per day. Of these respondents, 6 (28.6%) had poor sleep quality i.e. 1-2, 11 (52.4%) had sleep quality of 3 and only 4 (19%) had good sleep quality i.e. 4-5.

There is a negative correlation of -0.44 between the number of hours used of social media and the quality of sleep.

		SI					
		1	2	3	4	5	Total
Average number of hours for	0-2 hrs	0	0	1	2	0	3
social media use per day	2-4 hrs	0	1	2	3	0	6
	4-6 hrs	1	0	5	3	1	10
	more than 6 hrs	3	3	11	3	1	21

Table 3. Frequency of respondents based on sleep quality and hours of social media use

#### 4.4 Effect on Academic Performance and Coping Mechanism

The responses from the survey revealed that twenty-one (21) or 70% of the respondents believe that IT services are beneficial for their academic performance, as they provide access to online tools and digital resources. Such services are found to be helpful in finding information, improving task efficiency, and aiding in academic requirements, as well as serving as a source of motivation and stress relief. On the other hand, nine (9) respondents disagreed, stating that IT services had little positive impact on their academic performance. This group believe that IT services distracting, leading to a loss of focus on academic tasks and procrastination, ultimately hindering their performance.

The responses regarding the impact of IT services on sleep quality were diverse, with eight (8) or 26.67% of the respondents find them beneficial in aiding sleep, nineteen (19) or 63.3% stating negative effects due to addiction and staying awake, and three (3) or 10% expressing a neutral view. Some respondents acknowledged the positive aspects, noting

that technology helps them relax before bedtime and occasionally aids in falling asleep, improving sleep quality. However, many respondents raised concerns about the negative impact, citing addictive behaviors and exposure to disturbing content on social media as factors that disrupt their sleep. Despite recognizing the drawbacks, some found it difficult to reduce their reliance on IT services due to habit and accessibility. Conversely, a few respondents reported no significant impact on sleep quality, suggesting that technology played a relatively neutral role in their sleep patterns.

Based on the responses, 73.33% (22 of 30) of respondents view IT services and gadgets as a coping mechanism when facing sleep difficulty, using browsing, streaming services, and social media to relieve stress.

While technology can help some individuals relax and fall asleep, it can also exacerbate sleeping difficulties for others, leading to increased screen time and stimulation. Factors contributing to sleeping difficulties include academic stress, unbalanced sleep schedules, and anxieties related to schoolwork and deadlines. Further analysis reveals interesting patterns in students' perceptions of technology as a coping mechanism for sleep difficulties in relation to the timing of their responses.

During the PH period, 54.55% (5 of 11) responses do not consider technology as a coping mechanism likely due to the heavy workload and pressure. In contrast, during the PL period, 81.82% (9 of 11) responses considered technology as a coping mechanism maybe due \lighter workload and reduced stress. Although the percentage during midterms week in PMH period is higher, with 87.5% (7 of 8) responses indicating technology as a coping mechanism. The data suggests that students' perception of technology as a coping mechanism for sleep difficulties varies depending on the period being more beneficial during less intense periods and non-beneficial during stressful times.

## 5. Conclusions and Future Work

Information technology services are beneficial to students not only on their academic requirements but also in their general mental well-being. However, excessive usage of technology may negatively affect their academic performance, sleep quality and mental health. This could mean that technology may not always be safe to use by the learners especially when they cause distraction to sleep and technology addiction. This study aims to investigate how high usage rate of IT services such as social media, video games, and streaming services, can negatively affect the academic performance and sleep quality of high school students. Thirty (30) STEM high school students participated in this study over 3 different periods of academic workload. 63.3% believe that IT services are beneficial to their academic performance. 73% find such services as their coping mechanism during stress and heavy academic workload. Students' perception of technology as a coping mechanism for sleep difficulties varies depending on the period being more beneficial during less intense periods and non-beneficial during stressful times. 70% believe that IT services have negative impact on the quality of their sleep. However, there is small correlation between the number of hours used of social media and the quality of sleep, between the number of hours used of video streaming and the quality of sleep. There is no correlation at all with between sleep guality and number of hours of gaming per day. The results of this study show that IT services have no negative impact on the academic performance of the respondents and even beneficial since they are used as coping mechanism when faced with stress and heavy workload. However, most believe that they have negative impact on their quality of sleep. Future work may involve more respondents and more qualitative questions.

#### References

Alhadlaq, I. (2016). How Technology Influences Communication. *International Journal of Scientific & Engineering Research*, 7 (1), 1

- Al Kazhali, M., Shahwan, M., Hassan, N., & Jairoun, A. A. (2023). Social media use is linked to poor sleep quality: The opportunities and challenges to support evidence-informed policymaking in the UAE. *Journal of public health* (Oxford, England), 45(1), 124–133.
- Bhat, S., Pinto-Zipp, G., Upadhyay, H., & Polos, P. G. (2018). "To sleep, perchance to tweet": in-bed electronic social media use and its associations with insomnia, daytime sleepiness, mood, and sleep duration in adults. *Sleep health*, 4(2), 166–173.
- Charmaraman, L., Richer, A. M., Ben-Joseph, E. P., & Klerman, E. B. (2021). Quantity, Content, and Context Matter: Associations Among Social Technology Use and Sleep Habits in Early Adolescents. *The Journal of adolescent health* : official publication of the Society for Adolescent Medicine, 69(1), 162–165.
- Fuller, Caitlyn & Lehman, Eric & Hicks, Steven & Novick, Marsha. (2017). Bedtime Use of Technology and Associated Sleep Problems in Children. *Global Pediatric Health*. 4
- Gaultney, J. F. (2010). The prevalence of sleep disorders in college students: impact on academic performance. *Journal of American College Health*, 59(2), 91-97.
- Han, K. S., Kim, L., & Shim, I. (2012). Stress and sleep disorder. Experimental neurobiology, 21(4), 141–150.
- Moulin, K. (2015, March). Technology trumping sleep: Impact of electronic media and sleep in late adolescent students. In Society for Information Technology & Teacher Education International Conference (pp. 999-1001). Association for the Advancement of Computing in Education (AACE).
- Okano, K., Kaczmarzyk, J., Dave, N., Gabrieli, J. D. E., & Grossman, J. C. (2019). Sleep quality, duration, and consistency are associated with better academic performance in college students. *Npj Science of Learning*, *4*(1).
- Rose, N. N., Ishak, A. S., Sultan, N. H. H., Ismail, F., & Fahrudin, A. (2022). Effect of digital technology on adolescents. In Impact and Role of Digital Technologies in Adolescent Lives (pp. 1-18). IGI Global

# W05

# 3rd International Workshop on Embodied Learning: Technology Design, Analytics & Practices

## Embodied Learning of Integer Operations Using a Multitouch Design: Touchy Pinchy Integers

#### Priyadharshni ELANGAIVENDAN\*, Ashwin RAMASWAMY, Melwina ALBUQUERQUE & Sanjay CHANDRASEKHARAN

Learning Sciences Research Group, Homi Bhabha Centre for Science Education, Tata Institute of Fundamental Research, India \*priyadharshni@hbcse.tifr.res.in

**Abstract:** We report the design and pilot testing of a digital application to help middle school children learn integers, Touchy-Pinchy Integers (TPI). The design rationale for this touchscreen-based learning system was based on a neutralization model of integers. To design this embodied mathematics learning system, we associated commonly used touchscreen interaction gestures with mathematical meaning related to integers. Users could manipulate two collinear towers of stacked positive and negative blocks, which together represent an integer expression. The learner could add either type of block, by tapping on the appropriate part of the screen. They could also subtract a desired number of either type of block, by slicing across the blocks. They could also neutralize a pair of positive and negative blocks, by pinching in. The pinch-out actions generated a new pair of positive and negative blocks. A flipping gesture reversed the polarity of the integer expression, which corresponded to a multiplication by -1. Results from a pilot study with grade 7s students indicated that the application helped students resolve deep conceptual issues related to negative numbers, through embodied learning. We present some of our observations from the study and interpret these episodes using embodied learning theory.

**Keywords:** Integer Operations, Touchscreen Interactions, Tangible Gestures, Embodied Mathematics Learning, Representational Restructuration

#### 1. Introduction

Students around the world are usually introduced to integers and negative numbers in or around middle school. This is a challenging transition for students, as the idea of negative numbers is often a novel revelation, and requires children to drastically reinterpret familiar operations such as addition and subtraction. For instance, a child may have to unlearn previously useful generalizations, such as "subtracting makes the number smaller" or "you cannot subtract a bigger number from a smaller number" and replace these ideas with a richer, more nuanced understanding of the same operations.

Typically, issues arise with the learning of integers because of misconceptions in understanding negative numbers, or operations with negative numbers. (Fuadiah et al., 2016). While some students have an intuition for integer operations, and the idea of positives and negatives canceling out (Whitacre et al., 2012), some learners develop, and carry to higher grades, fundamental misconceptions about operations with negative numbers (Makonye & Fakude, 2016). It is not surprising that children find negative numbers challenging, as past mathematicians have struggled with epistemological issues in the process of integrating negative numbers to the number system (Whitacre et al., 2012).

One proposed explanation (for why negative numbers are challenging to understand) suggests that learners find it difficult to imagine a single perfect real-world analogy, or an intuitive model that gives readily interpretable meaning to negative numbers, as well as the possible operations on them. For example, while the number line is a very pedagogically

useful model, with high generality that extends up to real numbers, it does not support an intuitive explanation for what it means to subtract a negative number (Liebeck, 1990).

This difficulty could possibly be addressed using physical or virtual manipulatives or models, as these can act as concrete objects that follow predictable rules (Murray, 2018). Also, as argued by Wilensky & Papert (2010), new computational media allows radical 'restructurations' of existing ways to learn and understand difficult concepts. Mathematics educators use a class of intuitive models for integers, to help alleviate some of the issues with the subtraction of negative numbers. This approach involves establishing a conceptual metaphor that appeals to 'neutralization' or 'canceling out' of constituent positive units of an integer with their additive inverses (Flores, 2008; Murray, 2018).

Combining and extending these two model-based reasoning possibilities, as well as the representational restructuration view, we designed a digital learning application (Touchy Pinchy Integers, TPI) based on a neutralization model. The TPI system provides virtual manipulatives to learn integers. We outline some results from a preliminary study of the system with grade 7 students, who were asked to use the TPI system as a learning aid.

#### 2. Theoretical Background

Recent work in embodied cognition and human-computer interaction has outlined a number of theoretical approaches towards the design of novel learning interfaces. We outline some of these below.

#### 2.1 Re-imagining Human-Computer Interaction as Embodied Interaction

New computational media offers new ways of interacting with technology, overcoming some of the limitations of traditional interfaces such as keyboard-mouse input systems Touchscreens, with their ability to recognize and respond to sophisticated hand gestures such as pinching and dragging, are a good example of modern interaction technologies that engage users on a richer level, using the actions of the human body.

While touchscreens are almost ubiquitous today, more advanced interaction technology prototypes often leverage more aspects of the body to re-imagine the input-output interface. Tighter integration of the interface with the human body is a common thread in such technology designs, such as in the use of the human skin as a touch-responsive computer screen (Harrison et al., 2010). Authors such as (Wu & Hsu, 2011) have proposed methods to use human movement to help visualize large collections of information in 3 dimensional space.

One of the primary aims of developing novel interaction technology is to reduce the cognitive load experienced by the user. A possible explanation for the rapid rise of touchscreen devices is that interaction design with state-of-the-art touch-based interfaces is more intuitive and "natural" from a biological perspective, when compared to pointing and clicking using computer mice. A recent evolutionary view of Cognitive Load Theory suggests that cognitive load is significantly higher when learning tasks or information that are less optimized, in relation to the brain's evolutionary history (Paas & Sweller, 2012). From this perspective, the naturalness of common gesture-response combinations – such as drag-to-move or pinch-to-shrink – would alleviate some of the cognitive load associated with digital interactions.

Vygotsky's concept of tool-mediated cognition (Turner, 2016, p. 27–40) provides another perspective to understand the potentially profound consequences of some of the emerging technologies on human cognitive abilities. Keeping in mind Vygotsky's key assertion that tools interact bi-directionally with the human brain -- in other words our cognitive structures adapt to, and eventually mimic, the specific tools that we use to solve problems -- it is possible to argue that new computational media can significantly expand the scope of tool-mediated cognition, as it is now possible to quickly develop artifacts whose input-output relationships are not fully constrained by physical laws.

#### 2.2 Tool incorporation and extending it to screen spaces

Recent cognitive neuroscience studies on tool-use show that active and intentional use of tools lead to changes in the neural representation of the body (the body schema). Specifically, studies with macaque monkeys (Maravita, Iriki, 2004) show that tools are 'incorporated' into the body schema, as the brain treats the tool as part of the body', expanding the action space of effectors and the participant's peripersonal space ('actionable' space close to the body, in this case space close to the hand). These studies have been extended to humans as well (Farne et al, 2005). This study reports that the peri-hand space after intentional tool usage expands to include not just the tip of the tool, but also includes the peri-tool-space, which includes the functional space of the tool.

Studies have shown that tool incorporation occurs even when the tool is connected to a virtual space. Macaque monkeys were trained to retrieve food by manipulating a tool and observing the hand movements through a real time video monitor, and not directly looking at the hand. The monkeys were able to use this video as a tool for food retrieval (Maravita, Iriki, 2004), and the neural body schema changed similarly to physical tool use. Further, studies have shown that human users' peripersonal space is extended onto the screen, and the cursor is treated as a tool (Gozli, Brown, 2011). In these studies, there are two worlds, the physical world and the screen world. Changes in the real world continuously cause corresponding changes on the screen, with resulting changes in the body schema and the peri-personal (actionable) space of the body. This systematic neural relationship between the body schema and virtual tool use has become integrated into our everyday lives, with regular use of computers and touch based devices.

The theoretical framework of tool incorporation has recently been used to account for the development of computational thinking in kindergarten, through learners' active manipulation of a robotic toy (Sinha. R et al, 2023). Tool incorporation has also been used to account for the way students' actions on physical manipulatives lead to the learning of mathematical ideas such as area (Rahaman. J et al, 2018). The incorporation account could be further extended, to account for the way abacus users generate a mental abacus, and the resulting cognitive and neural changes, such as the activation of visuo-motor areas and related gestures while verbally solving arithmetic problems (Hanakawa, T. et al, 2003).

#### 2.3 The Role of External Representations in Cognition and Learning

Supporting the incorporation studies that show that tools have extensive cognitive effects, Kirsh (2010) has shown theoretically that external representations (ERs) – symbol-based tools used extensively during thinking, learning, and computing tasks – make cognitive contributions that are wider than optimizing internal memory load. For instance, ERs make learning processes interactive, by actively providing a stable external anchor, and also hints and prompts, which transform the reader's thoughts and ideas. This role of ERs allows users to think by altering the external representation, to arrive at a new set of insights. As ERs work as a tool that supports thinking, they transform the thinking task, by allowing building over, changing, removing or manipulating the tool (Kirsh, 2010).

Another characteristic of ERs – which also has a pedagogical benefit – is that external representations serve as a shareable object of thought. When the learner represents or works with the ERs, educators are able to observe the thinking and learning processes, which are otherwise not available. Given this structure, the educator is able to participate in the meaning-making process of the student, and intervene at the right moment to guide the learners' thinking. Abrahamson and Garcia (2016) discusses a similar idea – distributed coenactment – while examining the pedagogical outputs during the use of an embodied mathematics learning tool (The Mathematical Imagery Trainer).

Mathematics uses ERs extensively, in terms of numbers, symbols and diagrams, to capture the essence of the concepts and also lock-in the key ideas. These symbols once written down don't decay over time like mental structures. This available structure helps learners to build on their ideas. Even though ERs are symbol-based, they could be understood as functioning in ways similar to tools. It is possible that ERs are also

incorporated into the body schema, as indicated by empirical studies on the cognitive and neural effects of extensive abacus use (Hanakawa, T. et al, 2003).

#### 2.3 Embodied Learning of Mathematics

Embodied cognition theory emphasizes that the mind, body, and the environment are inextricably connected. Consequently, the body plays a significant role in thinking and acting in the world (Kosmas, Zaphiris 2018). Abrahamson and Garcia (2016) demonstrated through their instruction based embodied interaction design - Mathematical Imagery Trainer for proportion (MIT) – that learners can assimilate a mathematical concept through goal oriented hand movements. Through these actions, later reinforced using *symbolic artifacts* like cursors, grid lines and numbers on the screen, learners were able to learn the mathematical idea of proportion, making use of the movements and symbolic elements as both frames of action and reference. The symbolic artifacts allowed learners' engagement with the given tasks to be shifted closer to mathematical reasoning and visualization.

In a related thread, Sinclair (2014) argues that touch technologies allow the hand's actions to have immediate and unmediated impact, as changes in the screen are directly related to the intentional touch. She developed a novel interactive system based on multi touch - Touch counts - to develop and support basic numeracy in young children, including counting, addition and subtraction. In this system, touch media gestures – like tapping, pinching, flicking etc. – are given mathematical meanings. This structure allows learners to explore the nature of numbers, while interacting with the touch media using gestures and receiving real time feedback. Touchy feely vectors (Karnam, D et al, 2018) is another touch-based design to learn mathematics through embodied interactions. This system allows learners to learn vector concepts like addition and resolution, through the active manipulation of vectors directly using touch gestures.

Extending these theoretical ideas on embodied cognition and successful applications that help students learn mathematics in an embodied way, we have designed a system – Touchy Pinchy Integers (TPI) – to help students explore the concept of integers through embodied interaction. Apart from the embodied interaction, having TPI as a dynamic external representation reduces the cognitive load of learners, by helping them visualize the operations, and reducing their reliance on memorized rules of integer operations. TPI can be used as a tool during problem solving sessions related to integer operations. After some experience working with TPI, the system also works as an imagination tool, helping students visualize integer operations, especially when solving problems using static media, such as text and figures. This use is similar to the way experienced abacus users generate and manipulate a 'mental abacus', when trying to verbally solve arithmetic problems (Hanakawa, T. et al, 2003).

#### 3. Methods

#### 3.1 Learning System Design

The TPI application is presented as a single screen, with a horizontal line dividing the top and bottom halves. On either side of the line, there is a positive and negative tower, each consisting of red and blue stacked blocks respectively. Each block represents a unit of the appropriate sign, and all the blocks additively represent the integer expression. Thus, 3 red and 2 blue blocks would stand in for the expression 3 + (-2).

The user can manipulate the number of blocks in each tower, using a number of touch gestures, each of which is mapped to an operation on the integer expression. Performing a gesture triggers an animation, which can result in a change of state (i.e. change in the number of blocks in either or both towers). On the screen at all times is the expression m + (-n), in which m and n represent the current number of positive and negative blocks respectively. The system is available for interaction at the following link:

#### https://integercountsdevelopment.melwinaalbuque1.repl.co/

The following touch gestures lead to operations on the expression:

- *Tapping*: Tapping either the positive or negative side of the screen with k fingers at a time leads to addition of k blocks on the same side of the screen. The expression is incremented by +k or -k depending on the side of tapping. See Figure 1 (i) for a concrete case.
- *Slicing:* Swiping vertically across the top face of the i<sup>th</sup> positive block removes the base-most i blocks from the positive tower and subtracts i from the expression. Slicing across the top face of the i<sup>th</sup> negative block removes the base-most i blocks from the negative tower and subtracts -i from the expression. See Figure 1 (ii).
- *Flipping*: Swiping vertically across the division line causes the red and blue towers to swap colors and polarity on the screen. The overall expression is multiplied by -1. See Figure 1 (iii).



*Figure 1 (i).* Addition by tapping. Left: Adding +3 by tapping the top half with three fingers. Right: Adding -4 by tapping the bottom half with four fingers. Brown dots represent fingertip positions and are only for visualization - they do not appear to the user. *(ii).* Subtraction by slicing. Left: State before subtraction (subtraction animation has just begun). Right: State after neutralization. Gesture details shown for visualization (not visible to user): brown dot represents swipe start position, purple dot represents swipe end position, and dotted line represents linearized downsampled swipe trajectory used for computation. *(iii)*Flipping. Left: State before flipping. Right: State after flipping.

- *Pinching In*: Pinching in with one finger each on the positive and negative sides of the screen neutralizes one pair of blocks from each side assuming there is such a pair. The overall value of the expression is conserved, but both the positive and negative terms have their magnitudes decreased by 1. See Figure 2 (i).
- *Pinching Out*: Pinching out with one finger on each on the positive and negative sides of the screen generates a one positive and one negative block. The overall value of the expression is conserved, but both the positive and negative terms have their magnitudes increased by 1. See Figure 2 (ii).



*Figure 2 (i).* Neutralization by pinching in. Left: State before neutralization (pinching animation has just begun). Right: State after neutralization.

*(ii).* Reverse Neutralization by pinching out. Left: State before reverse-neutralization (pinching animation has just begun). Right: State after reverse-neutralization.

#### 3.2 Pilot Study Design

A pilot study using the TPI system was conducted with 5 grade 7 students in a residential school in Maharashtra, India. TPI was given to them for 15-20 minutes on a daily basis, for 1 week. The students were selected by their mathematics teacher based on their performance in their math classroom. All students were from different performance backgrounds. They were already introduced to integers and integer operations in class 6. The sessions started with a pretest, where the students were asked to solve integer-based addition and subtraction problems. At the end of the session the students were asked to explain how and why they got their answer. After the pretest, students were introduced to TPI using an ipad. They were then asked to practice the basic gestures, resulting in changes in the screen.

Following this phase, students engaged in a series of gradually advancing tasks, for 5 to 6 days. The tasks involved various operations, including: adding numbers with the same sign and opposite signs, subtracting a small number from a big number with the same sign (on day 1); subtracting an integer from zero, subtracting a larger number from a smaller one with the same sign, subtracting integers with opposite signs (on day 2); multiplying a given integer by (-1), multiplying an arithmetic expression by (-1), performing integer operations using TPI for a written expression (on day 3); exploring the triple nature of the minus sign through TPI (on day 4). In each session, the students practiced tasks from previous sessions. Day 5 was dedicated to practicing all the tasks learned using the TPI system. The students also performed integer operations for written expressions on this day. Since this was an initial study, the tasks and instruction were adjusted based on students' interactions and difficulties. For instance, the concept of screen value and multiple representations [like 2 = 3 + (-1) = 4 + (-2)] emerged during a session with one of the students.

The final stage was a post-test, where students were encouraged to perform arithmetic operations using an imaginary simulation of TPI. The questions in the post-test were similar to those in the pre-test, but this time students were immediately asked to explain their reasoning after each question. This change was made because we noticed that students had difficulty recalling their reasons when asked at the end of the test. Data collected during the study were: audio of students explaining their reasons for the integer operation (during pre-test and post-test), images of answer sheets or rough sheets created by students during pre and post test, videos of students' usage of TPI with aerial view of the software, field notes.

#### 4. Results

After the pretest, when students were asked to support their answers with reasons, all children provided the rules they learnt in grade 6 as a reason. A few children were able to give reasons using the number line. When two negative signs were next to each other, every child instantly used the rule 'minus into minus plus', and that remained the reason for their answer. After the post-test analysis, we found that students made use of the mental image of TPI and gestures, similar to abacus experts relying on the mental image of the abacus and gestures while solving verbal problems. Among the 5, three used pinching (pinching in for addition of opposite integers) gestures, and one child used the tapping gesture, when asked to add integers of the same sign.

Two other students, who were comfortable with the rules and got all the answers right during the pre-test, used the rules again during the post-test. When asked why they chose to work and reason in terms of rules even though they worked with TPI, they mentioned that they have been practicing a long time with rules. Even while using rules to solve problems, some children used TPI to confirm their answers, using the system as an additional tool. This allowed them to explain their answers more clearly and with greater confidence, compared to the pretest. The TPI system allowed students to avoid blindly applying rules and to prevent the misuse of rules in inappropriate situations.

#### 5. Discussion: Embodied Learning of Integers Enabled By TPI

A number of interesting embodied learning episodes were seen during students' interactions with the TPI and later. We discuss these here, based on the theoretical frameworks outlined in section 2.

• Thinking outside with the new representation

External representations allow us to execute some cognitive processes outside. These processes cannot be executed 'inside', i.e. with just our minds (Kirsh, 2010). In the case of the interactive external representation created by the TPI system, the value 3 can be decomposed into three units, and split across three fingers and when tapped on the screen. This leads to three boxes summoned on the screen, as 1+1+1. This representation of individualized units allows learners to appreciate the neutralization process when a positive and negative unit come together (added). Similar to other neutralization models used in traditional mathematics classrooms (two color button model, token model, or card model), TPI also provides a visual element for the neutralization process. However, in the other models, the user has to manually remove the two coloured units for the neutralization operation. Even after removal, the units are still kept aside somewhere in the real world. But in TPI, pinching-in positive and negative units with two fingers, the two units are animated in such a way that they actually come together to neutralize and vanish into thin air, to generate the value zero. Children were surprised to see this animation of addition of opposite integers. A child noted that during this process, when the boxes on either side are not the same, one kind of box runs out. This process allowed students to observe the reason for one of the rules – to add integers with opposite signs, find the difference and put the sign of a big number. Sign of the bigger number is the same as the kind of boxes that remain on the screen after the pinching-in process. In this way, the TPI external representation, based on interactive media, allowed students to think about things that were unthinkable without the new media system (Kirsh, 2010).

Similarly, with the button or token model, the reverse neutralization operation is done by manually introducing a pair of opposite coloured units. But in TPI, pinching-out with two fingers from the central horizontal line leads to two opposite units resurrecting on either side. This illustrates the decomposition of 0 into +1 and - 1, which provides an implicit avenue to understand a rather deep idea – zero can now be seen as potentiality, rather than emptiness.

Students mentioned that TPI allowed them to think and manipulate with the boxes present outside in the system, instead of relying on rules. This allowed them to not worry about forgetting rules, and not fear the use of rules in the wrong context. Students also mentioned that while they were using the cards for integer operation, the process became messy and confusing, particularly for subtraction. TPI was easier to work with, and made the operation easy to understand.

• Thinking clearly with varied touch screen gestures

The TPI system assigns mathematical meanings to gestures used on touch screen. This structure allows children to explore mathematical concepts easily, by manipulating and controlling the TPI system through simple touch screen gestures. The screen split into two halves, and operations in the two halves, gave students an embodied understanding that positive units belong to the upper half and negative units belong to the lower half. Based on this embodied understanding, students moved their hands to the appropriate half of the screen quickly during tasks and explanations, while mentioning positive or negative units. This sense of direction helped reduce mistakes students make while adding integers of the same sign or different sign. Subtraction in integers can be viewed as a change in value, movement on a number line, or taking away some units from a given unit. TPI follows the taking away idea. There are 3 different functions of minus sign when it comes to integers - unary, binary, symmetry (Vlassis, 2004). Unary defines a negative number -2, binary refers to the binary operation of subtraction 5 - 2, symmetry corresponds to the resultant additive inverse - (-2). Children seem to have internalized the concepts of unary and binary, and were able to use the appropriate gestures when they were asked to demonstrate -2 or 5-2 using TPI. They tapped 2 fingers on the lower bottom to represent the unary function, they tapped 5 and sliced off 2 boxes for the subtraction operation.

The function of symmetry required facilitators demonstrating the flipping gesture, and discussion about this operation. Students observed that flipping got them the additive inverse of the number that was flipped. With this observation, students were able to predict the answer when an expression with positive and negative numbers (units on both halves of the screen) was flipped. Students were also able to observe that on flipping both the numbers were replaced by their respective additive inverse. This allowed them to understand that although the sign '-' is the same, the function it possesses depends on the position in the expression.

The change in the expression caused by various minus signs is thus captured with the right nuances in TPI, with appropriate gestures for unary, binary and symmetric function. A question like -((-5) - (-2)) will require the student to use 3 different gestures for the similar looking minus sign in the expression.

#### • Thinking with a new concept – screen value

The concept of screen value came up during one of the sessions as a key idea, in the context of subtraction of big numbers from small numbers and subtracting positive numbers from negative numbers or vice versa. The idea of screen value is similar to the idea of net worth in the credit-debt model of integers, where the sum of credit and debit is the net worth. Screen value is the value one would obtain on pinching-in boxes from either side till there is only one color left on the screen. For a question like 3 - 5, the student could take away 5 from a screen value of 3 i.e, take 5 from 5+(-2) which results in -2. In this process the students pinches out (reverse neutralization) to keep the screen value constant. The idea of reverse neutralization, generating a pair of +1 and -1, helped students with the operations of integers. This idea is not used in conventional classrooms for such problems. Since pinching in and pinching out are core to the TPI system, after using these operations, students start thinking in terms of a new concept – the screen value.

#### • Transformation in thinking with TPI, and incorporation

The intervention tasks had three stages, spread across 5 -6 days. Students practiced the previously learned tasks during this period, and their thinking with TPI got refined across these stages. The first stage was performing operations using the TPI system. The second stage involved predicting outcomes by mentally simulating the system, and then confirming the answers by performing the operations on the TPI system. In the third stage, during the post test, students had to mentally simulate TPI to solve the given integer operation. This was observed through students' gestures of tapping and pinching in the space in front of them, while solving the problems given in the post test. As students used gestures even when there was no system in front of them, the children were using TPI as an imagination tool for integer operation, similar to expert abacus users' use of a mental abacus. This use of TPI as an imagination tool suggests that TPI was incorporated into students' body schema, through their actions on the touch interface.

#### 6. Redesigning TPI

Apart from the above embodied learning patterns, we also found some usage patterns that could be helpful in redesigning TPI, for better learning and usability. We outline two classes of patterns below.

• *Redesigning TPI to think mathematically* 

Abrahamson and Garcia (2016) discusses how symbolic artifacts like grid and numerals allowed students to bring their action-based engagement and manipulation strategies closer to mathematical reasoning. In the TPI system, the central horizontal line allows students to use it as a frame of action for pinching-in, pinching-out and tapping, to obtain positive and negative numbers. In addition to this structure, TPI also has an expression on the screen m+ (-n), for quick reference to the number of boxes present on either side of the horizontal line. We observed students counting slowly while they are expected to subtract a value. We are planning to include grid lines, or parallel horizontal lines with corresponding integers above and below the existing central horizontal line. This structure will allow the students to identify the boxes without counting them.

#### • Acclimatizing to the TPI environment

Almost all the students initially found the pinch-out and pinch-in gestures difficult. With regular practice with the TPI system, the child was able to figure out how to make a particular gesture work with the system. The teacher could only demonstrate how it is done, but to actually make a particular gesture work, practice was needed. The practice helped students adapt the system to their organismic idiosyncrasies, such as the length of the hand, suppleness of the fingers, or the pressure and distance between the fingers the student uses over the screen (Abrahamson and Garcia, 2016).

Students also had difficulties with the slicing gestures. While subtracting 2 from 5 using the TPI system, one needs to slice off the second box from the center line. But a few students naturally moved their fingers towards the tip of the tower, and sliced the second box from the top. This gesture has a rationale: if we go according to the cartesian plane, then the second box is counted from the reference line. And irrespective of the length of towers, if 2 units have to be subtracted, users would prefer the slice to happen at the same place. By slicing the second box from the center learners preserve both these features.

Gesture usage in TPI had many other behavioral patterns. Even after repeated instructions about the possibility of using two hands for operating TPI, some students preferred to use one hand. The way students decomposed a particular number was also interesting: some students use 3+3, 4+4 for 6 and 8 while some use 5+1, 5+3 for the same.

We plan to further study these and other similar behavioral patterns, and explore their potential for designing new pedagogical elements and tasks, which could improve the use and effectiveness of TPI.

#### Acknowledgements

We extend our gratitude to the study participants and their teachers. We also thank Pranshi Upadhyay for valuable inputs and technical support during the pilot study.

#### References

- Abrahamson, D., & Sánchez-García, R. (2016). Learning is moving in new ways: The ecological dynamics of mathematics education. *Journal of the Learning Sciences*, 25(2), 203-239.
- Farnè, A., Iriki, A., & Làdavas, E. (2005). Shaping multisensory action–space with tools: evidence from patients with cross-modal extinction. *Neuropsychologia*, *43*(2), 238-248.
- Flores, A. (2008). Subtraction of positive and negative numbers: The difference and completion approaches with chips. *Mathematics Teaching in the Middle School*, *14*(1), 21–23.
- Fuadiah, N. F., Suryadi, D., & Turmudi, T. (2016). Some difficulties in understanding negative numbers faced by students: A qualitative study applied at secondary schools in Indonesia. *International Education Studies*, *10*(1), 24.
- Gozli, D. G., & Brown, L. E. (2011). Agency and control for the integration of a virtual tool into the peripersonal space. *Perception*, *40*(11), 1309-1319.
- Hanakawa, T., Honda, M., Okada, T., Fukuyama, H., & Shibasaki, H. (2003). Neural
correlates underlying mental calculation in abacus experts: a functional magnetic resonance imaging study. *Neuroimage*, *19*(2), 296-307.

- Harrison, C., Tan, D., & Morris, D. (2010). Skinput: appropriating the body as an input surface. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 453–462.
- Karnam, D., Agrawal, H., Chandrasekharan, S., Yang, J. C., Chang, M., Wong, L. H., & Rodrigo, M. M. T. (2018). 'Touchy Feely Vectors' changes students' understanding and modes of reasoning. In *Proceedings of the 26th International Conference on Computers in Education* (pp. 143-152). Philippines: Asia-Pacific Society for Computers in Education.
- Kirsh, D. (2010). Thinking with external representations. AI & society, 25, 441-454.
- Kosmas, P., & Zaphiris, P. (2018). Embodied cognition and its implications in education: An overview of recent literature. *International Journal of Educational and Pedagogical Sciences*, *12*(7), 970-976.
- Liebeck, P. (1990). Scores and Forfeits: An Intuitive Model for Integer Arithmetic. *Educational Studies in Mathematics*, *21*(3), 221–239.
- Makonye, J. P., & Fakude, J. (2016). A Study of Errors and Misconceptions in the Learning of Addition and Subtraction of Directed Numbers in Grade 8. SAGE Open, 6(4), 2158244016671375.
- Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in cognitive sciences*, *8*(2), 79-86.
- Murray, E. (2018). Using Models and Representations: Exploring the Chip Model for Integer Subtraction. In L. Bofferding & N. M. Wessman-Enzinger (Eds.), *Exploring the Integer Addition and Subtraction Landscape: Perspectives on Integer Thinking* (pp. 231–255). Springer International Publishing.
- Paas, F., & Sweller, J. (2012). An Evolutionary Upgrade of Cognitive Load Theory: Using the Human Motor System and Collaboration to Support the Learning of Complex Cognitive Tasks. *Educational Psychology Review*, *24*(1), 27–45.
- Rahaman, J., Agrawal, H., Srivastava, N., & Chandrasekharan, S. (2018). Recombinant enaction: Manipulatives generate new procedures in the imagination, by extending and recombining action spaces. *Cognitive science*, *42*(2), 370-415.
- Sinclair, N., & de Freitas, E. (2014). The haptic nature of gesture: Rethinking gesture with new multitouch digital technologies. *Gesture*, *14*(3), 351-374.
- Sinha, R., Swanson, H., Clarke-Midura, J., Shumway, J. F., Lee, V. R., & Chandrasekharan, S. (2023, June). From Embodied Doing to Computational Thinking in Kindergarten: A Punctuated Motor Control Model. In *Proceedings of the 2023 Symposium on Learning, Design and Technology* (pp. 1-10).
- Turner, P. (2016). HCI Redux. Springer International Publishing.
- Vlassis, J. (2004). Making sense of the minus sign or becoming flexible in 'negativity'. *Learning and instruction*, *14*(5), 469-484.
- Whitacre, I., Bishop, J. P., Lamb, L. L. C., Philipp, R. A., Schappelle, B. P., & Lewis, M. L. (2012). Happy and sad thoughts: An exploration of children's integer reasoning. *The Journal of Mathematical Behavior*, *31*(3), 356–365.
- Wilensky, U., & Papert, S. (2010). Restructurations: Reformulations of knowledge disciplines through new representational forms. *Constructionism*, *17*(2010), 1-15.
- Wu, L.-H., & Hsu, P.-Y. (2011). Embodied cognition of information visualization: Human–computer interaction with six degrees of freedom in movement of an information space. *Expert Systems with Applications*, 38(9), 10730–10736.

### Maker-Material Creative Embodiments in Collaborative Making

#### Alekh V<sup>\*</sup>, Chandan DASGUPTA

Indian Institute of Technology, Bombay, India \*alekhv@iitb.ac.in

**Abstract:** Making opens up dialogic learning spaces where makers engage with the material world via embodied enactments. The article aims to contribute to the understanding of embodied sociomaterial entanglements where makers and materials encounter to form collaborative networks. In this study, we follow the encounters between makers and materials in design problem solving scenarios to uncover embodied material experiences leading to creative movements. We specifically look at a making context where novice makers collaboratively work on design problem solving, through the combined theoretical aspects of constructionism and posthumanism. Our findings suggest that the makers' embodied experiences were entangled with material elements of the situated making contexts and shaped the opportunities for learning. Makers and materials composed the embodied actions through which domain related concepts manifested.

Keywords: Making, Creativity, Materiality, Embodied Learning

#### 1. Introduction

Making is recognized as a means to constructively engage learners in creative designs and bring out skills in tackling uncommon problems as it provides opportunities to solve real-world problems, following different ways to arrive at multiple solutions (Honey & Kanter, 2013; Martin, 2015). Learning through collaborative making involves iterative cycles of building, testing, modifying the co-inventions, and reaching common ground with the participating makers. In makerspaces individuals and groups engage in multiple forms of design activities, resulting in cross-pollination of ideas and skills (Gross & Do, 2009; Peppler & Bender, 2013; Halverson & Sheridan, 2014). The opportunity to play and engage with materials is considered to act as "a social glue" for people to converge and engage in collaborative and creative activities (Vossoughi & Bevan, 2014; Honey & Kanter, 2013; Ingold, 2013). Studies suggest that hands-on experimentation and production across different media and digital platforms support learner's creative and critical engagement in disciplinary and transversal learning (Hughes, 2017; Ratto, 2011). Makerspaces hold the potential for interdisciplinary connections, collaboration, creativity and innovation (Martinez & Stager 2013). The emergent outcomes of these collaborative networks are conceptualized to be spread across social, material, and temporal dimensions of the creative phenomenon (Glăveanu, 2014; Sawyer & Dezutter, 2009). Making being a materially grounded activity, makers and materials are in continuous entanglements, considering the theorization of "design as a conversation with the situation" (Schön, 1983). Here the makers listen to the material entities, and the material entities "talk back" to maker advancements. Such dialogic exchanges support the understanding of differences or similarities in approaching design problems among makers, learning of material characteristics, and opportunities to search through the problem and solution spaces (Shotter, 2006). Makers converse with fellow makers and materials which can result in being, doing, and becoming material conversations (Gravel & Svihla, 2021).

Being material conversations considers the materials as well-known and not improvable, whereas doing material conversations involve exploring known material affordances that can be utilized to make progress in solving a design problem. A becoming material conversation is characterized by transformed purpose, reconfigurations, and modifications of materials (Gravel & Svihla, 2021). In finding creative solutions to design problems, where makers encounter familiar and unfamiliar material elements, the being, doing, and becoming material engagements acknowledge, use, and transform the material affordance. These sociomaterial conversations can involve shared agency, which reiterates the material aspect of emergent creativity. Therefore, materials cannot be considered as dormant entities (Pickering, 1993), not only responding to the actions performed by makers but also directing the makers to creative actions. The reframed views on agency manifesting in these maker-material conversations can disrupt the longstanding human-centric explanations on learning, making, and creative practices. With Latour's conceptualization of agency and flattened ontology (1996), all material entities are possible actants and actively play roles in everyday activities. Hence, agency is neither rested solely on human nor nonhuman elements but through becoming of emergent manifestations, effects through certain configurations of situated entities (Suchman, 1987), and unfolds in practice (Pickering, 1993).

Making, especially in the context of design problem solving, celebrates the philosophies of learning by doing, failures, iterations, sharing knowledge and other resources with fellow makers (Halverson & Sheridan, 2014). In alignment with constructionism, learning is viewed to happen with the learners engaging and manipulating physical or digital materials to arrive at personally meaningful artifacts and share with the community. Constructionist ideas of "objects-to-think-with" and "body syntonicity" have been explored by the research community in understanding how learning can happen when learners engage with tangible materials. "Objects-to-think-with" can be considered as bridges that can connect abstract ideas and sensory knowledge intertwined with embedded participation and personal linkages. In terms of "body syntonicity", learning is viewed to surface as learners imagine their bodies in place and connect with the materials under manipulation (Papert, 1980, 1993). The learner's knowledge and awareness of their bodies in the learning environment can support the internalization of domain related concepts and abstract ideas.

Materials are significant in the process of learning as learners encounter a wide range of other-than-human elements in the learning environment, but the facets of these entities are not closely looked into (Engeström & Blackler, 2005). In the constructivist views of learning, humans are considered as observers and actors, whereas non-human entities are seen as malleable and controllable. Similarly in sociocultural studies, materials are seen as mediators of human practice (Sheridan et al., 2020). On the other hand, the posthumanist perspective can reveal the active ways in which digital and physical materials direct the learning process (Thiel, 2015). The notion of posthumanism looks at decentralizing humans from the sole source and center of actions and unsettling the concepts related to material neutrality (Barad, 2003, 2007; Pickering, 1993). It can bring the vitality of the matter, unpredictability, unfolding ontology, and signifies material turns as humans encounter non-human matter (Barad, 2003; Haraway, 1985). Further insights are required to understand the ways in which materials can influence learning through embodied experiences with the world.

We see the combined theoretical perspectives of constructionism and posthumanism as having immense potential to investigate the learning dynamics in rich sociotechnical spaces like makerspaces. In line with the fused theoretical approach, we follow the idea of "material syntonicty" which provides a material direction to look at the encounters between makers and materials resulting in learning opportunities (Keune, 2022). Material syntonicity points to how materials become actively engaged in the embodied intra-actions with learner bodies and provides the learners with opportunities for material ways of knowing. The analytical lens can advance the understanding the ways in which materials can influence domain learning and performing domain related concepts. Material syntonicity adds to the constructionist ideas in such a way that it reveals how materials direct embodied experiences to produce domain related phenomena. By acknowledging the active states of materials and learner bodies, material syntonicity supports further investigation into how learner experience themselves as active entities of the embodied learning process.

Based on these prior literature, we look into the broader research goal of understanding how maker and material encounters shape the emergent creativity in makerspaces. With this study, we attempt to explore *how novice makers and materials come together to shape the*  *learning opportunities and creative movements in a situated collaborative making context.* We think with constructionist and posthumanist ideas to analyze the making context where seventh-grade students collaboratively work on design problem-solving. We consider design as an inherent practice within the making process (Bevan, 2017; Dougherty, 2012) and position the creative aspects of the making process as emergent (Sawyer & Dezutter, 2009; Tangaard, 2013).

#### 2. Study details

Data for this study was collected as part of a maker workshop where two groups of seventhgrade students from an English-Medium school in the city of Mumbai, India engaged in engineering design problem. The maker workshop was organized at a leading engineering institute in India, and divided into two sessions, (1) Training session and (2) Design session. The first team-Team A- consisted of two female participants (A1 and A2) and a male participant (A3). The second team-Team B-consisted of one female participant (B1) and one male participant (B2). Teams were formed randomly and they worked on designing a cleaning robot in the design session. The design challenge given to the students is stated as: "Keeping our surroundings clean is a very important aspect of our life but, doing that requires a lot of manual work and can get boring sometimes. Wouldn't it be amazing if a robot does that for you autonomously? Your design challenge for today is to use the Lego Mindstorm kit to build a cleaning robot. Your robot should be able to clean at least two of the following trash materials- paper bits, Lego pieces, water droplets, eraser dust, and pencil dust. You can also use the provided supplementary material i.e., cleaning mop wipes, cardboard, and sponge. The robot designs of the two teams will be compared based on how many trash items the robot can clean, how well it cleans the trash material, and the cost of the robot." Along with the Lego Mindstorm kit and the supplementary materials, each team was also provided with a cost-calculation sheet and a workbook for taking notes and making sketches. A facilitator mentor was allocated to each team to take observation notes and provide technical and logistic support. For our analysis, we followed the making activities of Team B. Audio and video recorders were used for capturing the making. The design artifacts consisting of sketches, written notes, and the final robot designed by the team were also collected.

#### 3. Analysis

We adopted the case study methodology (Merriam, 2007) to unpack the processes involved in the making settings of Team B in the Maker activity-centered workshop. We combined constructionist perspectives with posthumanist standpoint to understand the ways in which makers and materials come together to configure the learning opportunities and creative movements (Keune, 2022). With constructionism pointing towards the ways in which learners realize ideas via design and embodied experiences to make meaningful expressions, and posthumanist consideration of material entities as active participants in phenomena, the dual theoretical approach augments our effort in following maker-material entanglements in learning and subsequent creative actions. We followed the dual theoretical perspectives and adopted the methodological process of thinking with theory (Jackson & Mazzei, 2012; St. Pierre, 2008). We viewed the assemblage of data sources simultaneously and iteratively. Here, we see ourselves entangled and becoming with the unfurling of research course. The research entanglement involved the encounter with data sources as video, photographs, design artifacts, field notes, and observations during making.

#### 4. Findings

We found that makers and material entities of the situated making context came together and

configured the entangled embodied enactments. We present representative episodes from the making context to show the embodied intra-actions enmeshing human and material entities where materials become actively engaged with learner bodies and provide the learners with opportunities for material ways of knowing.

B1 and B2 began design problem solving by rooting on the preliminary idea of a two wheeler robot with a cleaning unit attached so that the combination can drag the trash items along with it. The makers kept a close look at the cost component of the robot throughout the making process. The following episode illustrates how makers engaged with the materiality of the making context, where the embodied experiences lead to subsequent creative acts.

- B1: See [points to the robot].. we can move the robot forward and the sponge will push the paper bits out.
- B2: I will run it on the floor.
- B2: [Drives the two wheeler robot over the carpet] Its not good.. getting stuck at some place.. not smooth.
- B1: Lets try outside..
- B2: [changes the sponge orientation to horizontal]
- B2: [Drives the robot on the smoother test floor] better.. but it feels wiggly kind.



*Figure 1.* Team B makers driving the two wheeler robot with sponge attachment on different test floors.

Here the makers tried to make a two wheeler robot with a sponge that can push the trash materials as it travels forward. Once the sponge was fixed to the front end of the robot, B2 moved the robot over the carpet manually. The carpet resisted the robot motion as the sponge unit and carpet came into contact. B2 drove the robot manually and experienced the resistance through his body as the cleaning robot came in contact with the carpet. This encounter between the cleaning robot and the carpet pushed the makers to change the orientation of the sponge placement and the testing floor, as seen in Figure 1. The makers moved to a smoother floor to understand the motion of the cleaning robot unit. Again B2 drove the robot manually and felt the resistance in action. The makers perceived the notion of friction via embodied encounters where the different test floors, robot design, sponge and the maker bodies came together and arranged the spots for syntonicty. The particular instance from the making session is an example of how materials resist to maker actions and maker responds via embodied entanglements. Here we also read that the domain related concept of friction manifesting, and the makers acknowledging the same in the emergent learning opportunity. This opening helped them in understanding the notion of friction with varying surfaces in contact and how it affects the motion.

The makers proceeded with the session by considering changes in the robot design. From the previous tests, the makers ideated to use bigger wheels for the drive. As the next step, makers came up with the idea to eliminate sponge unit from the robot design and to have alternate cleaning mechanisms. In the following excerpt, we see that the makers make moves for a four wheeled robot, and the maker-material encounters provide insights to design stability.

- B1: Can we use bigger wheels or.. maybe four wheeled.. like a simple car with a broom.
- B2: We can.. [takes a lego motor and two wheels]
- B1: No sponge.
- B2: [Connects the lego motors and wheels to the brick with connectors]
- B2: [Exerts force on the robot from top] It's kind of weak.. moving up and down ..loose maybe.



Figure 2. B2 of Team B inspecting the robot design via embodied material ways

In the above episode, we can see that the makers looking for design possibilities and modifying design features. B2 examined the four wheeled robot for mobility, and checked whether the links are properly connected. As B2 tried to drive the robot on the glass table, he felt the robot clumsy. B2 then started to exert force on the robot from the top and followed by lifting the robot, as seen in Figure 2. The makers identified that the robot needs further reinforcements and design changes so that the intended purpose of cleaning is met. This unscripted play with the robot model gave the makers awareness of the stability of the robot. Here we see that the four wheeled robot, maker body, and knowledge from the previous experience are entangled in such a way that the collective embodiment of the making context paved way for the makers to be cognizant of the stability aspects of the robot design. The collective embodiment of the making context was not only shaped by the makers but also by its materiality and syntonicity.

Once the four wheeler robot structure was finalized, the makers turned their attention in making the cleaning unit. The makers envisioned having the cleaning unit to be attached to the rear end of the robot. As the idea of using sponge was eliminated, they looked for other possible material components for making the cleaning unit.

- B1: [Finds the cleaning wipe mop] this can be used.
- B2: I will check other things.
- B1: [Running fingers through the wipe mop] its soft..
- B1: [Finds the soft cardboard box and cuts out a piece]
- B2: Support?
- B1: [Attaches the cardboard piece with sponge using stapler pins] yeah.

B1: [Cleans the paper trash with the cleaning unit] it works.. needs to be dragged with it.



Figure 3. B1 of Team B testing the cleaning wipe attachment by wiping the paper trash

B1 encountered a cleaning mop wipe on the working table and inspected it. B1 interviewed the cleaning mop wipe by visual inspection and running fingers through the wipe. The bodily engagement with the wipe revealed that it needed a support element so as to meet the purpose of cleaning and suitability of attachment to the robot structure. B1 cut a soft cardboard box, flattened the pieces, and used stapler pins to attach the soft cardboard piece with the cleaning mop wipe. B1 then tried to clean paper piece trash from the working table using the modified wipe manually, as seen in Figure 3. Here the meaning of cleaning is co-created by the maker body, the modified mop wipe, paper trash and the glass table, in conjunction. These materially syntonic intertwined bodily engagements and actions were crucial in identifying the necessary modes of wiping acts that the robot should perform. Although not explicitly, the test pointed towards the momentum and the material ways by which the robot is expected to produce the cleaning actions.

As the robot structure and cleaning unit were ready to use, makers used double tapes to connect both components. B2 proceeded with testing the cleaning robot by driving on the test floor manually and found that the mop wipe attachment was having contact with the robot wheels. The makers changed the orientation of the cleaning wipe mop, and B2 drove the cleaning robot over the paper trash, as seen in Figure 4. As B2 drove the robot, the makers found that the cleaning contact had good contact with the trash, but was unable to drag the trash with it due to the lack of momentum of the manual drive. The makers then followed up with changes in number of turns and speed of rotations for the Lego motors using programming. Here the cleaning robot and the maker body were in tandem during the test run which helped the maker to make sense of the wiping action of the combined maker-material effort with respect to the trash cleaned and momentum generated.

- B2: [Drives the robot on the floor] Feels like the mop is touching the robot wheels.
- B1: Make it little.. go up [gesturing in upward direction].
- B2: [Drives the robot on the floor] mm.. better.
- B1: Ok.. lets try with paper bits [ spreads paper trash].
- B2: [Drives the robot on the floor] yeah.. its touching the pieces..
- B1: See.. but its not moving them.. going very slow
- B2: Yeah.. it's the floor and paper
- B1: We can change the speed and motor turns [connects the robot to programing interface]



Figure 4. B2 driving the cleaning robot to follow robot motion and cleaning actions.

#### 5. Discussion & Conclusion

In this article, we presented an empirical case study of collaborative making context from India where novice makers work on design problem solving with a variety of materials. We used a dual theoretical perspective of constructionism and posthumanism to examine the entangled embodied enactments when makers and materials come together and make together to shape learning opportunities. The constructionist approach helped us to see how makers enacted design ideas and domain related concepts via embodied ways, whereas the posthumanist lens revealed the active role of materiality in such entangled embodied experiences. With this standpoint, the study has put forward the need to reimagine the role of materials in embodied learning experiences. Here we advocate the point of departure from considering materials from mere mediators between learners and ideas, units for cognitive development to active co-participant in the relational field of encounters among learners and materials via bodily engagements. The active participation of materiality came into play as the material entities enmeshed with maker bodies in action, as in the case when the makers had to vary the test floor in accordance with the emergent meaning of resistance and motion. The intra-actions among the maker body, robot structure, sponge, test floors opened up opportunities for makers to get entangled with domain related concept of friction. In a similar way, the unexpected play modes of exerted and reactive forces among the maker body and robot structure brought out the ideas related to stability. The entanglements of the makers with cleaning mop, robot motion, paper trash, test floor, wiping moves, and programming directed the embodied experiences and material ways of knowing the momentum in action as the robot was driven manually to clear and carry the trash along the direction of motion. These instances highlighted how the emergent material syntonicity among materials and learner bodies provided the learners with opportunities for material ways of knowing.

Further, we found that that the creative acts featured in the maker workshop were not alone driven by makers. These creative movements were not only shaped by social entities but also by material entities of the making contexts. The findings signify that makers and materials molded the creative making that the processes span across the social and material elements of the situated environment, and flare out temporally (Sawyer & Dezutter, 2009; Tangaard, 2013). We noticed that maker-material and maker-maker encounters through dialogues and embodied experiences helped in knowledge co-construction that helped to investigate possible solution approaches for the design problem. With this article, we attempt to place the combined theoretical perspectives of constructionism and posthumanism as a way to look at embodied entanglements in collaborative making, following the concept of material syntonicity. We place collaborative making as a potential research domain to implement and advance the understanding of maker-material entangled embodied experiences and relations to enhance learning, where the creative phenomena evolve in a flat plane of contributing entities (Petrich et. al., 2013; Thiel, 2015; Timotheou & Ioannou, 2019).

The present study focused on data associated with a typical collaborative making context in which makers engage in solving design problems through making, however, advanced studies are required to understand and mark further modes of maker-material entanglements in collaborative creative making.

#### References

- Barad, K. (2003). Posthumanist performativity: Toward an understanding of how matter comes to matter. *Signs: Journal of women in culture and society*, *28*(3), 801-831.
- Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Duke university Press.
- Bevan, B. (2017). The promise and the promises of making in science education. *Studies in Science Education*, 53(1), 75-103.

Dougherty, D. (2012). The maker movement. *Innovations*, 7(3), 11–14.

Engeström, Y., & Blackler, F. (2005). On the life of the object. Organization, 12(3), 307-330.

Glăveanu, V. P. (2014). *Distributed creativity: Thinking outside the box of the creative individual.* Cham/Heidelberger: Springer International Publishing.

- Gravel, B. E., & Svihla, V. (2021). Fostering heterogeneous engineering through whole-class design work. *Journal of the Learning Sciences*, *30*(2), 279-329.
- Gross, M. D., & Do, E. Y. L. (2009). Educating the new makers: Cross-disciplinary creativity. *Leonardo*, 42(3), 210-215.
- Halverson, E. R., & Sheridan, K. (2014). The maker movement in education. *Harvard educational review*, *84*(4), 495-504.
- Haraway, D. (1985). A Manifesto for Cyborgs: Science, Technology, and socialist feminism in the 1980s. *Socialist Review*, *80*(1985), 65–108.
- Honey, M. (Ed.). (2013). *Design, make, play: Growing the next generation of STEM innovators*. Routledge.
- Hughes, J. M. (2017). Digital making with "at-risk" youth. *The International Journal of Information and Learning Technology*, *34*(2), 102-113.
- Ingold, T. (2010). The textility of making. Cambridge journal of economics, 34(1), 91-102.
- Ingold, T. (2013). Making: Anthropology, archaeology, art and architecture. Routledge.
- Jackson, A. Y. (2013). Posthumanist data analysis of mangling practices. *International Journal of Qualitative Studies in Education*, 26(6), 741-748.
- Johri, A. (2011). The socio-materiality of learning practices and implications for the field of learning technology. *Research in Learning Technology*, *19*(3), 207-217.
- Keune, A., & Peppler, K. (2019). Materials-to-develop-with: The making of a makerspace. *British journal of educational technology*, *50*(1), 280-293.
- Keune, A. (2022). Material syntonicity: Examining computational performance and its materiality through weaving and sewing crafts. *Journal of the Learning Sciences*, *31*(4-5), 477-508.
- Kirchhoff, M. D. (2009). Material agency: a theoretical framework for ascribing agency to material culture. *Techne: research in philosophy and technology*, *13*(3), 206-220.
- Latour, B. (1996). On Interobjectivity. Mind, culture, and activity, 3(4), 228-245.

Martin, L. (2015). The promise of the maker movement for education. *Journal of Pre-College* Engineering Education Research (J-PEER), 5(1), 4.

Martinez, S. L., & Stager, G. (2013). Invent to learn. *Making, tinkering, and engineering in the classroom. Torrance, Canada: Construting Modern Knowledge.* 

- Merriam, S. B. (2007). Qualitative research and case study applications in education. Jossey-Bass.
- Papert, S. (1980). Mindstorms: Children, computers, and powerful ideas. Basic Books.
- Papert, S. (1993). The children's machine: Rethinking school in the age of the computer. Basic Books.
- Patton, R. M., & Knochel, A. D. (2017). Meaningful makers: Stuff, sharing, and connection in STEAM curriculum. *Art Education*, *70*(1), 36-43.
- Peppler, K., & Bender, S. (2013). Maker movement spreads innovation one project at a time. *Phi Delta Kappan*, *95*(3), 22-27.
- Petrich, M., Wilkinson, K., & Bevan, B. (2013). It looks like fun, but are they learning?. In *Design, make, play* (pp. 68-88). Routledge.
- Pickering, A. (1993). The mangle of practice: Agency and emergence in the sociology of science. *American journal of sociology*, *99*(3), 559-589.

- Ratto, M. (2011). Critical making: Conceptual and material studies in technology and social life. *The information society*, 27(4), 252-260.
- Sawyer, R. K., & DeZutter, S. (2009). Distributed creativity: How collective creations emerge from collaboration. *Psychology of aesthetics, creativity, and the arts, 3*(2), 81.

Schön, D. A. (1983). The reflective practitioner: How professionals think in action. Routledge.

- Sheridan, M. P., Lemieux, A., Do Nascimento, A., & Arnseth, H. C. (2020). Intra-active entanglements: What posthuman and new materialist frameworks can offer the learning sciences. *British journal of educational technology*, *51*(4), 1277-1291.
- Shotter, J. (2006). Understanding process from within: An argument for 'withness'-thinking. *Organization Studies*, 27(4), 585-604.
- St. Pierre, E. A. (2008). Decentering voice in qualitative inquiry. *International Review of Qualitative Research*, 1(3), 319-336.
- Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge university press.
- Svihla, V., Tucker, M., & Hynson, T. (2020, April). What Gaze Data Reveal About Material Agency: Resilient makers, materials and ideas. In *Proceedings of the FabLearn 2020-9th Annual Conference on Maker Education* (pp. 54-60).
- Tanggaard, L. (2013). The sociomateriality of creativity in everyday life. *Culture & Psychology*, *19*(1), 20-32.
- Thiel, J. J. (2015). Vibrant matter: The intra-active role of objects in the construction of young children's literacies. *Literacy Research: Theory, Method, and Practice*, 64(1), 112-131.
- Timotheou, S., & Ioannou, A. (2019). On making, tinkering, coding and play for learning: A review of current research. In *IFIP Conference on Human-Computer Interaction* (pp. 217-232). Springer, Cham.
- Vossoughi, S., & Bevan, B. (2014). Making and tinkering: A review of the literature. *National Research Council Committee on Out of School Time STEM*, 67, 1-55.

# Virtual Reality and Embodied Learning: Unraveling the Relationship via Dynamic Learner Behavior

#### Antony PRAKASH<sup>\*</sup>, Ramkumar RAJENDRAN

Interdisciplinary Programme in Educational Technology, Indian Institute of Technology Bombay, India

\*antony80004@gmail.com

**Abstract:** Embodied Cognition theory asserts that the physical body, its environment, and the interplay between them hold a pivotal role in the process of embodied learning. In comparison to traditional and computer-based learning methods, Virtual Reality (VR) enhances embodied learning, primarily owing to its capacity to offer a sense of situatedness and engage learners through physical interactions. Furthermore, the incorporation of haptic sensations in VR-based learning introduces tactile sensory memory, complementing the existing auditory and visual sensory dimensions. Despite the evident advantages of VR in facilitating embodied learning, the current body of literature has yet to delve into the dynamic behaviors exhibited by learners within Virtual Reality Learning Environments (VRLEs) during embodied interactions, and their inherent relationship with the embodiment phenomenon. In this paper, we take a significant step by integrating the Interaction Behavioral Data (IBD) collection mechanism, a development from our previous work, into a VRLE. This integration is achieved by adopting a structured framework for embodied learning activities within a VR context. Through an empirical study involving 14 participants, we meticulously logged their interaction traces using the IBD logger. Subsequently, we effectively extracted the diverse embodied learning activities carried out by these participants within the VRLE, from the interaction trace data collected in the IBD logger. Our endeavor aims to establish meaningful connections between these embodied interaction activities and the overarching concept of embodiment. These correlations, once identified, will serve as invaluable insights for VR content developers, offering clear guidelines for the design and creation of VRLEs that optimally facilitate enhanced learning experiences. Ultimately, this knowledge transfer will not only empower instructors and learners to harness VR as a potent educational tool within their regular teaching and learning practices but also foster the seamless integration of VR into mainstream education.

**Keywords:** Embodied Cognition, Embodied Learning, Virtual Reality, embodied interactions, haptic feedback

#### 1. Introduction

In line with the principles of embodied cognition theory, cognition emerges from the dynamic interplay between the body and its surroundings, closely tied to sensorimotor engagement (Rambusch & Ziemke, 2005). Both the fusion of the body with the mind and the surrounding environment contribute to cognitive processes (Wilson, 2002). The practical application of the body within educational practices (Georgiou & Ioannou, 2019). In the context of embodied learning, learners acquire knowledge not solely through conceptual comprehension, but also through bodily postures, gestures, and facial expressions (Foglia & Wilson, 2013). Embodied learning encompasses sensorimotor activities, where sensory inputs such as touch, vision, and hearing highly influence the learning process. The literature presents substantial evidence of brain activation in sensorimotor regions during cognitive tasks, underscoring the intertwined nature of cognition and bodily actions (Lindgren & Johnson-Glenberg, 2013).

The emergence and progression of Information and Communication Technology (ICT) have precipitated a swift evolution of technology-facilitated embodied learning within the realm of educational practice over the past decade (Georgiou & Ioannou, 2019). Among these technological advancements, Virtual Reality (VR) stands out as a particularly sophisticated tool that fosters embodied learning (Georgiou & Ioannou, 2019; Li et al., 2020; Chatain et al., 2023). This article focuses on the utilization of VR and delves into its inherent attributes within the context of embodied learning.

Within the domain of VR, users are immersed in a simulated three-dimensional environment from a first-person perspective, engaging with virtual objects in a manner akin to real-world interaction. This immersion is a key factor underpinning embodied learning in VR, encapsulated by the trio of qualities known as the 3Is - Immersion, Interaction, and Imagination (Zhang, 2017; Yang et al., 2019).

- Immersion pertains to the VR system's capacity to envelop users within the virtual world to such an extent that their awareness of the external physical environment diminishes.
- Interaction denotes the VR system's ability to provide immediate responses and feedback based on the user's actions within the simulated setting.
- Imagination refers to the VR system's capability to induce the perception that the virtual environment is genuine, even if it does not exist in the tangible real world.

Leveraging the immersive quality of VR, learners find themselves able to completely engage with and encounter the virtual environment through a first-person perspective, momentarily suspending the acknowledgment of their real physical surroundings. This phenomenon allows learners to grasp a sense of situatedness, an aspect seldom replicated consistently within conventional classroom setups (Li et al., 2020). Situatedness stands as a key factor in the realm of learning, a notion widely recognized (Fors et al., 2013). At the heart of situated learning lies the principle that a substantial portion of learning is intricately tied to the context in which it unfolds (Anderson et al., 1996). The tasks undertaken by learners within specific contexts, aimed at achieving particular objectives, serve as vessels for the transmission of situated cognition. Herein, the interactive facet of VR offers the opportunity for learners to execute tasks within fabricated scenarios within the virtual realm, utilizing bodily movements. As a result, VR effectively cultivates conditions conducive to situated learning, affording experiences that combine perception and action, thereby enriching the landscape of embodied learning (Li et al., 2020).

While Virtual Reality (VR) contributes to the advancement of embodied learning through its immersive and interactive attributes, previous research has yet to explore the correlation between learners' action-oriented behaviors within Virtual Reality Learning Environments (VRLE) and the embodiment facilitated by VR technology. This is because, there were no effective mechanism to log data related to the behavior of the learners while doing activities in VRLE. To bridge this gap, we devised a mechanism known as Interaction Behavioral Data (IBD) logging, designed to capture learners' real-time interactions, along with timestamps (Prakash & Rajendran, 2022). This IBD logging mechanism was seamlessly incorporated into MaroonVR (Pirker et al., 2019), a Virtual Reality Learning Environment utilized for the acquisition of physics concepts (Prakash et al., 2023). In this paper, we aim to find the relationship between the dynamic behavior and the related embodied interaction activities extracted from the IBD logging mechanism and the embodiment.

The paper is organized as follows: In Section 2, we delve into diverse categories of VR systems and their capacity to facilitate embodied learning. Additionally, this section challenges prevailing notions about interaction to foster an environment conducive to embodied learning. The IBD logging mechanism is detailed in Section 3, which also offers an in-depth explanation of the VRLE embedding the IBD logging mechanism. Section 4 provides a comprehensive portrayal of the framework underpinning embodied interaction activities. This section also enumerates the multimedia principles that have been integrated into the VR experience.

Moving to Section 5, we present the methodology of a study conducted with a cohort of 14 participants. In Section 6, we present the results of the study highlighting the extracted embodied interaction activities logged via IBD. Section 7 encapsulates the article by summarizing the current research, delineating its limitations, and casting a glimpse into the potential avenues for future exploration.

#### 2. Literature Review and Background

In this section, we first describe the different types of VR and the level of embodiment supported by them and then various interactions favored in VR for embodied learning.

#### 2.1 VR Device for Embodied Learning

The level of embodiment facilitated by VR is contingent upon the extent of immersion and interaction characteristics (Li et al., 2020). Consequently, VR can be categorized as follows: 1. Non-immersive VR, and 2. Immersive VR (iVR) (Freina & Ott, 2015).

Non-immersive VR involves viewing the virtual world on a desktop screen, promoting an allocentric viewpoint rather than an egocentric one. Hence, Immersion requires contextualization to compensate (Li et al., 2020). Furthermore, instead of physical bodily movements, interactions rely on keyboard and mouse controls (Li et al., 2020; Hamilton et al., 2021), making non-immersive VR less effective for embodied learning.

Within immersive Virtual Reality (iVR), users undergo spatial immersion through the aid of VR gears such as headsets and controllers. Accordingly, iVR can be further categorized into three distinct types: 1. Low-end VR (Freina & Ott, 2015; Radianti et al., 2020), 2. High-end VR (Freina & Ott, 2015; Radianti et al., 2020), and 3. Enhanced VR (Radianti et al., 2020).

Low-end VR is mobile VR, using devices like Google Cardboard and Samsung Gear VR. Interaction involves keyboards, handheld Bluetooth controllers, and gaze controls (Radianti et al., 2020). It offers three degrees of freedom (3DOF) allowing head rotation in roll, pitch, and yaw (Freina & Ott, 2015). Whereas, in high-end VR, users engage with the virtual environment using HMDs and integrated Hand Held Controllers (HHCs). High-end VR systems enable users to perform actions such as walking, touching, grabbing, and dropping objects, conducted through HHCs. These high-end VR headsets offer six degrees of freedom (6DOF), incorporating additional translational movements like forward-backward (surge), left-right (sway), and up-down (heave) within the simulated virtual environments (Freina & Ott, 2015). Devices like HTC Vive and Oculus Quest exemplify high-end VR technology.

Enhanced VR combines HMD, HHC, data gloves, and bodysuits with haptic sensors, offering immersive touch experiences (Radianti et al., 2020). Research indicates that learning can be significantly enriched when the sense of touch is engaged in addition to visual stimuli, highlighting its relevance (Magana et al., 2017). For optimal embodied learning, haptic-enhanced VR is preferred over high-end and low-end options. However, VR systems equipped with such haptic devices can be notably expensive (Magana et al., 2017). To address this, we opted for a high-end VR setup that integrates HHCs with haptic capabilities, thereby providing tactile feedback during the VR experience.

#### 2.2 Interactions For Embodied Learning

Embodied learning is supported by embodied interactions (Chatain et al., 2023). Within Virtual Reality (VR) learning environments designed to cultivate egocentric perception, interactions can manifest through visual, manual, or bodily means. However, certain VRLEs may not fully facilitate interactions. For instance, Sun et al. (2018) conducted a study using a virtual museum as a VR learning environment (VRLE). Learners navigated the environment with a first-person perspective and engaged physically with the virtual objects (Magana et al., 2017). Visual cues above objects directed attention to relevant information. Similarly, Sharma et al. (2018) created LearnDNA, a VRLE for understanding DNA concepts. LearnDNA allowed learners to explore a virtual DNA representation. However, both lacked interactions, leading to

limited hands-on experience. Although immersive, these environments lacked interaction, resulting in limited embodiment.

Certain VRLEs with serious games as educational content blend heightened immersion and interaction. Sankaranarayanan et al. (2018) created an immersive simulation for healthcare training, involving extinguishing a fire in an operating room. Learners navigate around flammable objects, interacting through touch, grab, and drag actions. However, it's worth noting that there is no evidence of haptic feedback being utilized in this context. The absence of a haptic device detracts from the tactile sensations that would normally accompany interactions with various objects. Consequently, the absence of tactile feedback could potentially result in a diminished sense of embodiment within the immersive and interactive environment.

Tactile engagement through hands-on activities holds a pivotal role in shaping cognition. Magana et al., (2017) explored hands-on learning of electricity and magnetism using visuohaptic simulation. Learners interacted with virtual electric charges through a haptic device, merging touch sensation while learning. Results showed no significant differences among visuohaptic, visual-only, and conventional groups in post-test scores. Although visuohaptic learning introduced embodied learning elements, the interactions were mediated by an external haptic device rather than being inherently tied to the learners' own bodily movements. Additionally, the portrayal of the virtual environment on a computer screen adopted an allocentric perspective, contributing to a diminished sense of embodiment. Considering our literature analysis, we selected a VRLE that promotes embodied interaction through auditory, visual, and tactile sensory channels to enrich embodied learning in VR.

#### 3. Virtual Reality Learning Environment (VRLE)

We developed a comprehensive Interaction Behavioral Data (IBD) collection mechanism to capture interactions in the Virtual Reality Learning Environment (VRLE) (Prakash & Rajendran, 2022; Prakash et al., 2023).



Figure 1. VRLE Scenes

The Interaction Behavioral Data (IBD) collection was integrated into MaroonVR (Pirker et al., 2019), a VRLE focused on physics learning. We applied the IBD logging mechanism to scenes involving Faraday's law and a falling coil experiment, enhancing comprehension of electromagnetic induction—a phenomenon characterized by the induction of electromotive force (emf) within a coil when subjected to the magnetic influence of a moving magnet. We added VR interfaces (virtual buttons and sliders) for varying coil turns, coil diameters, coil resistance, and magnetic strength and toggle the graph plot between the voltage curve and current curve to enhance the interactive experience in the VRLE. A new scene placed learners in the magnet's perspective, inducing emf and plotting real-time data on a virtual graph based

on the whole physical body movement of the learners. The scenes present in the VRLE are shown in Figure 1.

#### 4. Embodied Learning Activities Framework in VR

After reviewing the literature, we opted for the Oculus Quest 2 VR system developed by Meta. It offers heightened immersion with a 3D VRLE experienced through an HMD providing a first-person view and 6 DoF. Additionally, the integrated HHCs offer vibrotactile haptic feedback, alerting users in specific situations through vibrations.



*Figure 2.* Representation of Framework for Embodied Learning Activities in VR (adopted from Chatain et al. (2023))

We adopted the embodied interaction design framework represented by Chatain et al. (2023) as shown in Figure 2. According to this framework, both physical and social contexts influence embodied interaction through the physical movements of the part of the body or whole body to interact with the virtual objects with the goal of meaning-making and ultimately learning. In the MaroonVR VRLE, the physical context such as the length of the cable used for tethering the HMD and the obstacles present in the real world make the learners teleport from one place to another by using the thumbstick present in the HHC rather than making any physical locomotory movements to navigate in the digital world. The social context such as the perception of various virtual objects (magnet and iron bar) present in the VRLE make the learners grab, drag, and drop using the grip button of HHC. Similarly, the virtual interfaces (buttons and sliders) present in the VRLE make the learners use the trigger button of HHC to manipulate the variables. The learners immersed in the VRLE extend their real physical hands (with HHC) to establish contact with the virtual objects and other interfaces (indicated by a color change) and interact with them constituting embodied interaction. The physical movement of the hands of the learners is represented by digital HHC (avatar) which the learners have to contextualize as their hands. The learners also move their physical heads to look at the plot (voltage or current) in the virtual graph, make meanings, and manipulate the parameters (coil turns, coil diameters, coil resistance, and magnetic strength) for the learning to occur.

In tandem with the creation of the VRLE to cultivate enriched embodied interactions, the design also adeptly incorporated a selection of multimedia learning theory principles, further enhancing the potential for improved learning outcomes. Numerous studies exploring the learning impact within VR environments have yielded a blend of results (Hamilton et al., 2021). This variance in outcomes might stem from the potential cognitive overload introduced

by the multifaceted nature of learning in VR. Therefore, the VRLE and its interactions were meticulously structured to align with multimedia learning principles, aimed at mitigating the effects of cognitive overload linked to intrinsic, extraneous, and germane cognitive loads. Several multimedia learning principles were deliberately woven into the design:

- 1. The Pre-training Principle: This principle entailed the integration of a VR game named 'First Touch,' which allowed learners to familiarize themselves with the VR system and the corresponding interaction controllers.
- 2. The Self-pacing Principle: Learners were empowered to engage with the VRLE at their individualized pace, fostering a sense of self-guided progression.
- 3. The Signaling Principle: The learners were provided with several cues during VR intervention. For instances, the HHC avatar present in the VRLE signals the spatial movement of the hands of the learners, the change in colour of the virtual objects to blue colour indicate that the particular object is being touched and ready for interaction, and the vibrotactile haptic feedback indicate that the emf is induced.

Furthermore, the VRLE seamlessly incorporated the principle of temporal contiguity, ensuring that the aforementioned cues were temporally aligned signaling during the specific actions. The principles of spatial contiguity were equally applied, whereby textual information, including virtual object names and VR interfaces, were strategically co-located to their corresponding objects. This judicious amalgamation of multimedia learning principles into the VRLE's design is poised to optimize the learning experience, promoting comprehension and engagement while mitigating cognitive challenges associated with the immersive, multisensory nature of VR environments.

#### 5. Methodology

The IBD collection mechanism was put into operation within the MaroonVR VRLE, and a study was executed involving a group of 14 undergraduate engineering students. As the phenomena of electromagnetic induction is very familiar among the electrical engineering students, all the participants invited belonged to non-electrical engineering to avoid bias occurring due to prior knowledge. Prior to the commencement of the Virtual Reality (VR) intervention, we gathered relevant data encompassing participants' self-efficacy, self-regulation, and their existing understanding of electromagnetic induction phenomena (pre-test). Subsequently, during the VR intervention phase, the developed data collection mechanism automatically recorded comprehensive behavioral data. Following the conclusion of the VR intervention, we proceeded to amass data concerning learning outcomes (post-test).

#### 6. Results and Discussion

The impact of VR intervention has resulted in a significant increase in the normalized learning gain and the results are already published in one of the previous articles (Prakash et al., 2023). Based on the mean value of the post-test scores the participants were categorized into high and low performers. Also, the embodied action events performed by the participants are extracted using the data logged in the IBD logger. An excerpt of the IBD data logger is depicted in Figure 3 for reference.

The data extracted from the VRLE involved the utilization of specific columns from the IBD logger, including 'Controller Index', 'Button', 'Button Action', 'Objects', and 'Timestamp'. These columns facilitated the extraction of action events executed by the participants. The action events such as Navigation, Magnet Handling, Iron Bar Handling, Setting Coil Turns,

Setting Coil Diameter, Setting Magnetic Strength, Perspective Walking are extracted from the aforementioned IBD logger columns as shown in Table 1.

Controller Index	Button	Button Action	Button Pressure	Touchpad Axis	-	Touchpad Angle	Touchpad 2 Axis	-	Touchpad 2 Angle	Object	Time stamp
1	GRIP	unclicked	0	(0.0	0.0)	90	(0.0	0.0)	90	Magnet (VR1	6:14:13
1	GRIP	released	0	(0.0	0.0)	90	(0.0	0.0)	90	Magnet (VRT	6:14:13
1	GRIP	axis changed	0	(0.0	0.0)	90	(0.0)	0.0)	90	Magnet (VR1	6:14:13
1	GRIP	untouched	0	(0.0)	0.0)	90	(0.0	0.0)	90	Magnet (VR1	6:14:13
1	GRIP	touched	0	(0.0)	0.0)	90	(0.0	0.0)	90	2 Turns Coil	6:14:15
1	GRIP	pressed	1	(0.0)	0.0)	90	(0.0	0.0)	90	2 Turns Coil	6:14:15
1	GRIP	clicked	1	(0.0)	0.0)	90	(0.0	0.0)	90	2 Turns Coil	6:14:15
1	GRIP	axis changed	1	(0.0)	0.0)	90	(0.0	0.0)	90	2 Turns Coil	6:14:15
1	TRIGGER	untouched	0	(0.0)	0.0)	90	(0.0	0.0)	90	2 Turns Coil	6:14:16
1	GRIP	unclicked	0	(0.0)	0.0)	90	(0.0	0.0)	90	2 Turns Coil	6:14:16
1	GRIP	released	0	(0.0)	0.0)	90	(0.0	0.0)	90	2 Turns Coil	6:14:16
1	GRIP	untouched	0	(0.0)	0.0)	90	(0.0	0.0)	90	2 Turns Coil	6:14:16
1	GRIP	axis changed	0	(0.0)	0.0)	90	(0.0	0.0)	90	2 Turns Coil	6:14:16
1	BUTTON ONE	untouched	0	(0.0	0.0)	90	(0.0	0.0)	90	2 Turns Coil	6:14:16

*Figure 3.* An Excerpt of Interaction Behavioral Data Logger

The 'Controller Index' in the IBD logger identifies whether the interaction occurs with the right HHC or left HHC. The time duration of a specific action event is indicated by the entry in the 'Timestamp' between the 'Pressed' and 'Released' states in the 'Button Action' field. These action events, outlined in Table 1, encompass a wide range of interactions initiated by learners through their physical movements. These interactions vary, from actions like Magnet Handling, Iron Bar Handling, Setting Coil Turns, Setting Coil Diameter, and Setting Magnetic Strength—performed using specific body parts—to the comprehensive bodily engagement required for Perspective Walking. Magneti Atandling, Iron Bar Handling, Setting Coil Turns, Setting Coil Diameter, and Setting Coil Turns, Setting Coil Diameter, and Setting Magnetic Strength primarily contribute to the goal of inducing emf and are categorized under 'Interaction.'

Button	Objects	Action Event		
	Perspective Scene	Perspective Walking		
Grip	Magnet	Magnet Handling		
	Iron Bar	Iron Bar Handling		
	2, 4, and 6 Turns	Setting Coil Turns		
Trigger	2, 4, and 6 Diameter	Setting Coil Diameter		
	Magnetic Field Slider	Setting Magnetic Strength		
Touchpad	-	Navigate		

Participants were categorized as high or low performers based on post-test scores (Prakash et al., 2023). Analysis revealed no significant differences in pre-test scores (Mann-Whiteney U = 16.5, p = .368) or VR intervention duration (Mann-Whiteney U = 13, p = .174) at p < .05 between the two groups. However, there was a significant disparity in post-test scores (Mann-Whiteney U = 0, p = .002) at p < .05. Therefore, the high performers' ability to comprehend electromagnetic induction was examined using their bodily interactions in the VRLE. The percentage of embodied action event duration (Interaction, Navigation, and Perspective Walking) was calculated as  $\frac{Duration of Action Event}{Total Duration of VR intervention}$  for both high and low performers. Figure 4 displays a comparison of the duration of embodied action events (Interaction: Magnet Handling, Iron Bar Handling, Setting Coil Turns, Setting Coil Diameter, Setting Magnetic Strength), Navigation, and Perspective Walking between high and low performers. Mann-Whitney U tests found no significant differences in the percentage of Interaction duration (U = 15, p = 8), Navigation (U = 16, p = 8), or Perspective Walking (U = 15.5, p = 8). Although Figure 4 indicates that high performers engaged in action events for a longer period, the no significance in the duration of embodied action events is likely due to the limited sample size.



Figure 4. Distribution of Duration of Embodied Action Events

Interaction events like Magnet Handling and Iron Bar Handling involve tactile actions—touching, grabbing, dragging, and dropping virtual objects like magnets and iron bars in the VRLE. Meanwhile, events like Setting Coil Turns and Setting Coil Diameter entail adjusting coil parameters (number of turns, diameter, and resistance), and Setting Magnetic Strength focuses on modifying magnet field strength. To execute these actions (Magnet Handling, Iron Bar Handling, Setting Coil Turns, Setting Coil Diameter, and Setting Magnetic Strength), learners physically extend and manipulate their hands in the real world, activating corresponding events in the virtual environment. Perspective Walking, on the other hand, engages the entire body as learners physically walk to trigger the event. These actions enhance hands-on learning (Magana et al., 2017). The immersive VRLE experiences, combined with a first-person perspective, create a sense of situatedness. Additionally, vibrotactile feedback, triggered during emf generation, adds a tactile dimension. The interplay of bodily interactions in a highly situated environment, along with visual and tactile cues, results in embodiment during the conceptual learning of electromagnetic induction.

Learners control the virtual magnet, grabbing and moving it within a coil to induce emf, displayed in real-time on the virtual graph in the VRLE. Two avatar controllers, serve as the representation of the learner's hands, and learners perceive and use them as their own hands. The VR interfaces such as VR buttons and VR sliders are used to change the parameters of the coil and magnet. The VR buttons are pressed by the avatar touching by extending the physical hand and activating the trigger button of HHC. The VR sliders are operated by sliding movement of the physical hand after it is touched by the avatar. In addition, magnetic field lines with arrowheads are visible, originating from the magnet's north pole and extending toward the south pole. The magnetic field lines adapt dynamically as learners move the magnet. When learners move their hand (holding the virtual magnet) and observe the virtual graph for changes in induced emf, learning occurs through the sequence of *Context* => *Interaction* => *Bodily Action* => *Learning* (Figure 2). By extending their actual hand to make the avatar touch the virtual magnet (indicated by the magnet turning blue), grabbing it, and moving it in and out of the coil, learners experience vibrotactile haptic feedback in their hand, signifying that emf has been induced in the coil due to the magnet's movement (or hand

movement). Consequently, learning in this scenario follows the path of *Context* => *Interaction* => *Bodily Action* => *Interaction* => *Meaning Making* => *Learning* (Figure 2). Thus the objects within the VRLE create a physical context that facilitates interactions aligned with bodily actions, fostering learning as illustrated in the framework shown in Figure 2.

#### 7. Conclusion

We determined that embodied learning within the VRLE is facilitated through the engagement of embodied interaction activities. However, a deficiency exists within the current literature concerning the connection between embodied learning and these interaction activities, primarily due to the absence of an effective mechanism for logging such interactions. Consequently, we endeavored to devise an IBD mechanism capable of comprehensively recording all interaction traces, complete with timestamps. This data logging mechanism was seamlessly integrated into a VRLE, guided by the framework established for embodied interaction activities within a virtual environment. Furthermore, we successfully retrieved the embodied interaction traces following VR intervention. Although, there was no significant difference in the duration of embodied action events, it was higher for the high performers. Our ongoing efforts will extend to extracting temporal and spatial attributes related to these activities, subsequently delving into the exploration of their interrelationship with embodied learning.

The implemented IBD logger has effectively captured bodily interactions, yet it has not encompassed visual interactions mediated through participants' gaze (noticing the virtual graph). Consequently, there is a necessity to incorporate provisions within the IBD logger to record virtual objects viewed by learners. Such an enhancement would facilitate the acquisition of more comprehensive embodiment data, enriching the analytical capacity. While the experiment successfully extracted action events stemming from learners' bodily actions, it's important to acknowledge certain limitations. The experiment was conducted with a relatively modest sample size, comprising only 14 participants. Additionally, the VRLE's instructional content centered around the physics concept of electromagnetic induction. Consequently, establishing the generalizability of any assertions derived from this study would necessitate further experimentation with larger and more diverse participant groups, encompassing a broader spectrum of subject matter.

In our forthcoming endeavors, we are poised to delve into a comprehensive analysis of the interplay between the extracted action events and the measured learning outcomes. Given that the action events are a direct outcome of interactions driven by bodily actions, this correlation analysis will effectively elucidate the intricate connection between embodiment and the resulting learning outcomes. This investigation into embodied action events will be instrumental in the development of a learner model tailored for optimizing performance, with the added capability of early prediction of learning outcomes based on learners' dynamic bodily interactions. The insights gleaned from these analyses carry significant implications for multiple stakeholders. Notably, they will offer valuable guidance to VRLE content developers, instructors, and learners alike, facilitating the effective integration of VR-based learning and teaching within classroom settings. As we venture into this exploration, we anticipate enriching the understanding of how embodied actions shape learning outcomes, thereby enhancing the overall effectiveness of VR-enabled educational experiences.

#### References

- Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. *Educational* researcher, 25(4), 5-11.
- Chatain, J., Kapur, M., & Sumner, R. W. (2023, April). Three Perspectives on Embodied Learning in Virtual Reality: Opportunities for Interaction Design. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1-8).
- Foglia, L., & Wilson, R. A. (2013). Embodied cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(3), 319-325.

Fors, V., Bäckström, Å., & Pink, S. (2013). Multisensory emplaced learning: Resituating situated learning in a moving world. *Mind, Culture, and Activity*, *20*(2), 170-183.

Freina, L., & Ott, M. (2015, April). A literature review on immersive virtual reality in education: state of the art and perspectives. In *The international scientific conference elearning and software for education* (Vol. 1, No. 133, pp. 10-1007).

Georgiou, Y., & Ioannou, A. (2019). Embodied learning in a digital world: A systematic review of empirical research in K-12 education. *Learning in a digital world: Perspective on interactive technologies for formal and informal education*, 155-177.

Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*, *8*(1), 1-32.

Hindmarch, I. (2014). Psychomotor Performance in Humans. In: Stolerman, I., Price, L. (eds) Encyclopedia of Psychopharmacology. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-27772-6 181-2

Li, P., Legault, J., Klippel, A., & Zhao, J. (2020). Virtual reality for student learning: Understanding individual differences. *Human Behaviour and Brain*, *1*(1), 28-36.

Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by Embodiment. *Educational Researcher*. <u>https://doi.org/10.3102/0013189X13511661</u>

Magana, A. J., Sanchez, K. L., Shaikh, U. A., Jones, M. G., Tan, H. Z., & Benes, B. (2017). EXPLORING MULTIMEDIA PRINCIPLES FOR SUPPORTING CONCEPTUAL LEARNING OF ELECTRICITY AND MAGNETISM WITH VISUOHAPTIC SIMULATIONS. Computers in Education Journal, 8(2).

Pirker, J., Holly, M., Lesjak, I., Kopf, J., & Gütl, C. (2019). Maroonvr—an interactive and immersive virtual reality physics laboratory. *Learning in a digital world: Perspective on interactive technologies for formal and informal education*, 213-238.

Prakash, A., Shaikh, D. S., & Rajendran, R. (2023, March). Investigating Interaction Behaviors of Learners in VR Learning Environment. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 931-932). IEEE.

Prakash, A., & Rajendran, R. (2022, July). Mechanism to Capture Learners' Interactions in Virtual Reality Learning Environment. In 2022 International Conference on Advanced Learning Technologies (ICALT) (pp. 335-337). IEEE.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778.

Rambusch, J., & Ziemke, T. (2005, July). The role of embodiment in situated learning. In *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 1803-1808). Mahwah, NJ: Lawrence Erlbaum.

Roo, J. S., Gervais, R., Frey, J., & Hachet, M. (2017, May). Inner garden: Connecting inner states to a mixed reality sandbox for mindfulness. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 1459-1470).

Sankaranarayanan, G., Wooley, L., Hogg, D., Dorozhkin, D., Olasky, J., Chauhan, S., ... & Jones, D. B. (2018). Immersive virtual reality-based training improves response in a simulated operating room fire scenario. *Surgical endoscopy*, *32*, 3439-3449.

Sharma, L., Jin, R., Prabhakaran, B., & Gans, M. (2018, April). LearnDNA: an interactive VR application for learning DNA structure. In *Proceedings of the 3rd International Workshop on Interactive and Spatial Computing* (pp. 80-87).

Sun, L., Zhou, Y., Hansen, P., Geng, W., & Li, X. (2018). Cross-objects user interfaces for video interaction in virtual reality museum context. *Multimedia tools and applications*, 77, 29013-29041.

Wilson, M. (2002). Six views of embodied cognition. Psychonomic bulletin & review, 9, 625-636.

Yang, L. I., Huang, J., Feng, T. I. A. N., Hong-An, W. A. N. G., & Guo-Zhong, D. A. I. (2019). Gesture interaction in virtual reality. *Virtual Reality & Intelligent Hardware*, *1*(1), 84-112.

Zhang, H. (2017). Head-mounted display-based intuitive virtual reality training system for the mining industry. *International Journal of Mining Science and Technology*, 27(4), 717-722

## Enacting Biomolecular Interactions in VR: Impact on Student Conceptual Understanding in Biochemistry

#### Prajakt PANDE<sup>a&b\*</sup>

<sup>a</sup>Southern Methodist University, Dallas, United States of America <sup>b</sup>Roskilde University, Denmark \*ppande@smu.edu

**Abstract:** Through the theoretical lenses of embodied and enactive learning, this paper reports the instructional effectiveness of immersive virtual reality (VR) in helping undergraduates learn core biochemistry concepts. The reported pretest-posttest quasi-experimental study investigated how student enactment and learning of biochemical interactions in a VR simulation designed for embodied and enactive learning compared with traditional slideshow lecture-based instruction in terms of student learning outcomes across a number of cognitive (conceptual learning) and affective (intrinsic motivation, self-efficacy, perceived learning) measures. Thirty-eight undergraduates (17 females) who volunteered to participate in the study randomly received either the VR simulation (19), or slideshow-based traditional (18) instruction. Preliminary statistical analyses revealed that embodied and enactive interactions in VR: (i) had a significant positive impact on conceptual understanding in contrast to traditional instruction, and (ii) significantly helped in improving self-efficacy and confidence among the students, but was indifferent from the traditional instruction on student scores in intrinsic motivation and perceived learning tests.

**Keywords:** Virtual Reality, Embodied Learning, Science Education, 4E Cognition, Technology-enhanced Learning, Immersion

#### 1. Introduction

Immersive virtual reality (VR) – a technology that affords immersive, experiential, and real-life like interactions with dynamic virtual 3D models of objects in simulated environments – is becoming increasingly popular in STEM education (Radianti et al., 2020). VR allows learners to perceive, interact with, enact, and experience abstract STEM content in newer and innovative ways (Johnson-Glenberg, 2018; Skulmowski & Rey, 2018). However, research on the efficacy of VR as a tool for the learning-teaching of abstract scientific concepts is still evolving and varies considerably in relation to the educational level(s) at which VR was used (e.g. school, undergraduate; Jensen & Konradsen, 2018); the target content (e.g. visualization of abstract concepts; Checa et al., 2021), the degrees of VR's effectiveness (positive, mixed; Buttussi & Chittaro, 2017; Luo et al., 2021), and the nature of the VR experience (e.g. as facilitated by the hardware and software used; Luo et al., 2021).

In science education, the incorporation of VR as an instructional technology has generally led to favorable outcomes across several affective aspects of student learning (e.g. motivation, self-efficacy, enjoyment; Buttussi & Chittaro, 2018; Han, 2020; Pande et al., 2021) as well as knowledge-retention (e.g. Moro et al., 2017; Parong & Mayer, 2020), either in isolation, or in comparison with more traditional forms of instruction and/or other media. However, results related to content learning (e.g. particularly, conceptual understanding) have not been conclusive (e.g. Azer & Azer, 2016; Concannon et al., 2019; Klippel et al., 2020; Moro et al., 2017; Radianti et al., 2020; Zhao et al., 2021). Moreover, since majority of studies document the use of VR in controlled environments, the performance of VR in authentic education settings or classrooms, particularly over extended periods, remains

unclear (ecological contexts; Pande et al., 2021; Zhao et al., 2021). The bulk of studies that documented the instructional applications of VR over the last ten years are concentrated within the more applied domains in higher science education such as medical, dentistry, and engineering education and have largely focused on procedural learning, training, or skills acquisition (Roy et al., 2017; Tang et al., 2020). Finally, the reports of instructional uses and effectiveness of VR are predominantly from studies conducted in (relatively) richer countries and regions, and such research among socio-economically underprivileged populations in developing nations (e.g. Brazil, India, South Africa) is still scarce.

This paper reports results from a pre-test post-test quasi-experimental pilot study involving undergraduates, from predominantly socio-economically underprivileged student population, at a semi-urban university college in India, which investigated cognitive (conceptual understanding) and affective effects of learning university biochemistry topics through an interactive VR simulation as compared to traditional slideshow-based scripted lecture.

#### 2. Theoretical Framework: Embodied Learning

This study draws its theoretical inspiration from the recent cognitive mechanism-accounts of science learning, especially the embodied and enactive (or what are often collectively referred to as 4E) cognition perspectives. These accounts of STEM learning assert that, regardless of how complex the target content (e.g. complex scientific concepts) is, one's learning (of that content) is constitutive of the various physical interactions they have with the different interactable forms of that content (e.g. scientific models like Bohr atomic model, representations such as equations; Landy et al., 2014; Pande & Chandrasekharan, 2017 & 2022). Put differently, these approaches systematically illustrate how seemingly abstract learning (e.g. conceptual understanding) originates from "doing", or from active bodily participation that engages our sense and movement capacities in relation to the nature of the content (Nathan, 2021; Kersting et al., 2023; Pande, 2021).

In the VR and technology-enhanced learning contexts, embodied and enactive learning theories imply that richer sensorimotor interactions and realistic bodily engagement with various forms of the content further enhance the chances of effective learning of scientific concepts (McGowin et al., 2022; Lindgren & Johnson-Glenberg, 2013; Kothiyal et al., 2014; Skulmowski & Rey, 2018), particularly in relation to other less active and immersive modes of instruction.

Consistent with theories of embodied and enactive learning which primarily account for content learning, theoretical and empirical works on affective aspects of learning have long emphasized how active (e.g. hands-on or bodily) engagement and participation are positively intertwined with one's interest, intrinsic motivation, self-efficacy, overall enjoyment, and other non-cognitive factors (e.g. Bandura, 2001; Maresky et al., 2019). Representing scientific concepts in accessible forms that facilitate bodily interaction, or actually doing or enacting a certain scientific phenomenon (e.g. chemical interactions between biomolecules) or set of procedures (e.g. lab protocol) that are otherwise difficult to access and imagine in the real world helps invoke the feeling of active engagement. High degrees of immersion and sense of presence, as in case of VR, further help enhances this feeling (Büssing et al., 2022; Chessa et al., 2019; Johnson-Glenberg, 2018). For instance, when navigating VR and interacting with virtual objects, many VR users feel as if they were interacting with the real world. Considerable amount of research in learning with VR has demonstrated that such a feeling often results in an increased interest, intrinsic motivation, engagement, and enjoyment (Maresky et al., 2019; Stepan et al., 2017; Teranishi & Yamagishi, 2018), which in turn has been linked to higher self-confidence/efficacy (Jang, 2008; Kahu et al., 2017).

#### 3. The Study

This study explored whether enacting biochemical interactions in an immersive VR simulation designed for embodied learning yields better learning outcomes on abstract biochemistry concepts as compared to traditional slideshow-based traditional instruction.

A quasi-experimental two-group pretest–posttest design was employed.

A total of 38 undergraduate students (17 female) enrolled in various life sciences programs at a semi-urban university college, with a large proportion of students coming from socio-economically disadvantaged populations, in western India volunteered to participate in this study.



The study was conducted in accordance with the Declaration of Helsinki and the general ethical guidelines of the university/college where the study participants were recruited. All participating students were adults and were informed about the study's objectives prior to recruitment. Written consents were obtained from the participating students before the study.

Figure 1 above captures the study protocol. Due to the limited availability of VR equipment, the study was performed over a stretch of six days. On the first day, two sets of unique codes (e.g. VR1, VR2; SS1, SS2, etc.) were randomly distributed among all the 38 students. A code beginning with VR would assign a student to the VR biochemistry simulation (experimental/treatment) group (total 19 students; 9 female) whereas a code with SS would assign the student to the slideshow-based traditional lecture group (total 19 students; 8 female). The first day was also used to create a schedule that accommodated student availability. Between days 2-6, 3-4 pairs of students (with each pair composed of a VR student, and a slideshow lecture student) would visit the study site at different hours as scheduled. At the site they would: (1) first respond to a pre-test comprising of a brief demographic survey (age, gender, VR experience, major), a set of biochemistry questions (11 multiple-choice items to capture their conceptual understanding of the topic of the lesson) and affective questions deployed as 7-point Likert scales on intrinsic motivation (8 items; Monteiro et al., 2015) and self-efficacy (6 items; Makransky et al., 2016). After pretest, the VR student in the pair would go to another room where they would put on the VR HMD and play the pre-loaded biochemistry simulation. The other student from the pair would, at the same time, receive a one-on-one detailed slideshow-based scripted lecture delivered by a university teacher presenting the same biochemistry content as the simulation.

On completion of their respective instruction, each student responded to a posttest comprising of the same set of items as in the pre-test plus 5 additional items on a 7-point perceived learning Likert scale (Lee et al., 2010). In addition, students in the VR group answered 18 items (7-point Likert scale; adapted from Franco & Peck, 2018) that measured

the extent of embodiment they experienced when playing the simulation (analysis not included in this report).

#### 3.1 Instructional material

A highly interactive VR simulation supporting the embodied learning and conceptual understanding of the biochemical mechanisms (specifically Nucleophilic Substitution of SN1 and SN2 reactions) involved in an enzyme-substrate interaction was designed and developed as part of a larger design-based research preceding this study.

Specifically, this VR simulation narrates the classic case of lysozyme – an enzyme/protein in human mucosal secretions that neutralizes harmful bacteria by breaking down peptidoglycan molecules in their cell walls. Through visual prompts (flashing arrows), text-based instructions, and dynamic haptic feedback (vibration), the VR simulation we designed allows a learner to (a) embody, i.e. become, lysozyme in its molecular form (visible as molecular arms whose movement is synced with the learner's; see figure 2a), (b) zoom into bacterial cell wall molecular structures (see Figure 2b), and (c) enact and learn about two types of reaction mechanisms, nucleophilic substitution 1 and 2 (SN1 and SN2), comprising lysozyme's biochemical function (see Figure 2c). The learners also see dynamic graphs and chemical equations that reflect the changing states of the biochemical reaction in real-time (see Figure 2d).



*Figure 2.* Snippets from the VR simulation: (a) On the right, a learner can be seen lifting their arms (controllers) to embody (i.e. become) the enzyme. On the left is the learner's view of

embodied enzyme arms in a skeleton/stick molecular model. (b) A zoomed-in colorful molecular view of peptidoglycan polymer bacterial cell wall (with learner's view of embodied enzyme arms in a skeleton/stick molecular model in the foreground). (c) Learner's enaction of the two chemical reaction mechanisms using two different body movements: (Top) Nucleophilic Substitution 1 (SN1) reaction by an abrupt and jerky twisting the right hand/controller in clockwise direction. (Bottom) Activation of Nucleophilic Substitution (SN2) reaction mechanism by abruptly twisting both hands/controllers in anti-clockwise direction. (d) Dynamic graph and chemical equations as seen on a white panel in the background.

The VR simulation was deployed using an HTC Vive Cosmos headset tethered to a DELL gaming laptop.

The traditional slideshow-based traditional instruction presented this lysozymepeptidoglycan interaction, biochemical mechanisms, and other relevant concepts as covered in the VR simulation through a set of static diagrams, animations, and textual information in a self-paced format (i.e. students could choose to view/review the slides at their will), accompanying a scripted verbal lecture. The lecturing teacher was specifically trained to avoid using iconic gestures and/or body language related to the presented content/concepts, and strictly adhere to the verbal instruction provided in the script. This would prevent the students in this group from getting additional cues or support in imagining concerned biochemical phenomena/concepts.

#### 3.2 Expected Outcomes

Following from the theoretical framing of embodied, enactive, and action-based learning, we expected the VR simulation to outperform slideshow-based lecture instruction by helping the students dynamically visualize, enact, and experience the enzyme-substrate interactions and thus conceptually understand the content better. Similarly, we hoped that this would make the VR group more motivated as well as self-confident for navigating and learning about the complex/abstract biochemical concepts after the intervention in comparison to the other instruction (where students would be unsupported in imagining how lysozyme-peptidoglycan interactions would "look" and "feel" like). We also expected the average perceived learning scores for the VR group to be higher than those for the slideshow lecture group.

#### 3.3 Analysis

One student identifying themselves as a man dropped out from the VR group soon after trying the simulation for several minutes due to cybersickness-related issues. This rendered the effective sample size to a total of 37 participants (19 and 18 in the VR and Slideshow-based lecture groups, respectively).

All statistical analyses were performed using IBM SPSS Statistics 28.

First, descriptive statistics were calculated for pre and post test scores (independent and dependent variables) for each group. Baseline tests showed that the groups did not differ significantly in terms of any of the collected demographic variables (age, gender, prior VR experience).

ANCOVA was employed to compare means of the two study groups to assess their comparative learning effectiveness, while controlling for the groups' respective pre-test performances to make sure that the differences, if any, did not arise from pre-existing group differences, across the mentioned conceptual understanding and affective learning outcomes. Assumptions of ANCOVA were assessed, including linearity, homogeneity of regression slopes, independence, and normality. These assumptions were found to be sufficiently met for the analysis.

The main ANCOVA analyses were conducted with each dependent variable (i.e. posttest scores for conceptual understanding, intrinsic motivation, and self-efficacy) as the outcomes, controlling for the respective covariate (i.e. pre-test scores for conceptual understanding, intrinsic motivation, and self-efficacy, respectively).

Finally, an independent samples t-test was performed to compare the mean perceived learning scores of the two groups. Analyses of VR group's responses to the self-report embodiment questionnaire are not included in this paper.

#### 4. Results

#### 4.1 Descriptive statistics

The means and standard deviations are presented in Table 1.

Table 1. Descriptive statistics

	Variables	Pre-	test	Post-test	
	variables	Mean	SD	Mean	SD
VR	Conceptual understanding	0.21	0.14	0.45	0.22
	Intrinsic motivation	4.41	0.89	4.99	1.09
	Self-efficacy	4.23	0.72	5.73	0.81
	Perceived Learning	-	-	5.37	1.18
Slideshow-	Conceptual understanding	0.22	0.13	0.27	0.17
lecture	Intrinsic motivation	4.53	0.65	4.87	0.99
	Self-efficacy	4.05	0.41	5.02	0.96
	Perceived Learning	-	-	5.16	1.15

#### 4.2 ANCOVA Results

On conceptual understanding-related learning outcomes, the ANCOVA model revealed a significant effect of group on the dependent variable, mean post-test conceptual learning score (F(1, 34) = 5.884, p = 0.021), while controlling for the covariate, mean pre-test conceptual learning score. The effect size as indicated by the partial eta squared was considerably big with a value of 0.148.

For intrinsic motivation, ANCOVA did not result in a significant group effect for the dependent variable, mean post-test intrinsic motivation rating (F(1, 34) = 0.445, p = 0.509), while controlling for the covariate, mean pre-test intrinsic motivation rating.

For self-efficacy, the ANCOVA model revealed a significant effect of group on the dependent variable, mean post-intervention self-efficacy rating (F(1, 34) = 5.086, p = 0.031), when controlling for the covariate, mean pre-test self-efficacy rating before the intervention. The effect size as indicated by the partial eta squared was considerably big with a value of 0.13.

Finally, a two-tailed independent samples t-test assuming equal variances revealed that the two groups did not differ in terms of post-intervention perceived learning scores at p = 0.582.

#### 5. Summary and Conclusion

This research examined the learning effectiveness of a highly interactive VR simulation designed to support embodied learning of complex biochemical phenomena and concepts

through enactive experiences of certain chemical interactions between two biomolecules. The simulation learning-teaching efficacy was tested in contrast to the traditional PowerPoint slideshow-based instruction concerning the same conceptual content.

ANCOVA analyses were employed to allow for a more robust assessment of the instructional impact of embodied and enactive VR interactions. Additionally, a t-test was used to evaluate differences between groups on post-intervention perceived learning scores.

Conceptual understanding-related results underscore the effectiveness of embodied interactions in VR to foster enhanced conceptual understanding among learners. The substantial observed effect size ( $\eta^2 = 0.148$ ) suggested that embodied VR interactions had a noteworthy impact on post-test conceptual learning scores; thus, strongly aligning this finding with theoretical expectations. It is probable that deliberately controlling the slideshow group's exposure to embodied cues (e.g. natural/accidental iconic gestures performed by a teacher during instruction) by training the teacher to stick to the verbal script contributed to this sizable group effect. This opens up a potential research direction.

The analysis did not yield a significant group effect on post-test intrinsic motivation ratings, implying that VR and slideshow-based scripted lecture did not affect participants' intrinsic motivation directly to the same extent. Nor was there any group difference in terms of post-intervention perceived learning. Yet, there was a significant effect of group on post-intervention self-efficacy ratings, particularly suggesting that VR positively impacted participants' self-efficacy beliefs and their confidence. The effect size ( $\eta^2 = 0.13$ ) indicated a noteworthy magnitude of the VR's impact on post-intervention improvements in self-efficacy. It is not clear why results on intrinsic motivation and perceived learning were inconsistent with our expectations, particularly considering that interest and intrinsic motivation are often believed to be the roots of self-efficacy (Jang, 2008; Kahu et al., 2017); previous work has demonstrated an overall positive relationship between VR experience and affective engagement in comparison to traditional instructional methods.

This preliminary set of analyses merely scratch the surface of the complex relationships between VR-facilitated embodied and enactive interaction, conceptual understanding, and affective aspects related to learning. Further analyses of data may yield important insights into these relationships, and help provide a comprehensive picture on embodied VR's instructional effectiveness. More probing is needed specifically to understand how enactment in VR is affecting intrinsic motivation and perceived learning, and if prior VR experience and/or novelty effect are playing any roles in such settings.

#### Acknowledgements

Sincere thanks to my former research team (Biljana Mojsosksa, Morten Erik Moeller, Praveen Ramasamy, Per Meyer Jepsen, Jonathan Hvass) at Roskilde University Denmark for collaborating with me on the design and development of the embodied VR simulation used in this study. Special thanks to Søren Larsen for his administrative support during the simulation design and development. Special thanks to the undergraduate students for participating in this study, as well as their amazing professors and college staff for making this pilot study possible.

#### References

- Azer, S. A., & Azer, S. (2016). 3D Anatomy Models and Impact on Learning: A Review of the Quality of the Literature. *Health Professions Education*, *2*(2), 80–98. https://doi.org/10.1016/j.hpe.2016.05.002
- Bandura, A. (2001). Social Cognitive Theory: An Agentic Perspective. *Annual Review of Psychology*, 52(1), 1–26. https://doi.org/10.1146/annurev.psych.52.1.1
- Buttussi, F., & Chittaro, L. (2018). Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE Transactions on Visualization and Computer Graphics*, 24(2), 1063–1076. <u>https://doi.org/10.1109/TVCG.2017.2653117</u>

- Büssing, A. G., Thomsen, D., & Braun, N. (2022). Linking technology usage to instructional quality: Immersive virtual reality increases presence, positive emotions, and cognitive activation. Authors Preprint Accessed at: https://osf.io/preprints/p9esz/.
- Checa, D., Miguel-Alonso, I., & Bustillo, A. (2021). Immersive virtual-reality computer-assembly serious game to enhance autonomous learning. *Virtual Reality*. <u>https://doi.org/10.1007/s10055-021-00607-1</u>
- Concannon, B. J., Esmail, S., & Roduta Roberts, M. (2019). Head-Mounted Display Virtual Reality in Post-secondary Education and Skill Training. *Frontiers in Education*, *4*. <u>https://doi.org/10.3389/feduc.2019.00080</u>
- Gonzalez-Franco, M., & Peck, T. C. (2018). Avatar embodiment. towards a standardized questionnaire. *Frontiers in Robotics and AI*, 5, 74
- Han, I. (2020). Immersive virtual field trips in education: A mixed-methods study on elementary students' presence and perceived learning. *British Journal of Educational Technology*, *51*(2), 420–435. https://doi.org/10.1111/bjet.12842
- Jang, H. (2008). Supporting students' motivation, engagement, and learning during an uninteresting activity. *Journal of Educational Psychology*, *100*(4), 798–811. https://doi.org/10.1037/a0012841
- Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, *23*(4), 1515–1529. https://doi.org/10.1007/s10639-017-9676-0
- Johnson-Glenberg, M. C. (2018). Immersive VR and Education: Embodied Design Principles That Include Gesture and Hand Controls. *Frontiers in Robotics and AI*, 5. <u>https://doi.org/10.3389/frobt.2018.00081</u>
- Kahu, E., Nelson, K., & Picton, C. (2017). Student interest as a key driver of engagement for first year students. *Student Success*, *8*(2), 55–66. https://doi.org/10.5204/ssj.v8i2.379
- Kersting, M., Amin, T. G., Euler, E., Gregorcic, B., Haglund, J., Hardahl, L. K., & Steier, R. (2023). What is the role of the body in science education? A conversation between traditions. *Science & Education*, 1-40.
- Klippel, A., Zhao, J., Oprean, D., Wallgrün, J. O., Stubbs, C., la Femina, P., & Jackson, K. L. (2020). The value of being there: toward a science of immersive virtual field trips. *Virtual Reality*, *24*(4), 753–770. <u>https://doi.org/10.1007/s10055-019-00418-5</u>
- Kothiyal, A., Majumdar, R., Pande, P., Agarwal, H., Ranka, A., & Chandrasekharan, S. (2014). How does representational competence develop? Explorations using a fully controllable interface and eye tracking, In C.-C. Liu, Y. T. Wu, T. Supnithi, T. Kojiri, H. Ogata, S. C. Kong & A. Kashihara (Eds.), *Proceedings of the 22<sup>nd</sup> International Conference on Computers in Education,* 738-743. Japan: APSCE.
- Landy, D., Allen, C., & Zednik, C. (2014). A perceptual account of symbolic reasoning. *Frontiers in psychology*, *5*, 275.
- Lee, E. A. L., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education*, 55(4), 1424-1442.
- Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by Embodiment. *Educational Researcher*, 42(8), 445–452. <u>https://doi.org/10.3102/0013189X13511661</u>
- Luo, H., Li, G., Feng, Q., Yang, Y., & Zuo, M. (2021). Virtual reality in K-12 and higher education: A systematic review of the literature from 2000 to 2019. *Journal of Computer Assisted Learning*, 37(3), 887–901. https://doi.org/10.1111/jcal.12538
- MCGOWIN, G., FIORE, S. M., & ODEN, K. Towards a Theory of Learning in Immersive Virtual Reality: Designing Learning Affordances with Embodied, Enactive, Embedded, and Extended Cognition. *Bridging the XR Technology-to-Practice Gap: Methods and Strategies for Blending Extended Realities into Classroom Instruction*, 35.
- Maresky, H. S., Oikonomou, A., Ali, I., Ditkofsky, N., Pakkal, M., & Ballyk, B. (2019). Virtual reality and cardiac anatomy: Exploring immersive three-dimensional cardiac imaging, a pilot study in undergraduate medical anatomy education. *Clinical Anatomy*, *32*(2), 238–243. https://doi.org/10.1002/ca.23292
- Makransky, G., Thisgaard, M. W., & Gadegaard, H. (2016). Virtual Simulations as Preparation for Lab Exercises: Assessing Learning of Key Laboratory Skills in Microbiology and Improvement of Essential Non-Cognitive Skills. *PLOS ONE*, *11*(6), e0155895. https://doi.org/10.1371/journal.pone.0155895
- Monteiro, V., Mata, L., & Peixoto, F. (2015). Intrinsic Motivation Inventory: Psychometric Properties in the Context of First Language and Mathematics Learning. *Psicologia: Reflexão e Crítica*, 28(3), 434–443. https://doi.org/10.1590/1678-7153.201528302

- Moro, C., Štromberga, Z., Raikos, A., & Stirling, A. (2017). The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anatomical Sciences Education*, *10*(6), 549–559. https://doi.org/10.1002/ase.1696
- Nathan, M. J. (2021). Foundations of embodied learning: A paradigm for education. Routledge.
- Pande, P. (2021). Learning and Expertise with Scientific External Representations: An Embodied and Extended Cognition Model, *Phenomenology and the Cognitive Sciences*, (20), 463-482. doi: https://doi.org/10.1007/s11097-020-09686-y
- Pande, P., & Chandrasekharan, S. (2017). Representational competence: towards a distributed and embodied cognition account, *Studies in Science Education*, 53(1), 1-43. UK: Routledge. https://doi.org/10.1080/03057267.2017.1248627
- Pande, P. & Chandrasekharan, S. (2022). Expertise as Sensorimotor Tuning: Perceptual Navigation Patterns Mark Representational Competence in Science, *Research in Science Education*, 52, 725-747, https://doi.org/10.1007/s11165-020-09981-3
- Pande, P., Thit, A., Sørensen, A. E., Mojsoska, B., Moeller, M. E., & Jepsen, P. M. (2021). Long-term effectiveness of immersive vr simulations in undergraduate science learning: Lessons from a media-comparison study. *Research in Learning Technology*, 29. https://doi.org/10.25304/rlt.v29.2482
- Parong, J., & Mayer, R. E. (2020). Cognitive consequences of playing brain-training games in immersive virtual reality. *Applied Cognitive Psychology*, 34(1), 29–38. https://doi.org/10.1002/acp.3582
- Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778. <u>https://doi.org/10.1016/j.compedu.2019.103778</u>
- Roy, E., Bakr, M. M., & George, R. (2017). The need for virtual reality simulators in dental education: A review. *The Saudi Dental Journal*, 29(2), 41–47. <u>https://doi.org/10.1016/j.sdentj.2017.02.001</u>
- Skulmowski, A., & Rey, G. D. (2018). Embodied learning: introducing a taxonomy based on bodily engagement and task integration. Cognitive research: principles and implications, 3(1), 1-10.
- Stepan, K., Zeiger, J., Hanchuk, S., del Signore, A., Shrivastava, R., Govindaraj, S., & Iloreta, A. (2017). Immersive virtual reality as a teaching tool for neuroanatomy. *International Forum of Allergy & Rhinology*, 7(10), 1006–1013. https://doi.org/10.1002/alr.21986
- Tang, K. S., Cheng, D. L., Mi, E., & Greenberg, P. B. (2019). Augmented reality in medical education: a systematic review. *Canadian Medical Education Journal*. <u>https://doi.org/10.36834/cmej.61705</u>
- Teranishi, S. & Yamagishi, Y. (2017). Educational Effects of a Virtual Reality Simulation System for constructing Self-Built PCs. In J. Johnston (Ed.), *Proceedings of EdMedia 2017* (pp. 1214-1222).
  Washington, DC: Association for the Advancement of Computing in Education (AACE).
- Zhao, G., Fan, M., Yuan, Y., Zhao, F., & Huang, H. (2021). The comparison of teaching efficiency between virtual reality and traditional education in medical education: a systematic review and meta-analysis. *Annals of Translational Medicine*, *9*(3), 252–252. https://doi.org/10.21037/atm-20-2785

# W06

# The 16th Workshop on Technology Enhanced Learning by Posing/Solving Problems/Questions

# Evaluation of an Automatic Generation System for Tracing Tasks Based on Textbook Programs

Tomohiro MOGI<sup>a\*</sup>, Yuichiro TATEIWA<sup>b</sup>, Takahito TOMOTO<sup>c</sup> & Takako AKAKURA<sup>d</sup>

<sup>a</sup> Graduate School of Information and Computer Science, Chiba Institute of Technology, Japan

<sup>b</sup> Graduate School of Engineering, Nagoya Institute of Technology, Japan
<sup>c</sup> Faculty of Information and Computer Science, Chiba Institute of Technology, Japan
<sup>d</sup> Faculty of Engineering, Tokyo University of Science, Japan
\* tomo1238hi@gmail.com

**Abstract:** In programming education, it is important not only to create programs but also to reflect on the created programs in order to understand them more deeply. To that end, the authors developed a system that automatically generates tracing tasks from input programs. In this study, we used programs from a textbook as input in order to confirm that the automatic problem and feedback generation system can generate tracing tasks automatically without any problems. The results of the experiment showed that the system was able to automatically generate tracing tasks from 53.85% of the input program elements, and it was confirmed that functions such as 'for' and 'if' statements as well as assignments to variables were supported without issue. However, orthodox syntax and functions such as 'while' and 'switch' statements, as well as batch assignments to structures and arrays, were not correctly implemented, so it will be necessary to correct these implementations in the future.

Keywords: Automatic Problem Generation, Program Trace Learning

#### 1. Introduction

In recent years, the demand for programming education has increased. In general programming education, the teacher presents a task, and the learner creates a program to perform to the task. If a learner makes a mistake while creating the program, they can enhance their learning by identifying the mistake and correcting it. However, it is not always easy for learners to identify their mistakes or fix them.

Typical learning flows during a programming learning activity are shown in Figure 1. In normal programming learning, The learner receives the task (requirement) and creates a program satisfying the received task. The learners necessity to consider the expected results as they develop their programs. The learners (1) observe the execution results by running the program they have created, and then (2) detect the difference between the execution results and the expected results (correct answers). If the execution result differs from the expected result, the learner (3) looks back at the program code and corrects any mistakes, but it can be difficult to correct the program code simply by observing the execution results. Therefore, in advanced programming learning (i.e., trace learning: To consider the order in which programs are executed, the value of variables, output and conditional branching.), a debugger is used to trace the learner's code, making it easier to correct the mistake in the program code. In advanced programming learning, in addition to the normal programming learning flow, the learner (i) observes the execution process and the local results generated by the debugger, and (ii) detects the differences between the expected execution process

and the local results. In this way, trace learning makes it easier to modify the program compared with merely observing the execution results, thereby realizing more effective learning. However, because learners must learn how to use the debugger effectively, an environment in which feedback can be given is considered necessary.

We previously developed an automatic problem and feedback generation system (Mogi et al., 2023). By using this system to generate tracing tasks from programs created by learners, it is possible to quickly implement advanced programming learning using a debugger because (i)' Observe the process of execution and local results can be practiced and (ii)' Recognize differences from expected results can be visualized.

In this paper, we conduct experiments involving the automatic generation of tracing tasks, based on programs from a programming textbook (Shibata, 2021) in order to evaluate whether the developed system is able to generate appropriate tracing tasks.



Figure 1. Learner's activities in programming learning

#### 2. Related Work

In general programming education, emphasis is placed on the creation of programs, while activities involving reading programs are rarely conducted. However, in actual programming situations, the ability to read programs is also important. Also, when modifying (debugging) a program that you have created, it is necessary to understand the behavior of the program in detailed units. Several studies have proposed learning support aimed at developing the ability to read programs. These studies include support to estimate the functional requirements of programs (Kanamori et al., 2013; Arai et al., 2014), support to read the semantic coherence of programs (Watanabe et al., 2015), and support to make dedicated program tracing activities (Tomoto & Akakura, 2017). However, these studies dealt only with problems previously registered in the system and could not realize the advanced learning illustrated in Fig. 1.

#### 3. Program Tracing Task

The program tracing task asks the learner about the order of the lines to be executed and the content of the instructions to be executed in each line, with the aim of making the learner consider the behavior of the program. In this task, the learner is first given the source code. Next, the learner is asked to select the first line in the source code to be executed. Then, to make the learner think about the instructions to be executed on that line, the learner is asked to describe the change in the value of a variable and the content of the output. In repetitive processing and conditional branching, the program processing flow changes depending on the success or failure of the condition, so the program is also required to describe the

success or failure of the condition. In the same way, the learner selects the next line and describes the behavior of the program, repeating this process until the end of the program. In the case of sequential structures, the order of the lines in the source code and the order of execution are the same, but in the case of iterations, conditional branches, or functions, line jumps can occur, and thus the task is to check whether the processing flow can be traced appropriately. Although experienced users might perform these tasks on a daily basis, novice users may be unaware of these tasks or might not be able to perform them properly and thus require guidance.

It is assumed that novice trainees may make mistakes in the order of execution and in the input of the contents of each line of processing. Although the program tracing task itself can be carried out on paper, it is possible to proceed without realizing where mistakes have been made when learning by oneself on paper or by consulting reference books. In such cases, the error may not be noticed until the end, or the error may be noticed by looking at the correct answer at the final stage. However, because the correct answer is known when the error is noticed, the opportunity to correct the error oneself is lost. Even if the teacher provides guidance on correcting the error in an educational setting, sufficient support cannot be expected, given the one-to-many nature of the classroom. Some ordinary debuggers are equipped with a trace function, which can be used for learning. Many of these functions allow tracing while stopping the program line by line, and may even allow the values of variables in each line and the screen output to be known. However, even when learning using a debugger, the correct answer is known immediately, which precludes the possibility of noticing errors. Therefore, in this study, we considered it important to have the learner perform the tracing task and then provide feedback.



#### 4. Automatic Problem and Feedback Generation System

Figure 2. Interface of the Automatic Problem and Feedback Generation System

The system interface is shown in Figure 2. The system is largely divided into two phases: a phase for registering programs and automatically generating tracing problems, and a phase for working on the automatically generated tracing tasks. The learner first moves to the program registration screen by pressing the 'Register program' button shown in Fig. 2(I) and begins the first phase (registering a program and automatically generating tracing problems). On the program registration screen, the title and program are entered and the tracing problem is automatically generated by pressing the 'Register' button (Fig. 2(II)). After

registering the problem, the learner moves to the tracing screen by pressing the 'Exercise' button (Fig. 2(III)) at the top of the screen.

The learner then enters the second phase (working on the automatically generated tracing tasks). The program and the term-by-term options are presented on the left-hand side of the tracing screen. The learner considers which terms from the presented program are to be executed and in which order. If clicking on a selectable part changes a variable in that part, the learner will be asked to enter its value, using the keyboard (Fig. 2(IV)). For example, if the learner selects the term "i = 1", they would be asked "What is the value for i?" Because the value of i in this case is 1, they must enter "1". In the case of an 'if statement, both the value of the conditional judgement and the judgement result (T or F) are required, while in the case of a 'for' statement, the value and judgement result are required for the initial value, the conditional judgement, and the increment.

The system has a function for displaying the values of variables as feedback to the learner. The value based on the learner's answer is displayed at the top center of the screen, and the value in the case of a correct answer is displayed at the bottom center of the screen. The system also generates feedback automatically by analyzing the program, so that the value of the correct answer can be displayed without the teacher having to enter the correct answer beforehand. The learner aims to correct errors by observing the difference between the expected and executed results.

Here, we describe the flow when the learner's answer is correct is explained, using the program shown in the Fig. 2 as a concrete example. The learner first considers which part of the program is executed first. The choices are the parts with highlighted in light brown in this example, the six choices are ""i = 1"", ""i <= 4"", ""i++"", ""i %2 == 0"", ""print(""\%d"",i)"" and ""print(""\*"")"". The first part to be executed is considered to be ""i = 1"", so the learner clicks on that part. Then, as shown in (IV), the system asks ""What is the value for i?" Therefore, the learner enters "1" on the keyboard and presses the Enter key. This operation corresponds to the answer ""1 is assigned to the value of i"" and is therefore indicated as ""i=1"" in the variable summary of your the situation. Because this answer is equal corresponds to the actual system flow (correct answer), the variable summary of the correct situation changes in the same way.

Next, we describe the flow when the learner's answer is incorrect. For example, if the learner responds to the question "What value goes in i?" by entering '2', then the variable overview of the situation will show "i=2". The fact that this display differs from the variable summary of the correct situation, "i=1", allows the learner to realize their error. If the learner answered "*print("\*"*)" for the part that is executed first, a "*\**" would appear in the standard output of the situation, drawing attention to error.

#### 5. Evaluation

This paper describes an experiment conducted to investigate whether the developed automatic problem and feedback generation system is capable of generating appropriate tracing tasks. In the experiment, tracing tasks were automatically generated, using a program from a programming textbook (Shibata, 2021) as input.

#### 5.1 Method

Because there are 13 chapters in the textbook (Shibata, 2021), a total of 39 programs, three from each chapter, were extracted for the experiments. In addition, characters and functions that are not supported by the system were modified. First, the extracted programs contained multibyte characters, but this system does not support 2-byte characters, so they were changed to 1-byte characters. Next, because the system does not support the *scanf* function, it was changed to an appropriate numeric or character assignment. We manually imported the program with these changes into the system and confirmed that the tracing task could be generated automatically.

#### 5.2 Results

Table 1 summarizes the elements of the programs used and whether or not tracing tasks could be generated automatically for these elements.

Table 1. Elements of the programs used and whether tracing tasks could be generated automatically

Program elements	Number of programs included	Number of generated questions
printf function	39/39 (100%)	21/39 (53.85%)
puts function	13/39 (33.33%)	6/13 (46.15%)
sizeof function	1/39 (2.56%)	1/1 (100%)
atoi function	1/39 (2.56%)	1/1 (100%)
fclose function	3/39 (7.69%)	0/3 (0%)
user-defined function	12/39 (30.77%)	4/12 (33.33%)
four arithmetic operations	14/39 (35.9%)	11/14 (78.57%)
ternary operator	4/39 (10.26%)	2/4 (50%)
assignments to variables	23/39 (58.97%)	13/23 (56.52%)
Assignments to arrays	7/39 (17.95%)	1/7 (14.29%)
struct	3/39 (7.69%)	0/3 (0%)
if statement	10/39 (25.64%)	3/10 (30%)
for statement	14/39 (35.9%)	5/14 (35.71%)
while statement	9/39 (23.08%)	0/9 (0%)
switch statement	2/39 (5.13%)	0/2 (0%)
all	-	21/39 (53.85%)

Tracing tasks could be automatically generated for only 53.85% of the program elements. The reasons for this are thought to be due to several unsupported syntax and functions. For example, the developed system does not support 'while' and 'switch' statements as well as batch assignments to structures and arrays. Table 1 also shows that tasks for elements that include these statements cannot be generated automatically. Tracing tasks for functions such as for statements (35.71%) and if statements (30%), as well as assignments to variables (56.52%), had a low generation rate and at first glance appeared to be unsuccessful. This is because unsupported syntax and functions such as switch statements and array batch assignments tend to occur at the same time in these syntax and functions. Therefore, it was confirmed that the cause of failure was due to the currently unsupported syntax. And that, if did not contain unsupported syntax and functions was 100% generating of the programs successfully generated. It was also confirmed that tracing tasks for programs with inappropriate line breaks and indentation could also be generated automatically without problems. For these reasons and the fact that experiments in Mogi et al. (2023) confirmed the learning effect, we believe that the this system has a certain value. However, to realize debug-style learning using this system, it is necessary to be able to handle orthodox syntax and functions such as 'while' and 'switch' statements as well as batch assignment to structures and arrays. We do not consider the syntax and functions to be beyond the scope of the system, but rather the problem is that they are not implemented correctly. Therefore, we plan to correct their implementation in the future.

#### 6. Conclusion

In programming learning, it is important to acquire the ability to understand how a program works (tracing ability). We previously developed a system (Mogi et al., 2023) that realizes this aim as a learning method for acquiring tracing ability.

In the present paper, we conducted experiments to confirm that the system can be used for learning with debuggers. The experiments used programs from textbooks as input and investigated the degree to which the system could automatically generate tracing tasks based on these programs. The results showed that tracing tasks could be generated for 53.85% of the program elements. However, it was confirmed that orthodox syntax and functions such as 'while' and 'switch' statements as well as batch assignment to structures and arrays were not correctly implemented.

This study aims to improve the programming skills of learners. In particular, we focus on the ability to trace a program by statement. However, we consider it necessary to start working on other areas in the future.

When expert programmers debug a program, they first recognize the program in large block units and roughly identify the cause of problems by tracing. Then, by tracing in more detail, they identify and fix the cause of the problem. Expert programmers perform hierarchical debugging in this way, but this system does not support hierarchical debugging. However, we consider that the skill to recognize a program as a large block and debug it is important and should be considered in the future.

Elsewhere, The current system only uses values embedded in the code by the learner for tracing. However, for the learner to understand the behaviour of the program, the test data(input) needs to be appropriate. For example, in the problem "output the larger value of variables a and b", it is not possible to judge that it is an appropriate program with only "data in which b is larger than a", but at least "data in which a is larger than b" or "data in which a and b are the same size" etc. are considered necessary. Therefore, it is necessary to consider test data in the future.

#### Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers JP22K12322 and JP20H01730.

#### References

- Arai, T., Kanamori, H., Tomoto, T., Kometani, Y., & Akakura, T. (2014). Development of a learning support system for source code reading comprehension. Lecture Notes in Computer Science, 12–19. https://doi.org/10.1007/978-3-319-07863-2\_2
- Kanamori, H., Tomoto, T., & Akakura, T. (2013). Development of a computer programming learning support system based on reading computer program. Human Interface and the Management of Information. Information and Interaction for Learning, Culture, Collaboration and Business, 63– 69. https://doi.org/10.1007/978-3-642-39226-9\_8
- Mogi, T., Tateiwa, Y., Tomoto, T., & Akakura, T. (2023). Proposal for Automatic Problem and Feedback Generation for use in Trace Learning Support Systems. Lecture Notes in Computer Science, 310–321. https://doi.org/10.1007/978-3-031-35129-7\_23
- Shibata, B. (2021). Shin Meikai Cgengo Nyumonhen [New lucidity C Programming Language Introductory] (Vol. 2). SB Creative.
- Tomoto, T., & Akakura, T. (2017). Report on practice of a learning support system for reading program code exercise. Human Interface and the Management of Information: Supporting Learning, Decision-Making and Collaboration, 85–98. https://doi.org/10.1007/978-3-319-58524-6 8
- Watanabe, K., Tomoto, T., & Akakura, T. (2015). Development of a learning support system for reading source code by Stepwise Abstraction. Lecture Notes in Computer Science, 387–394. https://doi.org/10.1007/978-3-319-20618-9\_39
# Development of a Semi-Active Learning Support System with Operation Index for the Mathematics of Vectors

Tomohito JUMONJI<sup>a\*</sup>, Nonoka AIKAWA<sup>a</sup> & Takahito TOMOTO<sup>c</sup>

<sup>a</sup>Graduate School of Engineering, Tokyo Polytechnic University, Atsugi, Kanagawa, Japan <sup>b</sup> Faculty of Information Science, Chiba Institute of Technology, Narashino, Chiba, Japan \*tomohito067@gmail.com

**Abstract:** It is important for learners to understand the relationship between mathematical problems and their answers to those problems. One way for learners to recognize the significance of their answers method for a problem is through trial and error. However, in ordinary mathematics learning, the correct answer is generally fed back from the textbook or teacher after the problem is attempted, which does not give learners an opportunity to reflect on their answers. So far, we have developed a system that encourages trial and error by visualizing the learner's answer for the mathematics of vectors and having them interact with it. However, there was a possibility that some learners would only look at the visualized figures to determine correctness and would not reflect on the errors they were making. In this study, we proposed and developed a system that requires learners to interact with figures on the system side, thereby promoting greater understanding of the relationship between mathematical problems and their answers.

Keywords: high-school mathematics, mathematics of vectors, error visualization

## 1. Introduction

It is important for learners to understand the relationship between mathematical problems and their answers to those problems (Hayashi et al., 2021; Enomoto et al., 2018; Yang et al., 2014). Trial and error is one way for learners to recognize the significance of their answers method for the problems. However, in typical learning facilitated by textbooks and classroom teaching, if a learner gives an incorrect answer, they are then simply given (e.g., by the teacher or textbook) the correct answer and its explanation, and there is little adaptive feedback for the learner's specific error (Yang et al., 2012). If the instructor gives only correct answers and their explanations, the learner may become a "passive learner" who simply accepts the correct answers implicitly. This reduces the likelihood that they will reflect on their errors, and may encourage them to try to memorize the correct answers presented to them without understand how or why an error was made. Therefore, it is important to provide learners with an environment in which they can proactively reflect on their incorrect answers through trial and error.

To encourage learners to perform trial and error, it is necessary to provide them with information that will serve as clues to the correct answer, rather than just giving them the correct answer. However, just having them focus on the wrong part and giving them clues to the correct answer may make them correct what is wrong and how it is wrong, but it may not make them reflect on why it is wrong.

Hirashima et al. (2002) proposed the framework of error-based simulation (EBS) to enable error visualization. This presents the learner with an opportunity to notice the error in their answer by themselves, which in turn is expected to encourage them to reflect on why it is wrong and to engage in the trial-and-error process. Hirashima et al. developed EBS mainly for dynamics (Horiguchi et al., 2002; Hirashima et al., 1998), and verified its effectiveness through many classroom practice sessions (Hirashima et al., 2009; Horiguchi et al., 2014). In addition Kurokawa et al. (2018a; 2018b) developed and evaluated a system that encourages learners to perform trial and error by visualizing errors in loci in mathematics.

However, conventional dynamics EBS also does not guarantee that the learner will observe the strange behavior that is visualized. If they only check whether their answer is correct or incorrect, and do not fully reflect on their answer, then learning becomes passive. Not reflecting on the relationship between their answers and the visualized behaviors in the figures may result in learners not understanding why their answers are wrong.

Therefore, in this study, we propose a method that explicitly requires the learner to manipulate the figures, with the aim that they will become aware of how their answers relate to those figures. This method is expected to enable learners to "proactively" discover "why their answers are wrong" through operation of figures, rather than just passively accepting correct or incorrect answers. Since the activity itself is required by the system, we refer to it as "semi-active" in this study.

## 2. Error Visualization in Mathematics

## 2.1 Learning support targeting STEM education

There have been various studies on STEM education, Sengupta et al. (2018) took a discursive, perspectival, material, and embodied phenomenological approach to coding and computational thinking, rather than viewing it as the acquisition of computational logic and symbolic forms. Yang et al. (2012; 2014) supported learning in one-to-one classrooms using variation-based discovery learning and facilitating drawing representations, math representations, and answers explanations through peer instruction among students. Mehanovic et al. (2012) designed and investigated interactive learning environments to support mathematics learning, based on the claim that technological tools appropriately integrated into mathematical tasks can support understanding of a wide range of mathematical concepts. Slater et al. (2018) addressed the gap by investigating how motivational constructs manifest within online learning systems, measures of mathematical identity (self-concept, values, and interest in mathematics), and correlating them with behavior and performance within Reasoning Mind's Elementary Mathematics Foundations system. Wong et al. (2017) reviewed multiple aspects of the structure of motivation in mathematics learning in a technology enhanced learning (TEL) environment, not limited to a single motivational theory. Junus. (2018) investigated students' mathematical misconceptions using concept maps and discussion records regarding inner product spaces. Ya-Jing et al. (2017) investigated the impact of a collaborative peer evaluation system using concept maps on learning effectiveness in programming education. Nakamoto et al. (2021) generated a model using self-explanation and penstroke data to identify the causes of students' stuckness. Enomoto et al. (2018) developed and experimented with a learning support system for problem-posing exercises to promote the interpretation of relationships among mathematical formulas. Hayashi et al. (2021) proposed a practice of conversion between verbal, mathematical, and graphical representations for addition-subtraction sentence problems (Hayashi, Y., 2021).

Thus, learning support in STEM education is highly important, especially in the area of mathematics. In particular, how to provide feedback on math problems is an important consideration. An approach to make learners recognize their own misunderstandings and correct their errors is called "error visualization". Error visualization is a method of making errors visible by converting the learner's answers to other media. Error visualization encourages learners to correct their knowledge by making them aware of their own misunderstandings, thus enabling them to correct their own errors with a certain level of motivation. In this error visualization, the error visualization of dynamics is well known. It visualizes errors by transforming the learner's force inputs into phenomena. Therefore, this study examines the adaptation of error visualization in the mathematical domain. In particular, we consider the transformation of representations in mathematics to be important for error visualization.

### 2.2 Error visualization for linear algebra

As mentioned above, Kurokawa et al. (2018a; 2018b) developed and evaluated a system that encourages learners to perform trial-and-error by visualizing errors loci in mathematics. They focused on two of the five mathematical representations identified by Nakahara (Nakahara., 1995), namely, symbolic representations, which are representations of mathematical expressions often used in mathematics learning, and graphic representations, which are representations of figures and graphs. They developed a mathematical expression conversion system that encourages trial-and-error by changing the symbolic expression (mathematical equations) or graphic expression (graphs) that were given by learners into other expressions, thereby visualizing errors.

However, unlike in dynamics, the domain of mathematics is not concerned with phenomena, but with figures Unlike phenomena, there is no motion, and therefore it is usually impossible to generate behavior. EBS of dynamics, events are visualized as behaviors in situations such as "a block is pushed and sinks to the floor". However, the mathematical domain is composed of abstract concepts and operations, which must be independent of situations, and therefore cannot be visualized as in conventional EBS of dynamics Unlike in dynamics, the domain of mathematics is not concerned with phenomena, but with figures Unlike phenomena, there is no motion, and therefore it is usually impossible to generate behavior. EBS of dynamics, events are visualized as behaviors in situations such as "a block is pushed and sinks to the floor". However, the mathematical domain is composed of abstract concepts and operations, which must be independent of situations. and therefore cannot be visualized as in conventional EBS of dynamics. which made it more difficult to suggest errors to learners, which made it more difficult to suggest errors to learners. Their so-called mathematical representation transformation system visualizes graphs (graphic representations) according to mathematical rules called "constraints" that exist in the mathematical equations (symbolic representations) constructed by the learner, while at the same time suggesting errors to the learner by allowing operation of figures within the constraints that exist in the mathematical equations constructed by the learner.

As an example, suppose the correct answer to a question is "point P moves on y = 2x + 3," but the learner's answer is "point P moves on y = -2x + 3." When visualizing the correct answer, point P will move along the graph generated by the correct equation, but point P as determined by the learner's answer will move along the different graph -2x + 3. The range of motion of a point P on the graph of the system varies depending on the equation. The learner can check the range of movement by actually moving it on the system. Thus, by adjusting point P on the generated graph, the learner can confirm that the behavior of point P in the graph operating the correct answer differs from the behavior of point P in the graph operating them to recognize their error, and we expect the learner to engage in trial and error.

### 2.3 Error Visualization for the mathematics of vectors

We have previously developed a learning support system based on work by Kurokawa et al. (2018a; 2018c), using error visualization for mathematics of vectors and verified its learning effectiveness (Jumonji et al., 2022). In that study, we succeeded in visualizing errors by converting the direction, which is a characteristic of vectors, into a figure as an arrow.

However, as stated previously, since vectors involve only direction and magnitude, it is often not possible to uniquely plot a vector on a figure. For example, the component vector a=(3,4) is a vector with a magnitude and direction of 3 on the x-axis and 4 on the y-axis, but when placing it on a graph, it cannot be plotted uniquely because it has no restriction on the coordinate at which it is placed and can be placed at the origin O or at any other coordinate. However, it is important to understand that in the relation AB=AC+CB, the three vectors can be arranged in a triangular shape. Therefore, it is important to think about how to place the vectors that cannot be uniquely placed, making it more important for the learner to be able to "operate" them themselves. First, the system presents a problem statement and a figure that

is visualized when the correct answer. The learner observes the presented items and constructs an equation that he or she thinks is correct. The system visualizes the figure based on the equation constructed by the learner. Learners The learning support systems developed by the authors to date have shown a certain level of learning effectiveness.

Although the methods of Kurokawa et al. and our previous study allow learners to manipulate figures, those systems only present the figures that represent the answers (mathematical formulas) constructed by the learners, and learners are not required to actually perform the operations. Therefore, although we were able to visualize and present the learner's answer as a figure to the learner, the reflection of the answer using the operation was learner dependent. By giving the learner feedback on the correctness of their answer only, the student may simply accept the feedback and not fully reflect on their own answers, and thus learning becomes passive. If the learner does not fully recognize the relationship between their own answers and the visualized behaviors and figures, they may not understand why their answers are wrong. It is therefore important to make the learner aware of this relationship by providing specific interactions that highlight how their answers are incorrect. This problem has been discussed in past studies on error visualization. Free descriptions and concept maps were proposed as methods for making learners aware of how their answers relate to the visualized behaviors, and not understanding this will prevent them from fully reflecting on their errors. We hope to overcome this difficulty with the activity proposed in this paper, which we will elaborate upon in the next section.

### 3. Methodology

As stated previously, we are proposing a system that requires learners to explicitly manipulate figures, thereby allowing them to perform semi-active learning activities.

The operations required in our method are designed based on the model that is to be learned in the learning activity. For example, if we want the learner to understand a relationship like "in a vector addition-subtraction formula, a triangle can be drawn using the group of vectors composed by the formula," then the operation must be able to draw such a triangle. When the learner makes an error, the incorrect response is visualized in a strange figure. The learner then superimposes components from their answer onto those for the correct answer. In the vector addition/subtraction example, the learner visualizes each of the terms in the equation for their answers as a vector, then draws the triangles that are possible from that equation. By superimposing these onto the triangle associated with the correct answer, the learner verifies that their answer is incorrect. This then encourages them to proactively reflect on their answers.

Although we focus only on vector addition in this study, we expect that this method will be generally applicable regardless of field. Previous studies have left it up to the learner whether to interact with the figures, but our system requires it as part of the problem solving process, thereby promoting trial and error.

Operation index were presented to the learner as explicit demands for manipulating the figure. These index represent the goals to achieve by manipulating the figure, and it is left to them to experiment with the figure and deduce how to achieve them. This requires the learner to be proactive and encourages them to independently discover why their answer is wrong.

An operation index is presented in the form of a model of a certain activity that represents what is required for solving a given problem. In the addition and subtraction of vectors, which is the subject of the present study, a triangle can be formed from vectors on both sides of a given expression. As an example, suppose that we are given a regular pentagon ABCDE and are told to find XX for AC = XX + XX using point B. If the correct answer to this problem is AC = AB + BC, then a triangle can be formed from the vector AC on the left and the two vectors AB and BC on the right, as shown in Figure 1. The operation index for this problem would be the triangle consisting of these three vectors. By manipulating the visualized vectors in such a way that they fit into this triangle, the learner can learn the rule that the visualized vectors form triangles via the addition and subtraction of the vectors.



Figure 1. The operation index for vector addition.

Furthermore, if the learner gives an incorrect answer, then it will not be possible for them to form a triangle that matches the operation index associated with the correct answer. For instance, assume the learner incorrectly answers AC = AB + CB, then the resulting visualization and operation index will be shown as in Figure 2. It is possible to form this triangle using AC, AB, and CB by manipulating the visualized vectors (upper portion of Figure 2), but it is not possible form the triangle using AC, AB, and BC (lower portion of Figure 2). Trying to reproduce the operation index makes the learner aware of the errors in their answer and helps with knowledge retention.



Figure 2. Operation index in vector addition for an erroneous answer.

# 4. Development Systems

Figure 3 shows a screenshot of what is actually displayed by the system. The target range of this system is vector composition in mathematics. This allows the system to request explicit operation from the learner as a presentation of the operation index, since the model to be presented as the operation index is fixed to one model. It first provides the learner with a problem statement and asks them to find its answer. In previous learning support systems, the figure associated with the correct answer was presented at the same time. In contrast, the present system gives the triangle corresponding to the correct answer as an operation index. Once a learner inputs an answer, the system presents a visualization of their answer along with the operation index for the correct answer.



Figure 3. Screen output from the proposed system.

In the example shown in Figure 3, the learner presented with a regular pentagon ABCDE and asked to determine XX for AC = XX + XX using point B. Then, the triangles formed by the vectors AC, AB, and BC, which are generated when the correct answer in this question is AC = AB + BC, are presented as the operation index for the correct answer. We believe that the learner's search space can be expanded more with this system than with conventional learning support systems, which provide feedback based on the correct figure only.

The learner constructs the correct answer by looking at the question text and the operation index, then using a combo box from the answer input form at the bottom of the screen (Figure 3). The system visualizes errors by converting the learner's answers (mathematical expressions) into vectors (figures) (left portion of Figure 4). At the same time, the system also presents the triangles that can be formed by the vectors in the learner's answer as operation index (right portion of Figure 4).



Figure 4. The learner's operation index (correct answers).

Figures 4 and 5 show the operation index presented when the learner answers the question in Figure 3. Figure 4 shows the operation index for the correct answer, AC = AB + BC, while Figure 5 shows the one for the incorrect answer, AC = AB + CB. These visualizations make the difference between the correct answer and the learner's answer clear.

Figure 6 shows how the system specifies the vectors and allows the learner to select and manipulate them. This necessitates that the learner think more deeply about the individual vectors they are manipulating.



Figure 5. The learner's operation index (error).



Figure 6. Vector selection.

The learner is then asked to use the vectors to reproduce the operation index from their answer and the correct answer, an example of which is shown in Figure 7. We believe that this allows the learner to recognize the relationship between their answer given as a mathematical formula and the graphical representation. The learner will not be able to reproduce the operation index for the correct answer if their answer is incorrect (Figure 8). This process allows the learner to proactively discover why their answer is wrong and to reflect on it. Moreover, since the learner can only move on to the next answer by forming a triangle, we expect this to solve the problem found in conventional learning support systems in which the learner neither checks the visualization of the error nor reflects on it.



Figure 7. Output after vector manipulation (correct answer).



Figure 8. Output after vector manipulation (error).

### 5. Conclusion

In conventional learning such as learning from textbooks or classroom teaching, even if a learner gives an incorrect answer, they are often given only the correct answer and the corresponding explanation, and there is little adaptive feedback for the learner's specific error. This runs the risk of the learner will becoming a passive learner who only implicitly accepts correct answers, and the activity may encourage the learner to simply try to memorize correct answers without fully reflecting on their own wrong answers. Therefore, it is important to create an environment in which learners can reflect and fully consider why their answers were incorrect.

In the previously developed learning support system, errors were visualized by converting learners' answers into graphs. This provided an environment that allowed learners to think consider where they made mistakes and encouraged trial-and-error engagement.

However, in these systems, learners were only presented with figures, and the system did not require that they reflect on their own answers from the presented figures. If a learner only checks the given feedback on whether their answers are correct and does not fully reflect on their own answers, there was a risk of passive learning. In such a case, the learner may not understand why their answer was wrong, and they are likely to make the same mistake in similar problems.

In order for learners to fully reflect on their own answers, they need to recognize the relationship between their answers and the visualized figures. Therefore, it is important to provide specific activities that require learners to recognize the relationship between their answers and the figures.

In this paper, we proposed a learning support system for vector addition that requires learners to manipulate figures explicitly, thus encouraging them to perform semi-active learning activities. By explicitly requesting operation of figures from the system, learners can independently discover why their answers are wrong through the operation of figures and can reflect on their own answers.

In the proposed system, errors are first visualized by converting the learner's input answers into figures. Then, a triangle is formed according to the learner's answer and is presented to the learner as an operation index. This operation index is a model of the rules governing the problem given to the learner, and it is a criterion for how to perform the operation. In the addition of vectors, there is a rule that a triangle can be formed from the subtraction formula on the right-hand side and the vector on the left-hand side, so a triangle is presented as an operation index.

The learner manipulates the vectors generated by the system by dragging them and places them to form the triangle that represents the learner's answer, presented as the operation index, through trial and error. We believe that this allows the learner to recognize the relationship between their answers and the graph.

Next, the triangle formed by the correct answer is presented as an operation index. If the learner's answer is incorrect, when the vectors generated by the learner's answer are

superimposed on the triangle formed by the correct answer, the correct triangle is cannot replicated. This can help the learner to proactively discover why their answer is wrong through the operation of the figure, and to reflect on it.

We believe that this method is a new way to make learners aware of the relationship between their own answers and strange phenomena, which has been discussed in the field of error visualization.

As a task for future work, we will conduct evaluation experiments to verify the learning effectiveness of this system. We will also explore the extent to which the method of presenting operation index can be applied.

### Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers JP22K12322, JP21H03565, and JP20H01730.

### References

- Enomoto, H., Hayashi, Y., Hirashima, T. (2018). Experimental Use of Problem-Posing Exercise System for Efficient Calculation to Promote Relational Interpretation of Numerical Expression: Proceedings of the 26th International Conference on Computers in Education.
- Hayashi, Y., Tsudaka, N., Iwai, K., Hirashima, T. (2021). Learning the Condition of Addition and Subtraction Word Problems by Problem-Posing Based on Representation Conversion Model: *Proceedings of the 29th International Conference on Computers in Education*, 87–92.
- Hirashima, T., Horiguchi, T., Kashihara, A., Toyoda, J. (1998). Error-based simulation for errorvisualization and its management: *International Journal of Artificial Intelli- gence in Educa-tion*, 9(1-2), 17–31.
- Hirashima, T., Imai, I., Horiguchi, T., Tomoto, T. (2009). Error-based simulation to pro- mote awareness of error in elementary mechanics and its evaluation: *In Proceedings of Interna-tional Conference on Artificial Intelligence in Education* 409–416.
- Hirashima, T., Shinohara, T., Yamada, A., Hayashi, Y., Horiguchi, T. (2017). Effects of Error-Based Simulation as a Counterexample for Correcting MIF Misconception: *In Proceedings of the 18th International Conference on Artificial Intelligence in Education*, 90–101.
- Horiguchi, T., Hirashima, T. (2002). Simulation-Based Learning Environment for Assisting Error-Correction Management of Error-Based Simulation Considering the Cause of Errors: *The Japanese Society for Artificial Intelligence (Japanese edition), 17*(4), 462-472(2002), (in Japanese).
- Horiguchi, T., Imai, I., Tomoto, T., Hirashima, T. (2014). Error-based Simulation for error-awareness in learning mechanics: *An evaluation: Educational Technology and Society 17*(3), 1–13
- Jumonji, T., Aikawa, N., Tomoto, T. (2022). Development of a Learning Support System with Error Visualization for Learning for Vector in Mathematics: *The Japanese Society for Artifi-cial Intelligence (Japanese edition)*. 1-6, (in Japanese).
- Junus, K. (2018). Assessing students' mathematical misconceptions through concept maps and online discussion transcripts: Inner product spaces: *Proceedings of the 26th International Conference on Computers in Education.* 772–777.
- Kurokawa, K., Tomoto, T., Horiguchi, T., Hirashima, T. (2018a). Development and Evaluation of a Learning Support System with Functions for Conversion of Expression and Facilitation of Active Er-ror-Awareness in Locus Problem in Mathematics: *The Institute of Electronics (Japanese edition) J101-D*(6) 864-873, (in Japanese).
- Kurokawa, K., Tomoto, T., Horiguchi, T., Hirashima, T. (2018b). Development of a Mathematical Solution Environment to Understand Symbolic Expressions in Mathematics: *In Proceedings of Hu-man Interface and the Management of Information: HIMI2018,* 288-299.
- Mehanovic. S., Spikol, D. (2012). Investigating how to design interactive learning environments to support students' learning of upper secondary and university math: *Proceedings of the 20th International Conference on Computers in Education*, 782–786.
- Nakahara, T. (1995). Research on Constructive Approaches in Arithmetic and Mathematics: *Seibunsha*.

- Nakamoto, R., Flanagan, B., Takami, K., Dai, Y., Ogata, H. (2021). Identifying students' stuck points using self-explanations and pen stroke data in a mathematics quiz: *Proceedings of the 29th International Conference on Computers in Education*. 522–531.
- Sengupta, P., Dickes, A., Farris, A. (2018). Toward a Phenomenology of Computational Thinking in STEM Education: *Computational thinking in the STEM disciplines,* 49–72.
- Slater, S., Ocumpaugh, J., Baker, R., Li, J., Labrum, M. (2018). Identifying Changes in Math Identity Through Adaptive Learning Systems Use: *Proceedings of the 26th International Conference on Computers in Education.*
- Wong, S, L., Wong, S. (2017). Motivation towards Mathematics Learning in the Technology-enhanced Environment: Workshop Proceedings of the 25th International Conference on Computers in Education. 109–115.
- Ya-Jing, YU., Po-Han, WU., Yu-Sheng, SU. (2017). Effects of concept map based cooperative peer assessment system on students' learning outcomes on programming: *Workshop Proceedings of the 25th International Conference on Computers in Education.* 219–224.
- Yang, E. F. Y., Cheng, H. N. H., Ching, E., Chan, T. W. (2012). Variation based discovery learning design in 1 to 1 mathematics classroom: *Proceedings of the 20th International Conference* onComputers in Education, 811–815.
- Yang, E. F. Y., Cheng, H. N. H., Chan, T. W. (2014). Math Creation: Integrate Peer Tutoring for Facilitating the Mathematical Expression and Explanation: *Proceedings of the 22nd International Conference on Computers in Education*, 161–170.

# Using a Three-Problem Framework to Understand How Nursing College Students Learn to Design Healthcare Animations

**Chun-Hao CHANG** 

Allied Health Education and Digital Learning, National Taipei University of Nursing and Health Sciences, Taiwan chunhao@gm.ntunhs.edu.tw

**Abstract:** This study implemented a three-problem instructional framework to investigate the learning performance of nursing college students in healthcare animation design. Each genre of problem represents a distinct narrative animation design skill to acquire, including abstract design, actual design, and deconstruction design skills. A total of 34 nursing college students from northern Taiwan were recruited as participants. All participants sequentially engaged in the learning tasks presented by the three problems, with each problem unit spanning five weeks. Assessment took place at the end of each problem unit, utilizing a self-developed 5-point scoring rubric to analyze students' performance and variances across the three animation design problems. Preliminary findings revealed that students exhibited comparatively lower learning performance in abstract design and deconstruction design problems while excelling in actual design problems. Additionally, further analysis suggested no significant correlations between students' performance in abstract design and deconstruction design and deconstruction design, shedding light on future pedagogical design approaches. **Keywords:** Narrative animation, problem-solving, healthcare, nursing college students

## 1. Introduction

In the ever-evolving landscape of healthcare education, the acquisition of proficiency in crafting narrative healthcare animations holds profound significance for nursing college students (Parikh & Huniewicz, 2015; Korda & Itani, 2013). This significance has been magnified, particularly in the wake of the Covid-19 pandemic, which has reshaped the dynamics of public engagement with health information. The role of narrative healthcare animations as a pedagogical tool has transcended conventional methods of disseminating health knowledge. As the pandemic surged through communities, a noticeable shift occurred in how the public seeks, absorbs, and comprehends health-related information. The surge in digital consumption during lockdowns and social distancing measures prompted a surge in the popularity and effectiveness of narrative animations as a medium of communication.

Against this backdrop, nursing college students in Taiwan find themselves at the intersection of healthcare education and digital curation. The imperative to master the art of designing narrative healthcare animations has arisen not only from the need to convey complex, abstract healthcare concepts to a diverse audience but also from the recognition that animations possess a unique ability to bridge gaps in health literacy (Bétrancourt & Tversky, 2000; Preim & Meuschke, 2020; Weiss, Knowlton & Morrison, 2002; Sanchez & Wiley, 2010). The transformative potential of these animations lies not only in their ability to convey information but also in their capacity to evoke empathy and emotional connection (Lowe & Boucheix, 2008; Lowe & Boucheix, 2016). Storytelling through animation enables the presentation of healthcare scenarios in relatable and compelling ways, fostering a deeper understanding of health issues among the public. Consequently, nursing college students are not merely acquiring a technical skill; they are honing a form of expression that amalgamates creativity, medical acumen, and effective communication.

Thus, it becomes evident that a significant portion of nursing students exhibits the capacity to proficiently navigate the creation of healthcare animation projects, meeting the initial completion benchmark with relative ease. Nonetheless, it is noteworthy that despite this widespread capability, the resultant animation projects diverge markedly in terms of their quality and overall efficacy. While a substantial number of students can traverse the technical aspects of animation design, the variations in the final outcomes underscore a crucial distinction between merely producing animations and crafting ones that genuinely resonate with their intended audiences (Calderón, Shaheen, Hays, Fleming, Norris & Baker , 2014; Betrancourt, 2005). Ideally, a proficiently designed healthcare animation should adeptly transform abstract healthcare terminology, like "Myocardial Infarction," into a narrative scenario replete with relatable examples that enable the audience to swiftly grasp and apply the information to real-world scenarios. Thus, the audiovisual features of animations alone without a meticulous, structured design as well as thinking process fail to deliver the learning content.

## 2. Narrative Animation and Learning

Narrative animation can be perceived as a cross-national and cross-lingual teaching and communication tool. Relevant researches indicate that when integrating animation into learning, it primarily enhances learners' motivation through high interactivity and rich context-dependent storytelling. This approach reduces the extraneous cognitive load of knowledge content and reinforces learners' perception of the knowledge content, particularly aiding in the construction and organization of specific knowledge mental representations (Cavazza, Bandi & Palmer, 1999; Fleer, 2018; Mayer and Moreno, 1998). Among these, the convenient and narrative-rich nature of animation is common in health education animations, where creating learning situations through storytelling is a prevalent design approach. The narrative context of animation aids in knowledge transmission because storytelling itself assists learners in revisiting life experiences and organizing sequences and logical relationships among various experiences (Gottschall, 2013; Moen, 2006). Therefore, for abstract health education knowledge requiring contextual comprehension through storytelling, computer animations with narrative elements play an inseparable role in instructional applications and research.

Given the swift expansion of cloud-based, drag-and-drop animation tools like Vyond, Animaker, and Powtoon, the process of designing healthcare narrative animations appear less daunting for nursing college students than in the past. The technical learning curve of animation design has steepened for learners, yet the real challenges reside in the instructional design strategy that can effectively foster the assimilation of healthcare knowledge. The process of narrative animation design, however, is typically a linear flow in which learners engage from start to finish on their own. Subsequent refinement and adjustments to the animation work often occur beyond the confines of a classroom, posing a challenge for teachers to effectively monitor and provide timely instructional support. Consequently, within the context of animation design courses tailored for nursing college students, teachers encounter a number of unresolved challenges. Foremost, there lies the imperative to cultivate the adeptness in converting abstract, complex healthcare concepts into concrete examples—an endeavor demanding immersive practices, discussions, and introspective evaluation. This crucial facet often remains overshadowed within conventional animation design curriculum. Secondly, varied narrative plots begets a corresponding diversity in the demanded audiovisual effects for animations. Given the expansive spectrum of healthcare themes, guiding students towards identifying pertinent shortcomings within their narrative animation endeavors presents an intricate task for teachers. Thirdly, a majority of students possess the ability to create animations from the ground up, but they often lack the expertise to analyze and dissect an animation project that is already in progress. This skill becomes particularly crucial when designing healthcare animation projects that encompass multiple viewing paths.

# 3. Research Method

### 3.1 Participants

The study took place at a nursing vocational university located in Taipei, Taiwan. The participants comprised 34 sophomores, consisting of 27 females and 7 males, with an average age of 21.6 years. None of the participants had previously enrolled in any courses related to animation design prior to the study.

## 3.2 Procedure

In this 15-week study, each participant was required to attend a 100-minute animation design course in a computer classroom every week. The researcher and another nursing expert teacher jointly served as the instructors. The course was structured into three problem-based learning units, including (1) abstract design problems, (2) actual design problems, and (3) deconstruction design problems. Each problem unit had a duration of 5 weeks. All participants engaged in the three, problem-based learning units in a sequential order. Assessments were conducted after the final class of each learning unit, exploring participants' learning performance toward animation design.

## 3.3 Curriculum

The curriculum was designed and structured with respect to the three animation design problems as stated above. In the abstract design problem unit, participants were asked to transform textual health information into narrative animation scripts, endeavoring to present abstract healthcare knowledge concepts in a personified manner for easier comprehension. Figure 1 is an example of transforming text-based descriptions of Angina pectoris into visual illustrations. Participants used the Canva application to visualize and illustrate their ideas as individual scenes. Next, In the actual design problem unit, all participants attempted to utilize the Vyond animation design application to carry out the previously designed story script. Participants were tasked with contemplating how to adhere to multimedia design principles and effectively employ the audio-visual effects of animation to convey the narrative within the script. Through the use of audiovisual enhancements, participants aimed to guide the audience's attention effectively. Lastly, in the deconstruction design problem unit, participants exchanged their unfinished animation works with other groups and then attempted to complete each other's designs. Participants needed to first deconstruct the narratives step by step and then figure out how to resume the design work without deviating from the original goals.



Figure 1. An example of the abstract design problem

### 3.4 Assessment

The assessment primarily corresponds to the three design problems that guide the learning of animation design. The researcher created four measures for each of the three design problems along with one additional measure for assessing the overall design performance. These measures include Abstraction Measure, Design Measure, Deconstruction Measure, and Application Measure. The first three measures correspond to the three animation design problems while the last measure evaluates the final healthcare design project at the end of the class. Each measure consists of customized scoring rubrics. For each assessment indicator, the researcher designed customized scoring rubrics, serving as the basis for measuring students' learning outcomes before and after participating in the workshops.

A scoring rubric adheres to three application norms: (1) it focuses on measuring a clearly describable learning behavior, (2) it employs a scale to measure the outcomes of learning activities, (3) it employs text-based or descriptive representations of specific learning performances, which are then graded on a scale. In the current study, each dimension and detailed scoring criterion within the scoring rubric were reviewed by three domain experts to ensure reliability.

## 3.5 Narrative Animation Design Tools

Vyond is used as the animation design tool for all participants. Vyond is a cloud-based narrative animation design application, its interface is shown in Figure 2. Learners don't require to learn any modeling techniques to start their design work; instead, they can quickly create characters according to the animation script using an intuitive drag-and-drop approach and modular editing interface. Additionally, learners can even customize the facial expressions and subtle movements of their characters.



Figure 2. The user interface of Vyond animation design application

## 3.6 Research Questions

The proposed research questions are as follows:

- 1. What are the learning results when nursing college students engage in the three problembased animation design tasks?
- 2. Is there any interaction effect between the three animation design problems for nursing college students?
- 3. Does one's performances in the three problem-based design tasks affect their overall animation design results?

## 4. Preliminary Results

Multiple regression analysis indicated that participants' learning performance in the three design tasks explained 63.9% of their overall animation design results at the end. Participants showed the highest scores in the actual design problems, and they struggled with the abstract design as well as deconstruction design problems. When exploring the interactions among the four measures, the Repeated Measures ANOVA results indicated a significant interaction among the four assessment indicators, with Wilks' Lambda=0.77, F(2,

28)=4.08, p=0.028. This suggested that the three types of animation design problem implementations significantly impact the overall learning outcomes of nursing college students.

Further analyzing the correlations among the four measures, the Pearson correlation coefficient analysis revealed that the overall "Application Measure" had significant correlations with the other three indicators (p<0.01). The "Design Measure" showed significant correlations with both "Abstract Measure" and "Deconstruction Measure" (p<0.01), while the correlation between "Abstraction Measure" and "Deconstruction Measure" was not significant (p=0.247), see table 1. These analyses implied that nursing students' performance in abstract design problems and in deconstruction design problems might represent distinct problem-solving capabilities.

Measure	1	2	3	4
Abstraction Measure				
Design Measure	0.412*			
Deconstruction Measure	0.247	0.385**		
Application Measure	0.678**	0.639**	0.491**	

Table 1. Correlational analysis of the four measures

## 5. Conclusion

Integrating and implementing three distinct types of animation design problems into one class was a challenging experimental teaching attempt aimed at guiding nursing college students to develop healthcare narrative animation design skills. The results yielded fruitful pedagogical insights. Drawing from the perspective of Lave & Wenger (1990), learning is seen as a contextually meaningful social phenomenon, leading to the concept of Legitimate Peripheral Participation (LPP). This concept asserts that novice learners must gradually cultivate expert experiences through engagement in essential learning activities. The three-problem framework in this study essentially fostered LPP for one's animation design competence. Through problem-solving activities in diverse contexts, nursing college students were guided to actively engage in the entire process of narrative animation creation, from pre-production to design and refinement, while also assessing their own creative and explanatory skills. Lastly, on the technical front of animation creation, current healthcare narrative animations predominantly follow a linear, single-path viewing format where viewers watch from start to finish. This fixed viewing path limits viewer interaction with the knowledge content within the animation. The absence of an open script and interactive viewing paths hinders the customization of animation scripts according to individual viewer needs. Students' creative works also often lack an awareness of the limitations posed by a single viewing path. They may not know how to incorporate interactive paths into their creations. The challenge lies in guiding students to design customized viewing paths, allowing for interaction and engagement with the content. Integrating interactivity into animations requires students to be conscious of the possibilities and limitations posed by different viewing paths. In the future, facilitating the incorporation of customized viewing paths into students' creations will be a significant challenge.

## Acknowledgments

This material is based upon work supported by the National Science and Technology Council of Taiwan under Grant 111-2410-H-227-006.

### References

- Betrancourt, M. (2005). The Animation and Interactivity Principles in Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (p. 287–296). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.019
- Bétrancourt, M., & Tversky, B. (2000). Effect of computer animation on users' performance: A review. Le Travail Humain: A Bilingual and Multi-Disciplinary. *Journal in Human Factors, 63*(4), 311–329.
- Cavazza, M., Bandi, S. & Palmer, I. (1999). Situated AI in Video Games: Integrating NLP, Path Planning and 3D Animation. In *Proceedings of the AAAI Spring Symposium on Computer Games and Artificial Intelligence*.
- Calderon, J. L., Shaheen, M., Hays, R. D., Fleming, E. J., Norris, K. C., & Baker, R. S. (2014). Improving diabetes health literacy by animation. *Diabetes Education*, *40*(3), 361-372.
- Fleer, M. (2018). Digital animation: New conditions for children's development in play-based setting. *British. Journal of Educational Technology, 49*(5), 943–958.
- Gottschall, J. (2013). *The storytelling animal: How stories make us human*. New York, NY: First Mariner Books.
- Korda, H.& Itani, Z. (2013). Harnessing social media for health promotion and behavior change. *Health Promotion Practice, 14*(1), 15-23. https://doi:10.1177/1524839911405850
- Lave, J., & Wenger, E. (1990). *Situated Learning: Legitimate Peripheral Participation*. Cambridge, UK: Cambridge University Press.
- Lowe, R. K., & Boucheix, J.-M. (2008). Learning from animated diagrams: How are mental models built? In G. Stapleton, J. Howse, & J. Lee (Eds.), *Diagrammatic representation and inference* (pp. 266-281). Berlin: Springer.
- Lowe, R. K. & Boucheix, J. M. (2016). Principled animation design improves comprehension of complex dynamics, *Learning and Instruction*, *45*, 72-84.
- Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90(2), 312–320. <u>https://doi.org/10.1037/0022-0663.90.2.312</u>
- Moen, T. (2006). Reflections on the narrative research approach. *International Journal of Qualitative Methodology*, *5*(4), 56-69. https://doi.org/10.1177/160940690600500405
- Parikh, S. V., & Huniewicz, P. (2015). E-health: an overview of the uses of the Internet, social media, apps, and websites for mood disorders. *Current opinion in psychiatry*, *28*(1), 13-17.
- Preim, B. & Meuschk, M. (2020). A survey of medical animations, Computers & Graphics, 90, 145-168.
- Sanchez, C. A., & Wiley, J. (2010). Sex differences in science learning: Closing the gap through animations. *Learning and Individual Differences, 20*, 271–275. https://doi.org/10.1016/j.lindif.2010.01.003
- Weiss, R. E., Knowlton, D. S., & Morrison, G. R. (2002). Principles for using animation in computerbased instruction: Theoretical heuristics for effective design. *Computers in Human Behavior*, 18(4), 465–477. <u>https://doi.org/10.1016/S0747-5632(01)00049-8</u>

# Focal: A Proposed Method of Leveraging LLMs for Automating Assessments

#### Peter Meyers<sup>a</sup>, Annette Han<sup>a</sup>, Razik Grewal<sup>a</sup>, Mitali Potnis<sup>a</sup> & John Stamper<sup>a</sup> <sup>a</sup>Carnegie Mellon University. United States

\*jstamper@cs.cmu.edu

**Abstract:** In response to the growing need for frequent, high-quality assessments in the expanding field of online learning and the significant time burden their manual creation places on educators, this study proposes Focal, an end-to-end assessment generation pipeline. Focal employs large language models, notably Text-to-Text Transfer Transformers, fine-trained on diverse learning materials, to generate and evaluate pedagogically sound questions and their corresponding answers. The Focal pipeline is designed to integrate with Learning Management Systems, providing educators an automated means of creating assessments that align with their curriculum. This not only eases the task of creating and evaluating assessments but also frees educators to focus on other crucial responsibilities. The system is domain agnostic and its efficacy is continually improved by training and evaluating it using data from multiple subject areas. By automating the traditionally labor-intensive process of assessment production, Focal aims to increase efficiency in online education and enhance the learning experience for students.

Keywords: Question Generation, Assessment Generation, Large Language Models

### 1. Introduction

Frequent assessments bolster long-term retention and conceptual understanding, acting as "memory modifiers" (Bjork, 1975) aiding learners to apply information in novel situations (Storm et al., 2010, Kornell et al., 2009). This process enhances recall, cultivates adaptive expertise (Baroody, 2013, Chi et al., 1982), and promotes cognitive flexibility in procedural domains like mathematics (Hiebert and Lefevre, 1986, Salomon and Perkins, 1989).

Given assessments' critical role, educators invest considerable effort in crafting questions reinforcing course concepts (Zhao et al., 2005). However, curating a substantial question bank that caters to students of varied skill levels poses a considerable challenge—it is a time-consuming and labor-intensive process (Nguyen et al., 2022). Moreover, providing quality feedback, crucial to augment students' learning experiences, further compounds this burden on educators.

Addressing these challenges necessitates a novel approach, leading us to propose Focal—an innovative, end-to-end assessment generation pipeline. Designed to automate the process of creating educational assessments, Focal is built upon the principles of the Knowledge-Learning-Instruction (KLI) framework (Koedinger et al., 2012). This framework guides a systematic approach to educational practice, prompting detailed analyses of the knowledge students acquire in courses. Consequently, Focal auto-generates questions and answers based on provided curricula across various subject domains, thus embodying instructional principles with a high potential for generality and skill tracking.

The core of Focal lies in its machine learning pipeline that ingests texts from learning materials, generating pedagogically sound and logically coherent questions and answers that align with the curriculum. It further evaluates their quality to ensure the assessments are academically rigorous. Leveraging large language models (LLMs) and data from multiple subjects, Focal continually enhances its domain-agnostic capabilities, which helps maintain the relevance and quality of the assessments it generates.

Focal's integration into Learning Management Systems (LMS) not only alleviates the burden on educators but also democratizes access to high-quality assessments. These assessments, drawn from expert-curated curricula, are accessible to all learners, regardless of their socio-economic background. By effectively addressing scalability and quality issues inherent in assessments, Focal positions itself as a promising solution in the dynamic landscape of online education, propelling it towards new horizons of inclusivity and efficacy.

## 2. Related Work

The proliferation of LLMs significantly shapes the landscape of Natural Language Processing (NLP) and Question Generation (QG). These transformer-based encoder-decoder models have demonstrated efficiency in general QG (Xue et al., 2020, Yu et al., 2021) and in educational settings as well (Grover et al., 2017) significantly propelled by neural transformer-based methods and the BERT model (Vaswani et al., 2021).

In addressing the challenges initially faced by BERT in token generation (Lopez et al., 2020) advanced the robustness of QG by fine-tuning a pre-trained language model. Efforts have also been made to refine the relevance of the generated questions by incorporating common sense and domain knowledge into the QG process (Jia et al., 2021, Wang et al., 2020) and developing an attention-based sequence-to-sequence model that integrates target answer information (Liu, 2020).

Significant progress has been made towards automating the generation of educational questions, with models such as GPT-2 successfully generating mathematical word problems of varying difficulty (Cheng et al., 2021) and others producing questions reflective of realistic scenarios (Liu et al., 2022). Recently, researchers employed GPT-3 to create EduQuiz, an end-to-end educational quiz generator (Dijkstra et al., 2022). Despite certain limitations, including language specificity, domain specificity, and higher costs for fine-tuned models, EduQuiz generated reasonably high-quality quizzes. High-quality distractor generation, however, remains a challenge. The authors envisage potential for improvement and suggest integrating human input in future iterations to enhance the quality of generated quizzes.

QG's evaluation remains a formidable challenge, necessitating a blend of automated assessments using machine learning models and human evaluations (Kurdi et al., 2020). Automated assessment strategies often rely on metrics such as BLEU and ROGUE (Novikova et al., 2017). However, these approaches face scrutiny due to concerns surrounding their interpretability and weak correlation with human evaluations (Van Der Lee et al., 2019). In contrast, human evaluators examine factors like grammatical correctness, fluency of language, relevance to the topic, and the naturalness of the language employed in the questions (Amidei et al., 2018, Chen et al., 2018, Ruseti et al., 2018)

An additional measure of evaluation involves ensuring that the generated questions correspond with "ground truth" data or expert-crafted reference questions (Sai et al., 2022). This aspect becomes even more crucial in the realm of educational QG, where the assessments are expected to be a mirror reflection of the learning material's target skills. The creators of MOOCCubeX introduced an automated solution to address this concern, pioneering the path for adaptive learning research and concept-centric data organization (Yu et al., 2021).

Against this backdrop, the Focal pipeline emerged as a testament to the success of LLMs, further propelled by findings that key concept extraction could boost their usefulness (Yu et al., 2021, Stamper et al., 2023). Recent work (Nguyen et al., 2022) has established the potential of LLMs in generating questions that are not only coherent but also pedagogically beneficial. Our work builds upon this foundation, aspiring to optimize the Focal pipeline end-to-end, expand Focal's domain-agnostic capabilities, and refine the grading metrics for questions.

## 3. **Methodology**

## 3.1 Data

In our initial Focal pipeline testing, we used text data from a graduate-level data science course - Foundations of Computational Data Science (FCDS) and an undergraduate-level chemistry course, existing in XML format. The course content has hierarchical levels: Units, Modules, and Topics. We prepended these hierarchy titles to the course text based on the proven value of such an approach in QG (Nguyen et al., 2022). This is done for each hierarchy level, as illustrated in 1. For instance, the unit title gets concatenated with the text content before being fed into the QG model.



*Figure 1.* Graphical depiction of data pre-pending process demonstrated to be valuable in generating high quality questions.

Aside from course content, Focal employs the SQuAD 1.1 dataset for pre-training the QG model; this dataset comprises over 100,000 comprehension questions based on Wikipedia articles (Rajpurkar et al., 2016). This method improves the logic of generated questions (Nguyen et al., 2022). Further, the LearningQ dataset, with 230,000 document-question pairs and 7000 educator-crafted queries on assorted educational subjects (Chen et al., 2021), refines our evaluator model, enhancing the pipeline's accuracy in judging the logic of auto-generated questions.

# 3.2 Model Design

Figure 2 delineates Focal's QG Workflow. Initiated with Data Extraction and Pre-processing, the pipeline processes XML input data using BeautifulSoup to scrape and cleanse, including punctuation and stop word removal. It prepares the data for subsequent stages, including a crucial pre-pending phase that extracts headings, keywords, and main text from each section of course material. Post-preprocessing, the data splits, directing to the Concept Hierarchy Extraction and Question and Answer Generation stages.

The subsequent Concept Hierarchy Extraction, a vital part of the Focal pipeline, encompasses key concept creation and extraction. This stage relies on prior research showing LLMs' improved question generation ability with these key concepts (Stamper et al., 2023). We leverage the MOOCCubeX pipeline for this extraction, a platform rich in educational content and associated concept maps (Yu et al., 2021). It also filters invalid concepts like prepositions, indexing numbers, and generic verbs, currently needing a domain expert. As part of our long-term goals, we aim to develop a self-sufficient system capable of conducting this filtering process independently.

After Concept Hierarchy Extraction, the pipeline enters the Question and Answer Generation phase, creating multiple-choice and short-answer questions. This phase employs

a T5 transformer-based encoder-decoder model, fine-tuned based on prepending headings to the text (Raffel et al., 2020). The T5 model's current fine-tuning relies on the SQuAD 1.1 dataset (Rajpurkar et al., 2016), with potential for future adjustment. Crucially, it generates distractors for multiple-choice questions and extracts the correct answers via a dependency parse tree in a rule-based approach.



Figure 2. Complete Focal QG workflow diagram.

The final step before involving students is the Evaluation of Generated Questions and Answers. Here, three primary methods are used to evaluate the generated questions in terms of logical and pedagogical soundness: information score, GPT-3 model classification, and human expert evaluation. The information score is a custom metric designed to evaluate each question's relevance within the context of the identified key concepts from the Concept Hierarchy Extraction phase. By analyzing the overlap between tokens in a question and the extracted key concept tokens from the course text, the information score provides a robust assessment of question quality, while normalizing for question length ensures fairness across all questions.

$$IS = \frac{1}{|T(q)|} \sum_{t \in T(q)} 1(t \in C).$$
(1)

To assess question soundness, we utilize an information score and a GPT-3 model fine-tuned on the LearningQ dataset (Chen et al., 2018). Domain experts also evaluate question soundness, but with the information score's refinement, their involvement might become unnecessary.

Once questions are generated and evaluated, the Focal pipeline integrates student responses. Student answers are assessed for correctness using cosine similarity or exact matching, depending on the question type. For short-answer questions, a vector is generated from the input text using bag-of-words or TF-IDF methods, and its cosine similarity to the model-generated answer vector is measured using Equation 2. In this

equation, A denotes the original answer, S represents the student's response, and D is the dot product of the two.

$$CosineSimilarity(A, S) = \frac{D}{|A| * |S|}$$
(2)

For numerical, Yes/No, or multiple-choice type questions, we employ an exact matching approach between the student's response and the model-generated answer to determine the correctness of the student's answer as illustrated in figure 3.



Figure 3. User Answer Evaluation Workflow.

## 4. Initial Results

## 4.1 Question Generation

In order to evaluate the quality of several QG models, we generated questions for both the data science course and chemistry course using four different LLMs: bert2BERT, GPT-2, BART, and T5 (Chen et al., 2021, Radford et al., 2018, Xue et al., 2020). Table 1 provides examples of questions generated by Focal; additional examples can be found on Github<sup>1</sup>.

Table 1. Examples of questions generated for a paragraph of text content.

Unit	Module	Generated Question
Analytic Algorithms and Model Building	Data Science Patterns	What is the difference between an observed value and the fitted value given by a model?
Data Gathering and Wrangling	Data Wrangling Pipeline	What does omission involve excluding the missing values from the dataset?
Exploratory Data Analysis Performing Exploratory Data Analysis		What measures do you use to describe variability?

After generating questions using each of these models, we performed Model Classification, Information Score, and Perplexity Score to rate the quality of each generated question.

# 4.2 Why use Perplexity?

Perplexity is a metric commonly used in NLP to evaluate the quality of language models. It measures how well a language model predicts a sequence of words or tokens. In simpler terms, perplexity quantifies how surprised a language model is when it encounters a new

<sup>&</sup>lt;sup>1</sup> https://github.com/annettehan/focal

sequence of words. Perplexity is calculated based on the probability distribution assigned by the language model to a given sequence of words. The lower the perplexity score, the better the language model's ability to predict the next word in a sequence.

$$-1/N \sum_{i=1}^{n} \log P(x_i | w_1, \dots, w_{i-1})$$
(3)  
$$P(W) = b$$

Perplexity is particularly useful in assessing the quality of generated questions because it captures how well a language model understands the context and generates coherent and meaningful questions. A language model with a low perplexity score is more likely to generate questions that align with the desired context and exhibit grammatical correctness and relevance. However, it's important to note that perplexity alone may not capture all aspects of question quality, such as the relevance or informativeness of the generated questions. Therefore, it is often used in conjunction with other evaluation metrics, such as Model Classification and Information Score, to provide a comprehensive assessment of the generated questions' quality.

#### 4.3 Model Classification

For the Model Classification, we employed the GPT-3 classifier fine-tuned on the learningQ dataset to label each question as either sound or not sound (Radford et al, 2019, Chen et al., 2018). Here by soundness, we mean whether the questions are rationally valid, contextually relevant, and are effective such that they can be used for assessing the knowledge in regards with the topic. Following the GPT-3 classification, we randomly sampled 100 questions from each model type for both courses and manually annotated them as either sound or not sound. Figures 5 through 8 show the percentage of the 100 randomly sampled questions that each method of evaluation found to be sound for each model.









Figure 6: Human evaluator - Data Science course





Figure 8: Human evaluator - Chemistry course

Initial analysis reveals our fine-tuned GPT-3 classifier often overestimates soundness compared to human assessment, with bert2BERT producing notably more unsound questions. T5 and BART fared better in the subset evaluated. To understand this divergence, we constructed confusion matrices for each model and course.

Table 2 reveals that human evaluators are likely to agree with our GPT-3 classifier when it rates a questions as unsound, suggesting its potential as a tool for eliminating low-quality questions rather than selecting high-quality ones.

	Data Science Course		Chemistry Course	
	Expert: Sound	Expert: Not Sound	Expert: Sound	Expert: Not Sound
GPT3: Sound	161	148	158	183
GPT3: Not Sound	11	80	2	57

Table 2. Confusion Matrix for generated questions (400 random questions for each course).

Upon analysis of the Figure 9 plots, we first see that on average, the perplexity score points to the correct direction of ground truth as the average perplexity of sound questions (149.4) is less than the average perplexity of not sound questions (155.6). Additionally, it seems evident that GPT should outperform T5, as a lower perplexity score signifies that it possesses better prediction power, but when we empirically look at the quality of questions output by T5, they are much more coherent and sound as compared to those output by GPT-2 model, which signifies that perplexity score evaluation has some room for improvement, so it more closely aligns human labeling, which is considered the gold standard in question.

Total Average Perplexity Sound: 149.4 Not Sound: 155.6



*Figure 9.* Average Perplexity vs Human evaluation for Data Science and Chemistry courses. Total Average Information Score: Sound: 1.4 Not Sound: 1.2



*Figure 10.* Average Information Score vs Human evaluation for Data Science and Chemistry courses.

The information score is meant to indicate that a question overlaps with the key concepts of a given passage. As shown by the study of the Figure 10 plots, where the average information score of sound questions is 1.4 and the average information score of unsound questions is 1.2. Furthermore, it seems obvious that GPT should outperform T5, as a higher information score denotes that it has better question quality. However, when we

empirically examine the quality of questions produced by T5, they are significantly more coherent and sound than those produced by GPT-2 model, indicating that information score evaluation has some room for improvement so that it more closely aligns with human evaluation results.

# 5. Error Analysis

Throughout the development and testing of the Focal system, error analysis has been performed to identify issues and improve the quality of the generated questions and their evaluations. Some common error patterns observed in the experimental results are:

- Irrelevant questions: Sometimes, questions generated fail to reflect key concepts, possibly due to misalignment in the Concept Hierarchy Generation and Information Score. To counteract this, Information Score measures question relevance, guiding improvements in Concept Hierarchy Generation and keyword extraction. For instance, it could address instances of vague queries such as "What type of procedure does the procedure follow?" (Table 3).
- Question variability: The Focal system primarily generates "what" type questions and struggles with creating "why" or "how" questions. This might arise from the training data's lack of diversity. To enhance variability, the model can be fine-tuned using additional diverse datasets outside of SQuAD 1.1 or additional data with various question types. This would balance overproduction of "why" questions (Table 3).
- 3. Incorrect evaluation of student answers: The system may incorrectly assess student answers due to limitations in the cosine similarity and exact matching methods. Testing multiple evaluation models and selecting the most accurate would reduce such errors.

Table 3. Examples of question invariability.

Unit	Module	Generated Question
Exploratory Data Analysis	Feature Engineering	What kind of questions do you have?
Analytic Algorithms and Model Building	Model Selection	What is the model Mj with the best performance on the test set?
		What type of procedure does the procedure follow?
		What is a model model for Mi?
		What subset is split into train subset, validation subset and test set?
		What is the name of the best hyperparameters?

In each iteration of the Focal system, error analysis is conducted to identify these issues and guide the improvement of the system. By examining experimental results at various stages and observing patterns of errors, the system can be refined to generate more pedagogically sound questions and provide more accurate evaluations of student answers. This iterative process ensures continuous improvement of the Focal system, enhancing its educational value and potential for adoption in the classroom.

# 6. Conclusions and Future Work

In this study, we have outlined both the current progression and future vision for the Focal assessment pipeline. Our initial evaluations have shown promising potential in its role as an

assessment tool, though certain areas for improvement have been identified. For instance, we found that the Focal QG model occasionally generates questions that may seem logical but are not pedagogically sound or useful for assessments. Additionally, the current information score metric, which grades question quality, needs to be more nuanced to better encapsulate key concepts from course texts.

In our early testing, we noticed that Focal often generated questions that lacked complexity and only required students to recall specific facts from the material. Our future research will explore new QG models to improve this aspect, emphasizing questions that promote a deeper understanding of the course instead of mere recollection of details.

Moving forward, we intend to make the Focal pipeline more valuable by streamlining the assessment process and reducing the burden on educators. This includes automating the end to-end assessment process, from generating and evaluating questions (Moore et al., 2023) to generating answers and evaluating student responses. One potential challenge is the generation of a wide array of question types. For example, LLMs are adept at creating trivia-style questions but struggle with generating more analytical "how" or "why" questions, which are better for gauging a student's comprehensive understanding. We hypothesize that by fine-tuning these models on more diverse datasets with additional analytical questions, the models will produce more varied question types. This will be a key area of focus in our research.

Furthermore, refining the information score metric to more precisely predict the pedagogical appropriateness of questions is a crucial part of future work. This would allow the pipeline to be less dependent on domain experts and ensure that the generated assessments maintain a high level of quality. It's essential that Focal serves as a practical tool for educators and students, delivering timely and relevant feedback, and contributing to a more efficient and effective learning environment.

#### References

- Amidei, J., Piwek, P., & Willis, A. (2018). Evaluation methodologies in automatic question generation 2013-2018.
- Baroody, A. J. (2013). The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge. In *The development of arithmetic concepts and skills* (pp. 1-33). Routledge.
- Bjork, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. Information Processing and Cognition: The Loyola Symposium.
- Chen, C., Yin, Y., Shang, L., Jiang, X., Qin, Y., Wang, F., ... & Liu, Q. (2021). bert2bert: Towards reusable pretrained language models. *arXiv preprint arXiv:2110.07143*.
- Chen, G., Yang, J., Hauff, C., & Houben, G. J. (2018). LearningQ: a large-scale dataset for educational question generation. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 12, No. 1).
- Cheng, Y., Li, S., Liu, B., Zhao, R., Li, S., Lin, C., & Zheng, Y. (2021). Guiding the growth: Difficulty-controllable question generation through step-by-step rewriting. *arXiv preprint arXiv:2105.11698*.
- Chi, M., Glaser, R., Rees, E., & Steinberg, R. J. (1982). , Expertise in problem solving . Advances in the psychology of human intelligence.
- Dijkstra, R., Genç, Z., Kayal, S., & Kamps, J. (2022). Reading Comprehension Quiz Generation using Generative Pre-trained Transformers.
- Grover, K., Kaur, K., Tiwari, K., Rupali, & Kumar, P. (2021). Deep learning based question generation using t5 transformer. In *Advanced Computing: 10th International Conference, IACC 2020, Panaji, Goa, India, December 5–6, 2020, Revised Selected Papers, Part I 10* (pp. 243-255). Singapore.
- Hiebert, J., & Lefevre, P. (1986). Conceptual knowledge in mathematics—An introductory analysis. *Conceptual and procedural knowledge–The case of mathematics*, 1-27.
- Jia, X., Wang, H., Yin, D., & Wu, Y. (2021). Enhancing question generation with commonsense knowledge. In *China National Conference on Chinese Computational Linguistics* (pp. 145-160). Cham: Springer International Publishing.
- Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The Knowledge-Learning-Instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive science*, *36*(5), 757-798.

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 989.

Kurdi, G., Leo, J., Parsia, B., Sattler, U., & Al-Emari, S. (2020). A systematic review of automatic question generation for educational purposes. *International Journal of Artificial Intelligence in Education*, *30*, 121-204.

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... & Zettlemoyer, L. (2019). Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*.

Liu, B. (2020). Neural question generation based on Seq2Seq. In *Proceedings of the 2020 5th International Conference on Mathematics and Artificial Intelligence* (pp. 119-123).

Liu, T., Fang, Q., Ding, W., Li, H., Wu, Z., & Liu, Z. (2020). Mathematical word problem generation from commonsense knowledge graph and equations. *arXiv preprint arXiv:2010.06196*.

Lopez, L. E., Cruz, D. K., Cruz, J. C. B., & Cheng, C. (2020). Transformer-based end-to-end question generation. *arXiv preprint arXiv:2005.01107*, *4*.

Moore, S., Nguyen, H. A., Bier, N., Domadia, T., & Stamper, J. (2022, September). Assessing the quality of student-generated short answer questions using GPT-3. In ECTEL 2023 (pp. 243-257).

Nguyen, H. A., Bhat, S., Moore, S., Bier, N., & Stamper, J. (2022). Towards generalized methods for automatic question generation in educational domains. In *European conference on technology enhanced learning* (pp. 272-284). Cham: Springer International Publishing.

Novikova, J., Dušek, O., Curry, A. C., & Rieser, V. (2017). Why we need new evaluation metrics for NLG. *arXiv preprint arXiv:1707.06875*.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, *1*(8), 9.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. (2020). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *arXiv:1910.10683* 

Rajpurkar, P., Zhang, J., Lopyrev, K., & Liang, P. (2016). Squad: 100,000+ questions for machine comprehension of text. *arXiv preprint arXiv:1606.05250*.

Ruseti, S., Dascalu, M., Johnson, A. M., Balyan, R., Kopp, K. J., McNamara, D. S., ... & Trausan-Matu, S. (2018). Predicting question quality using recurrent neural networks. In Artificial Intelligence in Education: 19th International Conference, AIED 2018, London, UK, June 27–30, 2018, Proceedings, Part I 19 (pp. 491-502). Springer International Publishing.

Sai, A. B., Mohankumar, A. K., & Khapra, M. M. (2022). A survey of evaluation metrics used for NLG systems. *ACM Computing Surveys (CSUR)*, *55*(2), 1-39.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational psychologist*, *24*(2), 113-142.

Stamper, John, Gaind, Bharat, Thankachan, Karun, Nguyen, Huy, and Moore, Steven (2023). Hierarchical Concept Map Generation from Course Data. In AAAI 2023 Workshop on Artificial Intelligence in Education (AI4Edu).

Storm, B. C., Bjork, R. A., & Storm, J. C. (2010). Optimizing retrieval as a learning event: When and why expanding retrieval practice enhances long-term retention. Memory & Cognition, 38, 244-253.

Van Der Lee, C., Gatt, A., Van Miltenburg, E., Wubben, S., & Krahmer, E. (2019). Best practices for the human evaluation of automatically generated text. In *Proceedings of the 12th International Conference on Natural Language Generation* (pp. 355-368).

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Wang, S., Wei, Z., Fan, Z., Huang, Z., Sun, W., Zhang, Q., & Huang, X. J. (2020). PathQG: Neural question generation from facts. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 9066-9075).

Xue, L., Constant, N., Roberts, A., Kale, M., Al-Rfou, R., Siddhant, A., ... & Raffel, C. (2020). mT5: A massively multilingual pre-trained text-to-text transformer. *arXiv preprint arXiv:2010.11934*.

Yu, J., Wang, Y., Zhong, Q., Luo, G., Mao, Y., Sun, K., ... & Sun, M. (2021). MOOCCubeX: a large knowledge-centered repository for adaptive learning in MOOCs. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (pp. 4643-4652).

Zhao, Y., Lei, J., Lai, B. Y. C., & Tan, H. S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record*, 107(8), 1836-1884.

# Preliminary Use of a Learning Game for Arithmetic Word Problems with Elementary School Students

Yusuke HAYASHI<sup>a\*</sup>, Kohei YAMAGUCHI<sup>a</sup> & Tsukasa HIRASHIMA<sup>a</sup> <sup>a</sup>Graduate School of Advanced Science and Engineering, Hiroshima University, Japan \*hayashi@lel.hiroshima-u.ac.jp

**Abstract:** Although word problems can be challenging for many learners, they are an important part of mathematics instruction. This is because word problems often require students to transfer abstract mathematical concepts to real-world situations, which can be difficult for some individuals. In this paper, we propose a learning game names "Triprop-scrabble" where learners can try to create arithmetic stories related to addition and subtraction from among many choices and can obtain feedback. To prove the research questions: (1) Can elementary first grade students create and chain arithmetic stories? and (2) Are they able to enjoy this game?, we conduct a preliminary use of the game on 44 first grade elementary school students who are the target of learning of arithmetic word problems of addition and subtraction. they can make arithmetic stories and some of them have made complex structures of stories. The results shows that this game is feasible for first-grade elementary school students.

Keywords: arithmetic word problems, problem-posing, augmented-reality

## 1. Introduction

Word problems are often considered challenging for mathematics learners and a significant amount of research has been dedicated to exploring this topic over the past 50 years. Word problems are verbal descriptions of problem situations in a scholastic setting that raise one or more questions. The answer can be obtained by applying mathematical operations to numerical data found in the problem statement or derived from it (Verschaffel et al. 2000). Solving an arithmetic word problem involves reading sentences, extracting the quantity relationships within them, representing them mathematically, and deriving unknown numbers (Mayer, 1992). Researchers have investigated the process of solving word problems and divided it into two sub-processes: (1) comprehension phase and (2) solution phase (Cummins et al., 1988) (Hegarty et al., 1995) (Heffernan and Koedinger, 1998) (Riley et al., 1983). They have also noted that the comprehension phase plays a major role in the difficulty of the word problem. During this phase, a learner must interpret representations expressed in words and create quantitative relationships. This can also be related to mathematical modeling, one of the important aspects in mathematics education, is mathematical modeling, which involves modeling real-world problems as "mathematical models" in the world of mathematics and solving them (Kaiser 2017). Although word problems are in a scholastic setting, if appropriately designed and handled in a mathematics educational context, they can act as accurate and valuable "simulacra" of authentic mathematical modeling problems (Verschaffel et al. 2020). This requires well-defined domain and instructional model (Mitrovic and Weerasinghe, 2009) of arithmetic word problems.

"Monsa-kun" (Hirashima et al., 2007) is a learning environment for arithmetic word problem learning based on the triplet structure model. This is based on the triplet structure model (Hirashima et al, 2014), which is a domain model for describing the relationship between the text of arithmetic word problems and mathematical equations. The characteristic of this model is that it breaks down the verbal description of arithmetic word problems into simple sentences as propositions of each quantity element in the described situation. In

"Monsa-kun" learners can carry out problem-posing as the integration of given simple sentences and receive immediate feedback. Sentence-integration is an instructional model of problem-posing in which learner can concentrate on the structure of arithmetic story not on generating story content. Through this approach, learners can concentrate on thinking about the relation between quantities in the situation and can also get feedback immediately.

The goal of this study is to transform the approach to thinking about the relationship between quantities in given situations, making it a more "open-ended" task. Monsa-kun tackles unit arithmetic word problems that can be solved with a single arithmetic operation, requiring learners to think in a "closed-ended" way where only one correct answer exists within the given constraints. In contrast, this study aims to introduce tasks that are "more open-ended" and have "enlarged structures." This is based on the continuum between well-defined and ill-defined problems (Le et al., 2013). Triplet structure model and problem-posing as sentence-integration can be well-defined domain and instructional models because they clarify the constraints in the domain of arithmetic problems/stories and guidelines to pose problems/stories as well as can provide immediate feedback to learners. This study aims at proposing a learning task as pro-ill-defined educational problem from Monsa-kun.

This study develops a learning game called "Tri-Prop-Scrabble" (Yamaguchi, 2023) in which learners can construct a chain of stories with the rule like Scrabble (Scrabble, n.d). This environment allows learners to create various arithmetic stories in a chain made up of single sentences and to consider the relationship between the quantitative elements. Posing problems is an effective way to learn a domain and posing arithmetic word problems is beneficial to learn the structure of arithmetic word problems. Although the target of Tri-Prop-Scrabble is not word problem but story, the basic goal is the same as Monsa-kun, that is, to consider the relationship between the quantitative elements in stories. In addition to that, Tri-Prop-Scrabble supports the understanding of the relationship between stories. We call this exercise "chained arithmetic story creation" and it involves creating compound arithmetic stories by predicting patterns for various stories with less conditions than Monsa-kun. This activity encourages learners to think the relationship between the verbal description and the mathematical formulas in a "more open-ended" manner in "more enlarged structures" than Monsa-kun. The results of experimental use by high school students shows that the exercise using the system is available for learners older than the main target of elementary school students (Yamaguchi, 2023).

This paper reports the result of the use of Tri-prop-scrabble by first grade elementary school students. The structure of this paper is as follows. The second chapter describe a relationship understanding support system based on the triplet structure model. The third chapter shows the design and the develop of Tri-Prop-Scrabble that allows learners to explore the possibilities of arithmetic stories as the combination of single sentences and a chain of them. The fourth chapter analyze the result of the use of Tri-prop-scrabble by elementary school students and consider the feasibility of the exercise for learning the relation between quantities in the situation. Finally, the last chapter present the conclusion and future tasks.

### 2. Modeling Arithmetic Stories with Triplet structure model

The triplet structure model is a domain model of arithmetic word problems by describing the relationship between verbal description and the mathematical formula in a situation. This model can be a well-defined model that breaks down the problem description into simple sentences for each quantity element and associates their semantic quantity relationship with the corresponding mathematical formula.

Monsa-kun is a "sentence-integration" typed problem-posing learning environment based on the triplet structure model, where learners can create problems as a combination of simple sentences and receive immediate feedback. Sentence-integration typed problemposing is a well-defined instructional model. This defines problem-posing as the task: integrating sentences representing numerical propositions and satisfying the constraints defined in the triplet structure model. Nakano (2002), Kurayama (2010) and Yamamoto (2012) show the effectiveness of Monsa-kun in learning arithmetic word problems even if learners pose problems with provided sentences. Figure 1 shows the interface of the problem creation screen of Monsa-kun. Learning by problem-posing is a method of learning by posing problems rather than solving them. This can be used to support the understanding of the relationship between elements in stories, which is important in understanding the relation between the problem statements and mathematical formulas. Although learning by problem-posing is considered an effective learning method for understanding the relation, it has not been widely used in educational situations due to the difficulty for both of learners and teachers. Monsakun has been developed to solve such problems and can execute sentence-integrated problem creation on a tablet device.

Monsa-kun deals with unit arithmetic stories that can be expressed by only one arithmetic operation and allows learners to understand the relation between stories and mathematical formulas through sentence integration. In the triplet structure model, a unit arithmetic story is formulated as a composition of two types of sentences: an existence sentence that represents the existence of a quantity and a relational sentence that represents the relationship between two quantities. A unit arithmetic word problem is a story in which one value is set as unknown and the unknown value can be derived from known values in the story by only one arithmetic operation. Although Monsa-kun only deals with stories that can express quantity relationships in one operation, the triplet structure model can express complex arithmetic stories as a chain of unit arithmetic stories. Monsa-kun encourages learners to be aware of the quantity relationship expressed in the text.

While the task setting in Monsa-kun is "closed-ended" and "a unit structure" problemposing, this study aims to handle more "open-ended" and "expanded structures" than Monsakun. Closed-ended environments are beneficial for learning basic structure, whereas openended environments are valuable for developing creativity based on the basic structure. Monsa-kun deals with unit arithmetic word problems and can help learners understand the basic structure of them as the relationship of problems that meet the constraints by presenting the conditions for creating arithmetic word problems. This system requires learners to think in a "closed" manner about the "minimum unit structure" by creating problems that meet the conditions.



Figure 1 Monsa-kun.

To support the understanding of the relationship between the verbal descriptions and the mathematical formulas in a variety of situations, we design and development an exercise environment in which learners can try to create arithmetic stories from among many choices and can obtain feedback. In the triplet structure model used in Monsa-kun, it is possible to create multiple stories from a single sentence as shown in Figure 2 when considered without conditions (open end), and it is possible to predict various stories. In this study, we call the activity of assembling stories from three quantities and creating a complex arithmetic story by continuously (compounding) creating it without conditions (open end) using some of those quantities as "chain story creation." This activity encourages learners to think in a "more open" manner and prompts them to think about "expanded structures" by allowing them to predict

patterns for various stories without conditions. The purpose of this study is to create an exercise environment that allows learners to create chain stories and consider various arithmetic stories in a chain from one sentence that is "diagnosable by the system" and supports the understanding of the relationship between the "real world" and the "world of mathematics" in mathematical modeling.

# 3. Tri-Prop-Scrabble

We propose the learning game "Tri-Prop-Scrabble," in which, like the English word game "Scrabble," players connect single-sentence cards to create and chain arithmetic stories as shown in Figure 3. Players are expected to learn the structure of arithmetic stories through exploring the potential combinations of single sentences in a chain of stories. The winner of this game is the player who creates the greatest number of arithmetic stories. Currently, as the first step in the design of this game, Tri-Prop-Scrabble only refers the basic rules of Scrabble. Tri-Prop-Scrabble is designed by the fusion method for game design (Umetsu et al., 2003). In the design of Tri-Prop-Scrabble based on Scrabble, letters are replaced with single-sentences, and English words are replaced with arithmetic stories. The thinking required to win this game is to consider the relation between the combination of single-sentences and mathematical formula and it is also required learning activity for arithmetic word problems.



Figure 2 an example of Possibilities of combining simple sentences.



Figure 3 Tri-Prop-Scrabble.

For example, when playing with multiple people, at the start of the game each player draws six single-sentence cards from the deck. In each turn, players try to create arithmetic stories and then, if a player create a story, the player check the validity of created story using the application named Story-AR-Checker shown in Figure 4. If a player cannot create an arithmetic story using their hand, they can exchange one or two cards from the deck or skip their turn. The game ends when all players cannot create an arithmetic story.



Figure 4 Story-AR-Checker.

In Tri-Prop-Scrabble, learners are expected to generate as many stories connected to the cards on the board as possible. Figure 5 shows an example of possibility for connectable stories in the situation. The edge of the cards placed on the board can be connected to other single-sentence cards, and players can find a place to create arithmetic stories and take out two cards to create a story. The left side of Figure 5 shows an example situation of the field and the hand. There are cards "There are five oranges" and "There are five apples" as the edge on the board, and six cards are in the player's hand. Learners are asked to generate possible stories that can be connected to the edges on the board. In this case, an arithmetic story "There are six apples" can be created by connecting "There are five apples"  $\rightarrow$  "I received one apple"  $\rightarrow$  "There are six apples," as shown in the photo on the right, and learners are asked to create a story by connecting the cards. The player who creates the most stories using their hand and the cards on the board wins. Through this activity, the aim is to encourage learners to consider various possibilities of stories without any conditions (open-ended) by combining them in a chain, promoting "more open" thinking that deals with "expanded structures."

Each card represents a simple sentence and the color represent the types of sentences. Existence sentence has a gray frame, and relational sentence has a frame with another color. In this version, cards must be arranged in the order of existence sentences (gray frame), relational sentences (another colored frame), and the other existence sentence (gray frame). The difference in color only indicates whether the card is an existence sentence (gray) or a relational sentence (other colors). Also, each card has an illustration that characterizes the card for the Story-AR-Checker to recognize. This illustration is merely a technique to improve the recognition accuracy of the Story-AR-Checker and does not represent the content of the single sentence written on the card.

With Story-AR-Checker shown in エラー! 参照元が見つかりません。, players can diagnose the validity of the arithmetic stories they have made. Learners can connect single-sentence cards they have in hand to create arithmetic stories and compete with the number of stories created. Learners can determine the validity of the created story by scanning the cards with the Story-AR-Checker. The system recognizes the story created by arranging the cards in physical space by reading the marked single-sentence cards with a tablet camera. The validity of the story is judged based on three criteria: "problem structure," "object," and "formula." The problem structure confirms that the created problem consists of two existence

sentences and one relational sentence, the object confirms that the object corresponding to the relational sentence "one unit of number" is used, and the formula confirms that the created problem is valid as an equation. The system judges the validity based on whether the created story meets these three criteria. In this game, players take a picture of the arrangement of cards on the board using the camera of a tablet and Story-AR-Checker scans the cards to identify the story created by the cards. If the order of the cards that make up the story is identified, an indicator representing the order of the sentence is displayed on the scanned card image on the tablet screen, as shown in Figure 4. Vuforia (PTC, Inc., n.d) Augmented Reality SDK and Unity (Unity Technologies, n.d) are used as the development environment for this system.



Figure 5 an example of Thinking for chaining arithmetic stories

## 4. Analysis of the Preliminary Use by Elementary School Students

To verify whether exercises with Tri-prop-scrabble for elementary school students can be feasible, 44 first graders of an elementary school in Japan played Tri-prop-scrabble with Story-AR-Checker. Due to the influence of the COVID-19 pandemic, the students played not in group but individually. Before playing Tri-Prop-Scrabble, they had the exercises Monsa-kun and Monsa-kun Tape Block (Hayashi, 2021). Through the exercise they had time to understand problem posing by sentence-integration based on Triplet structure model. In addition to that, before and after the exercises they also took the tests to solve arithmetic word problems.

The exercise including the tutorial was conducted about 30 minutes. At the beginning of the use of Tri-Prop-Scrabble, the author explained the rules of Tri-Prop-Scrabble and how to use Story-AR-Checker in the game. However, since this exercise was conducted individually, there were no rules such as rotation of turns and winner determination, and participants were simply asked to create as many stories as possible using the provided cards. The participants were allowed to confirm the rules of the game and how to use the system by asking questions to the authors during the game.

The research questions of this exercise are as follows:

- (1) Can elementary first grade students create and chain arithmetic stories?
- (2) Are they able to enjoy this game?

To confirm (1), we analyze the numbers of operations with Story-AR-Checker, stories and chains created by the students. If the students keep trying to create stories and most of students can create and chain stories, we can say the learning exercises with Tri-propscrabble is feasible for the main target of it, first and second graders in elementary schools in Japan. To confirm (2) whether participants can enjoy this game, a questionnaire about system usage was conducted. Figure 6 is the boxplot of the numbers of operations with Story-AR-Checker and shows that students continuously kept playing Tri-prop-scrabble. From 6 minutes to 12 minutes later, this is just after the explanation in the tutorial and the students try to create the example shown by the authors. From 12 minutes to 18 minutes later, the authors made explanation of the rules again and then start the free exercise. From 18 minutes to 30 minutes later, they constantly used Story-AR-Checker.



Figure 6 Number of operations on Story-AR-Checker per 3 minutes.

Figure 7 and Figure 8 show the number of arithmetic stories and chains created by the learners, respectively. These data show that most of student could create more than one story and it is difficult for them to chain arithmetic stories. The median number of stories created was 1, with a maximum of 10 and a minimum of 1. As for the number of chains, the median number was 1.84, with a maximum of 5 and a minimum of 0. From these data, whereas most of students could create at least one or two arithmetic stories, it is difficult for them to chain stories. It can also be said that this is a valid result because in Japanese standard curriculum children learn arithmetic word problems requiring multiple calculations and a chain of arithmetic stories is the base of such problems. Although chaining arithmetic stories is the out of the curriculum, more than half of students could chain arithmetic stories. This shows the potential of learning for them. On the other hand, a few students created only one story including the example shown at the tutorial. However, even such students try to create arithmetic stories.

Furthermore, this result suggests the further potential of Tri-Prop-Scrabble, which promotes thinking for chaining arithmetic stories, for learners. Figure 9 shows the structure of arithmetic stories created by a learner who had the highest number of chains. The feature of this structure is that there is an arithmetic story loop in the center. In this structure, not only simply connecting stories but also more complex thinking is required. A loop is the minimum structure to compose four arithmetic stories and it requires to consider the constraints of arithmetic stories in a challenging situation. For example, by providing motivation for creating a loop, Tri-Prop-Scrabble has the potential to offer learners a different perspective. Learners can consider at a various level depending on their understanding on Tri-Prop-Scrabble.

To confirm (2) whether participants can enjoy this game, we conducted the questionnaire about this game. As a result, many positive evaluations were obtained regarding playing this game. The questionnaire results in Figure10 show that participants felt "want to try again," "understand how to use the system," and "the game is fun." However, some participants answered "difficult to use" in the free description section. This suggests that there are points to be improved in Story-AR-Checker.

From these results, we answer yes to the research question (1) and (2). For the research question (1): "can elementary first grade students create and chain arithmetic stories?", the students continuously kept playing Tri-prop-scrabble and most of students could create more than one arithmetic stories. Even the students could not create any arithmetic stories by themselves kept playing this game. For the research question (2): "are they able to enjoy this game?", most of students gave positive answers. This means learners could enjoyed Tri-prop-scrabble as a game even though it required thinking about the structure of arithmetic stories.



Figure 7 numbers of stories created by learners.



Figure 8 numbers of chain created by learners.

# 5. Summary

In this study, we developed a game called "Tri-Prop-Scrabble" that explores arithmetic stories as an integration of simple sentences, and a diagnostic system called "Story-AR-Checker" that supports it with AR technology. The game allows players to explore arithmetic stories that can be created using the simple sentences on the board and in their hands. Story-AR-Checker uses a camera to recognize cards and judge the validity of the created arithmetic stories. Th

use by the elementary school students as the main target of this game shows that the students were able to create arithmetic stories, and some were able to chain arithmetic stories. From these results, it was suggested that exercises with Tri-prop-scrabble are feasible for elementary school students and that this exercise can be one that encourages the consideration of the relationships between the verbal description and mathematical formula in arithmetic stories. Of course, this doesn't show that learning has happened, and we heed to investigate the effectiveness of this game for learning of arithmetic word problems. In the future, we plan to analyze practical data, expand the scoring rules to encourage consideration of relationships, and further verify the learning effects of the proposed game.



Figure 9 chained arithmetic stories with a loop



Figure 10 the questionnaire results

# Acknowledgements

I am grateful to Teachers of KuchitaHigashi elementary school in Hiroshima city for collaboration on the exercise of Tri-prop-scrabble.

#### References

- Cummins, D., Kintsch, W., Reusser, K., Weimer, R. (1988). The role of understanding in solving word problems. Cognitive Psychology, 20, 439-462.
- Hayashi, Y., Tsudaka, N., Iwai, K., & Hirashima, T. (2021). *Problem posing with representation conversion model for learning the condition of addition and subtraction word problems*. Proc. of ICCE2021, 2, 60-69.
- Heffernan, N., Koedinger, K. (1998). A developmental model for algebra symbolization: The results of a difficulty factors assessment. Proceedings of the twentieth annual conference of the cognitive science society, 484-489.
- Hegarty, M., Mayer, R.E., Monk, C.A. (1995). Comprehension of arithmetic word problems: a comparison of successful and unsuccessful problem solvers. Journal of Educational Psychology, Vol.87, No.1, 18-32.
- Hirashima, T., Yokoyama, T., Okamoto, M., Takeuchi, A. (2007). Learning by problem-posing as sentence-integration and experimental use. AIED 2007, 254-261.
- Hirashima, T., Yamamoto, S., & Hayashi, Y. (2014). Triplet Structure Model of Arithmetical Word Problems for Learning by Problem-Posing. *Human Interface and the Management of Information. Information and Knowledge in Applications and Services*, 42–50.
- Kaiser, G. (2017). The teaching and learning of mathematical modeling. In J. Cai (Ed.), Compendium for research in mathematics education (pp. 267–291). Reston, VA: National Council of Teachers of Mathematics
- Kurayama, M., & Hirashima, T (2010) Interactive Learning En-vi-ronment Designed Based on Task Model of Problem-Posing, Proceedings of ICCE2010, 98-100.
- Le, N.-T., Loll, F., & Pinkwart, N. (2013). Operationalizing the Continuum between Well-Defined and III-Defined Problems for Educational Technology. IEEE Transactions on Learning Technologies, 6(3), 258–270.
- Mayer, R.E. (1992) Thinking, problem solving, cognition, Second ed.,455–489, W.H. Freeman, New York.
- Mitrovic, A. and Weerasinghe, A. (2009). Revisiting III-Definedness and the Consequences for ITSs, Proc. 14th Int'l Conf. Artificial Intelligence in Education (AIED '09), 375-382.
- Nakano, A., Hirashima, T., Takeuchi, A (2002) An Evaluation of Intelligent Learning Environment for Problem Posing, Proceedings of ITS2002, pp.861-872,
- PTC, Inc. (2023) Vuforia, Retrieved May 21, from https://developer.vuforia.com
- Riley, M. S. Greeno, J. G. Heller, J. I. (1983). Development of children's problem solving ability in arithmetic, H. P. Ginsburg (Ed.). The development of mathematical thinking. New York: Academic Press, 153-196.
- Scrabble (Aug. 10, 2023, 22:52 UTC). Retrieved Aug. 14, 2023, In Wikipedia: The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Scrabble
- Umetsu, T., Hirashima, T., & Takeuchi, A. (2003). A computer-based learning game designed by fusion method. The Journal of Information and Systems in Education, 2(1), 7-13
- Unity Technologies, Retrieved May 21, 2023, from https://unity.com
- Verschaffel, L., Greer, B., & De Corte, E. (2000). Making sense of word problems. Lisse: Swets & Zeitlinger.
- Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: a survey. *ZDM: The International Journal on Mathematics Education*, 52(1), 1–16.
- Yamaguchi, K., Hayashi, Y., & Hirashima, T. (2023). Design and Development of a Card Game for Learning on the Structure of Arithmetic Story by Concatenated Sentence Integration. IEICE Transactions on Information and Systems. E106-D(2), 131-136.
- Yamamoto, S., Kanbe, T., Yoshida, Y., Maeda, K., & Hirashima, T. (2012) A case study of learning by problem-posing in introductory phase of arithmetic word problems. In Proceedings of the 20th International Conference on Computers in Education, 25-32.
## The Effect of Revision-Making of Contextualized Student-Generated Questions on University Students' English Grammar Learning Performance

Chih-Chung LIN & Fu-Yun YU\*

Institute of Education, National Cheng Kung University, Taiwan \*fuyun.ncku@gmail.com

**Abstract:** There is a great deal of research supporting the application of studentgenerated questions (SGQ) to English curriculums, and recent studies have found L2 learners to benefit further from an elaborated SGQ approach — contextualized SGQ (cSGQ), which leverages the provision of a context for students to base on their generated questions. Previous research has also proved the advantages of making revision on English learner's writing proficiency but its relevance to English grammar learning is yet to be exploited. Thus, the present study aimed to investigate the effect of cSGQ learning tasks combined with the practice of making revision of the given context on L2 learners' English grammar learning performance. A pre- and post-test quasi-experimental research design was adopted, and 88 university students enrolling in freshman English class were invited to participate. The results showed that students in the cSGQ-revision group outperformed those not given the opportunity to make revision in both linguistic knowledge and pragmatic understanding.

**Keywords:** contextual learning, English grammar learning, online student-generated questions, revision-making

#### 1. Introduction

English grammar has been considered one of the essential elements in teaching and learning foreign languages and Zhou (2018) suggested that English grammar is "a necessary framework in a language system". As basic as other components like vocabulary and pronunciation tend to be, English grammar plays an important role in monitoring if those basic components are in their correct places in sentences (Borjars & Burridge, 2019).

When it comes to teaching and learning English grammars, which involves linguistic knowledge and pragmatic understanding, researchers in functional schools in the field of linguistics highlight the significance of the communicative aspects of using languages rather than the memorization of the rules of the target languages (Feng, 2013). Learners are believed to develop communicative competence by actually using the linguistic knowledge and cultivate their pragmatic understanding with the provision of contexts (Schwarz & Hamman-Ortiz, 2020), which refers to the comprehension of the language used within contexts, where language users should accurately decode the contextual meaning of the used language in order to achieve the success of communication (Bachman & Palmer, 2010).

Student-generated Questions (SGQ) has been considered one of the innovative ways of teaching and learning English, which involves having students construct questions and corresponding answers on the learning topic (Yu, 2020; Yu, 2021a). Several advantages of SGQ on teaching and learning have been found, including demonstrating and verifying one's understanding and knowledge of the learning materials, resolving misconceptions, and filling knowledge gaps (Juan, 2021; Offerdahl & Montplawasir, 2014). Moreover, in the process of completing SGQ tasks, several effective learning strategies were activated, such as reviewing

the learning materials, finding core concepts, and transforming the learned knowledge into forming new cognition (Bangert-Drowns, et al., 2004; Yu, 2012; Yu & Liu, 2008).

Recently, in view of testlet increasingly prevalence in contemporary assessment practice and distinctive benefits (e.g., efficiency in terms of item development and test administration) (Keng et al., 2008; Lane et al., 2016), Yu (2021b) developed the world's firstever online student-generated testlets system to support students' formulation of a group of related question items on the basis of a given situation or scenario, and its learning potential as a meaningful extension of SGQ was preliminarily assessed. Leveraging the power of contextual learning (Ampa, et al., 2013; Surya, et al., 2017), 'contextualized student-generated questions' (cSGQ) was later coined (Cheng & Yu, 2021; Lin & Yu, 2021), and the importance of the provision of a context where meaningful and relevant information is provided for students to base their generated questions is stressed. That is, students are given a specific scenario where students generate questions, which should correspond to the given scenario and reflect the situational clues and details of the given context, such as the characters in the story and the timeline of the events. In so doing, students should not only detect and analyze the important information presented in the given context but also apply the learned linguistic knowledge and information embedded in the given context for successful attainment of the cSGQ learning tasks (Lin & Yu, 2021; Yu, 2021b).

Despite that a couple of follow-up preliminary studies supported and substantiated the learning effects of cSGQ (Lin & Yu, 2021; Lin & Yu, 2022), existing studies involved students generating questions based on a fixed scenario. In light of revision-making, defined as the process in which writers could go back and forth in order to move forward (Zamel, 1982) and also known as process-based approach, has been commonly practiced in L2 writing classroom and Goldstein (2015) suggested that revision in writing was a process with various factors interacting and mediating each other within the texts in which students' texts were created through cyclical processes. From the socio-cognitive view on revision, revision making refers to the process where learners make changes to the texts, learning the conceptual ideas of using words, phrases, sentences and paragraphs, and modifying both linguistic and non-linguistic forms within the texts so as to achieve the purpose of successful communication (Finkenstaedt-Quinn et al., 2021; Gere et al., 2019).

Previous research has proven the critical role of making revision on learners' writing process (Barkaoui, 2016; Hayes, 2012). In the process of revision, writers gain insights into and heighten awareness to both linguistic and non-linguistic communicative functions of the language used in the texts, noticing possible problems such as the spelling mistakes and incoherence of the texts, and coming up with solutions to the identified text problems, which may contribute to students' repeatedly read the writing passages, finding and addressing the potential problems within the texts, and thinking better ways to improve the quality of the work (Finkenstaedt-Quinn, et al., 2021).

With the advantages of making revision on the English writing, the present study aimed to investigate the possible influence of making revision combined with cSGQ learning activities on English learners' grammar learning performance in terms of linguistic knowledge and pragmatic understanding. To be more specific, two research questions are examined in the study:

RQ#1 Do students in the cSGQ-revised group have better English grammar learning performance in terms of linguistic knowledge than the group with a fixed, given scenario for SGQ (i.e., the originally conceived cSGQ)?

RQ#2 Do students in the cSGQ-revised group have better English grammar learning performance in terms of pragmatic understanding than the fixed-scenario SGQ group?

#### 2. Method

#### 2.1 The Participants and Study Design

The participants of this study were college students and their English proficiency level was B1 according to the Common European Framework of Reference for Language Learning,

Teaching, Assessment (CEFR). A pre- and post-test quasi-experimental research designs was employed, and two treatment groups were devised – the comparison group (the cSGQ-fixed group) (N = 44) and the experimental group (cSGQ-revised group) (N = 44).

The topics for cSGQ learning activities were tenses and subjunctive moods. Before the experiment, all participants in the two groups took a pre-test to assess their knowledge on the two targeted English grammars, tenses and subjunctive moods. In total, four cSGQ learning activities were arranged and took place right after instruction on tenses and subjunctive moods. Training on the essential skill and knowledge on SGQ, cSGQ and system use was arranged in consideration that all participants didn't have experience in SGQ before the study. Finally, all participants took the same set of tests to assess their competencies on the two targeted English grammars.

#### 2.2 The Learning System and the Design of Revision Tasks

The online student-generated testlets learning system developed by one of the authors was used in this study (Yu, 2021b). As shown in Figure 1, the half-bottom part is the area for both groups' participants to generate a set of questions on the basis of the given scenario (the half-top portion) for the cSGQ learning tasks. As mentioned, the participants didn't have prior experience in SGQ, a button placed at the upper-left corner above the SGQ field was incorporated as procedural prompts to link to the main idea with key ideas and example sentences for students to refer to when generating questions on the targeted English grammar.

As for system use, the only difference between the two groups lies in that for the participants in the cSGQ-fixed group, the given context (i.e., the top portion) field were fixed so making revision of the given context was not possible whereas the participants in the cSGQ-revised group were allowed to make revision of the given context at their discretion. With revision-making in place, the participants can edit the given scenario by deleting, re-arranging, or adding any words/phrases/sentences before proceeding to cSGQ learning tasks. Such revisions are expected to influence students' use of English vocabulary and grammars, which, in turn, is expected to influence students' English grammar learning performance. As scaffolding for making revision, the participants in the cSGQ-revised group can click on the button placed at the top of the context field, so as to refer to the hints and important notes, for example, the subjects, the time, the sequence of action.



Figure 1. A Screenshot of Testlet for cSGQ-revised learning tasks.

#### 2.3 The Instruments

There were 50 multiple-choice questions with 25 on tenses and with 25 on subjunctive moods. There were 36 questions for the test on linguistic knowledge and 14 questions for the evaluation on pragmatic understanding. The pre- and post-test were made by the authors and were validated by two professors who had more than 10 years of experience in teaching English at universities across southern Taiwan. They helped to correct the wording of the questions, choices and answers and provided feedback and suggestions on the refinement of the questions to make the questions congruent with the teaching objective and the experiment.

#### 2.4 Data Analysis

For the purpose of this study, one-way analysis of covariance (ANCOVA) was conducted on the pre- and post-test of students' English grammar learning performance in terms of linguistic knowledge and pragmatic understanding.

#### 3. Results and Discussion

#### 3.1 Results Regarding the Participants' Learning Performance in Terms of Linguistic Knowledge

Based on the results shown in Table 1, both groups improved their linguistic knowledge after the experiment, and the adjusted score of the cSGQ-fixed group was lower than that of the cSGQ-revised group.

 Table 1. The Descriptive Statistics of Pre-, Post- and Adjusted Post-test of the cSGQ-revised and cSGQ-fixed Groups in Terms of Linguistic Knowledge

		<u> </u>	0			
Groups	Pre-test		Pos	st-test	Adjusted post-test	
	Μ	SD	М	SD	Μ	
cSGQ- revised	30.31	7.59	40.59	10.29	40.14	
cSGQ-fixed	24.18	8.02	31.05	9.45	34.43	

With the *F* value of the interaction not reaching statistical difference, 1.53 (p = .22 > .05), it means that the regression coefficients contain homogeneity, and ANCOVA could be further performed. Using the pre-test scores as covariates, the results of the analysis of ANCOVA between the two groups shows that the *F* value between the two groups was 43.8 (p = .00 < .05). With statistically significant difference, it means that the two cSGQ activities influenced participants' learning in terms of linguistic knowledge.

#### 3.2 Results Regarding the Participants' Learning Performance in Terms of Pragmatic Understanding

Based on the results shown in Table 2, the two groups improved their pragmatic understanding after the experiment, and the adjusted score of the cSGQ-fixed group was lower than that of the cSGQ-revised group.

and cSGQ-inced Groups in Terms of Pragmatic Onderstanding									
Groups	Pre-test		Pos	t-test	Adjusted post-test				
	М	SD	М	SD	М				
cSGQ-	18.77	4.09	21.13	3.77	19.77				
revised									
cSGQ-fixed	10.18	7.05	15.18	6.05	18.03				

 Table 2. The Descriptive Statistics of Pre-, Post- and Adjusted Post-test of the cSGQ-revised and cSGQ-fixed Groups in Terms of Pragmatic Understanding

The *F* value of the interaction was 0.77 (p = .46 > .05), not reaching statistical difference, and suggesting that ANCOVA could be further performed. The results of the analysis of ANCOVA between the two groups show that, using the pre-test scores as covariates, the *F* value between the two groups was 40.26 (p = .00 < .05), reaching statistical difference. It means that the two cSGQ activities influenced participants' learning in terms of pragmatic understanding.

3.3 Discussion on Revision-making on Students' English Grammar Learning

#### Performance

As found in this study, the participants in the cSGQ-revised group had better learning performance in terms of linguistic knowledge and pragmatic understanding than those in the cSGQ-fixed group. From the of socio-cognitive perspective, learners in the process of making revision would gain insights and awareness of the used linguistic knowledge and situational information (Barkaoui, 2016; Hayes, 2012). That is, when making revision, learners would pay more attention and better figure out what and how to improve the quality of their written texts, noticing possible problems, thinking and trying to figure out solutions to the currently encountered problems. Such processes would provide learners with more opportunities to rehearse the learned content and consolidate their knowledge before, during and after making revision, which should contribute to better learning performance. For example, when leaners try to modify the sequence of the story, they need to familiarize themselves with the use of different tenses in order to successfully change the order of the events as depicted in the scenario.

#### 4. Conclusion

The present study elaborates cSGQ learning tasks by introducing the element of revisionmaking of the given context (cSGQ-revised) for EFL learners, hoping to further help develop their target language proficiency as compared to the cSGQ-fixed learning tasks. The obtained results demonstrated that students in the cSGQ-revised group outperformed those in the cSGQ-fixed group in terms of both linguistic knowledge and pragmatic understanding. The findings of the present study provide the empirical support on the effect of cSGQ learning tasks with making revision on L2 learners' grammar learning performance.

Despite the potential benefit of cSGQ-revised learning activities on L2 learners, some limitations of using this innovative teaching and learning approach should be noted. Specifically, the present study focused on English grammar learning performance, and future studies could explore the potential of cSGQ-revised learning activities, say, on L2 learners' receptive and receptive skills.

#### Acknowledgements

This study was supported by the Ministry of Science and Technology, Taiwan (MOST 108-2511-H-006-007-MY3).

#### References

- Ampa, A. T., Basri, M., & Andriani, A. A. (2013). The development of contextual learning materials for the English speaking skills. *International Journal of Education and Research*, *1*(9), 1-10.
- Bachman, L., & Palmer, A. (2010). Language assessment in practice. Oxford: Oxford University Press.
- Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-tolearn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, 74(1), 29–58.
- Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Journal, 100*(1), 320-340.
- Borjars, K., & Burridge, K. (2019). Introducing English Grammar (2nd ed.). London: Hodder Education.
- Cheng, W-W & Yu, F. Y. (2021). Leveraging context for computer-supported student-generated questions and EFL learning in grammar instruction: Its effects on task performance. In M. M. T. Rodrigo, S. Iyer, A. Mitrovic (Eds), *29th International Conference on Computers in Education Conference Proceedings II* (pp. 328~337), November 22-26, Bangkok, Thailand.
- Feng, Z. (2013). Functional grammar and its implications for English teaching and learning. *English Language Teaching*, *6*(10), 86-94.

- Finkenstaedt-Quinn, S. A., Polakowski, N., Gunderson, B., Shultz, G. V., & Gere, A. R. (2021). Utilizing peer review and revision in STEM to support the development of conceptual knowledge through writing. *Written Communication*, *38*(3), 351-379.
- Gere, A. R., Limlamai, N., Wilson, E., MacDougall Saylor, K., & Pugh, R. (2019). Writing and conceptual learning in science: An analysis of assignments. *Written Communication*, *36*(1), 99-135.
- Goldstein, L. M. (2005). *Teacher written commentary in second language writing classrooms*. Ann Arbor, MI: University of Michigan Press.
- Hayes, J. R. (2012). Modeling and remodeling writing. Written communication, 29(3), 369-388.
- Juan, S. (2021). Promoting engagement of nursing students in online learning: Use of the studentgenerated question in a nursing leadership course. *Nurse Education Today*, 97, 104710-104714.
- Keng, L., Ho, T-H., Chen, T-A. & Dodd, B. (2008). A comparison of item and testlet selection procedures in computerized adaptive testing. Paper presented at the *National Council on Measurement in Education*. March 24th, New York City.
- Lane, S., Raymond, M. R. & Haladyna, T. M. (2016). *Handbook of test development* (2<sup>nd</sup> Ed). NY: Routledge.
- Lin, C-C & Yu, F. Y. (2022). The effect of contextual student-generated questions on EFL learners' English learning performance, language learning strategy use, and perceived cognitive load. S. Iyer, J. L. Shih, W. Chen, M. N. MD Khambari (Eds), *30th International Conference on Computers in Education Conference Proceedings Volume II* (pp. 24~33), Kuala Lumpur, Malaysia, November 28~December 2.
- Lin, C. C. & Yu, F. Y. (2021). The design and effects of online contextual student-generated questions for English grammar learning. In M. M. Rodrigo, S. I., A. Mitrovic (eds), 29th International Conference on Computers in Education Workshop Proceedings, Volume I (pp. 92-101), November 22-26.
- Offerdahl, E. G., & Montplawasir, L. (2014). Student-generated reading questions: Diagnosing student thinking with diverse formative assessments. *Biochemwastry and Molecular Biology Education*, *42*(1), 29-38.
- Schwarz, V. S., & Hamman-Ortiz, L. (2020). Systemic functional linguistics, teacher education, and writing outcomes for US elementary English learners: A review of the literature. *Journal of Second Language Writing*, *4*9, 100727.
- Surya, E., & Putri, F. A. (2017). Improving mathematical problem-solving ability and self-confidence of high school students through contextual learning model. *Journal on Mathematics Education, 8*(1), 85-94.
- Yu, F. Y. (2021a). Student-generated questions with technology-mediated socio-constructivist learning approaches. Panel on 'Leveraging student-generated ideas (SGI) to facilitate socio-constructivist learning and conceptual change: The roles of technology in SGI learning trajectories. *Proceedings* of the 29th International Conference on Computers in Education (ICCE2021), November 22-26, Bangkok, Thailand.
- Yu, F. Y. (2021b). Development and preliminary evaluation of the learning potential of an online system in support of a student-generated testlets learning activity. In M. M. T. Rodrigo, S. Iyer, A. Mitrovic (Eds), 29th International Conference on Computers in Education Conference Proceedings I (pp. 638~643), November 22-26, Bangkok, Thailand.
- Yu, F. Y. (2020). Online student-generated questions 2.0—further enhancement of teaching and learning. Keynote speech delivered at the 24<sup>th</sup> Global Chinese Conference on Computers in Education (GCCCE2020), September 12~16, Lanzhou, China.
- Yu, F. Y. (2012). Keynote Speech: Learner-centered pedagogy + adaptable and scaffolded learning space design—online student question-generation. Keynote speech delivered at *International Conference on Computers in Education* (ICCE) 2012, November 26-30, Singapore.
- Yu, F. Y. & Liu, Y. H. (2008). The comparative effects of student question-posing and questionanswering strategies on promoting college students' academic achievement, cognitive and metacognitive strategies use. *Journal of Education & Psychology*, 31(3), 25-52.
- Zamel, V. (1982). Writing: The process of discovering meaning. TESOL quarterly, 16(2), 195-209.
- Zhou, C. (2018). Teaching model of college English grammar in intensive reading course. *Educational Sciences: Theory & Practice*, *18*(6), 2617-2632.

### Review of the Literature on Question-Posing-based Self-Directed Learning

Sumitra SADHUKHAN<sup>a\*</sup>, Shitanshu MISHRA<sup>b</sup> & Sridhar IYER<sup>a</sup> <sup>a</sup> Indian Institute of Technology, Bombay, India <sup>b</sup>UNESCO MGIEP, India <u>\*sumitra.sadhukhan@iitb.ac.in</u>

#The views and content presented in this paper are the sole responsibility of the individual author and do not represent the views or opinions of the respective organization.

**Abstract:** Self-directed learning presents a promising approach to enable students to achieve self-reliance and cultivate greater ownership over their learning pursuits. In the current teaching-learning scenarios, the implementation of classroom-oriented self-directed learning (SDL) strategies does not achieve their desired goal. The primary reason may be that the teachers are not trained to implement these strategies. According to the literature, learners face difficulties in following different SDL strategies. Question posing (QP) serves as a tool that can help learners understand the concept, clear misconceptions, and collaborate with peers to construct their knowledge. We have analyzed existing papers on question-posing-based classroom-oriented self-directed learning strategies to find how question-posing may help students overcome their difficulties and find ownership in their learning. We find that QP-based teaching-learning approaches require well-designed activities that can help students elicit questions and with which students can navigate in achieving ownership of their learning. The paper concludes with the identification of a QP-based strategy that conforms to the essential set of requirements of self-directed learning.

Keywords: Self-directed learning, question posing, learner difficulty.

#### 1. Introduction

The 21<sup>st</sup> century is a technology-driven era. It has changed the conceptions of the way to learn and teach, the characteristics of classroom environments, students' thinking processes, the availability, and type of learning resources, and the teacher's role. In these rapidly changing and challenging environments, students want to learn topics according to their interests and follow strategies that will be under their control. Self-directed learning empowers students by giving them the agency to drive their learning. The self-driven learning environments should be based on the individual students' perspectives and allow students to interact with the educational design constructs as per their interests (Lorås et al., 2021). As pointed out by Kim & Zitzer (1999), in a self-directed learning process, students take charge of their own learning, and teachers are facilitators who understand the different needs of students in terms of students' knowledge and interests. Charoula (1998), emphasized this learning process satisfies individual students' needs at different rates.

Self-directed learning helps students adopt a thorough and deep method of studying (Bell et al., 2019) and enhances their educational experience (Pott et al., 2021). It is a powerful approach for increasing student engagement. It is an approach in which students construct and reconstruct knowledge dynamically and their growing engagement is supported by motivational, cognitive, and social aspects (Martinetti, 2020). Student-directed learning is based on constructivist theory, which defines knowledge as "temporary, developmental,

socially and culturally mediated and non-objective" (Brooks & Brooks, 1993). Within this paradigm, students are active knowledge constructors driven by their goals and curiosities. Students take control and ownership of their learning, (Vyas, 2018). In such classrooms, the students are at the center of their learning. As a descriptor for constructivist instructional approaches, within this paradigm, the teacher gives up the role of the transmitter of knowledge and interacts with the students as a facilitator with the perspective that learning involves the active construction of meaning (Biggs, 1989) by the students. In the SDL environments, students develop deep approaches to learning. The deep approaches are based on student interest in the subject matter. The strategy here is to maximize understanding so that curiosity is satisfied. It instills a sense of autonomy and mastery in the student, and a sense of purpose into the educational process.

Knowles (1975), defined self-directed learning as a five-phase process, in which students (i) recognize their learning needs, (ii) formulate learning goals, (iii) select learning resources, (iv) employ appropriate learning strategies, and (v) assess learning outcomes. (In this paper, we refer to them as SDL processes). Conceptions of learning vary from superficial to deep and the progression from surface to deep is aided by classroom processes that include student-driven learning and meta-learning (Biggs, 1989). To achieve self-directed learning, several strategies like project-based learning, problem-based learning, inquiry-based learning, etc. were studied by the researchers, which state, inbuilt complexities about these strategies to apply in a classroom setting. These complexities and challenges are faced both by the students and the educators (Hwang et al., 2020).

Though self-directed learning and student-directed learning are similar, there is a slight difference between them. While both self-directed learning and student-directed learning focus on individual autonomy and independence, student-directed learning involves more guidance and support from teachers or facilitators. In self-directed learning, learners have the freedom to make decisions and choices about their learning, however, in student-directed learning process. Self-directed learning skills continue to develop as individual learning progresses through early adolescence to adulthood (Brandt, 2020). Brandt (2020) proposed a self-directed learning continuum, where a student's self-directedness gradually grows from a dependent student to a self-determined student.

Due to the multifaceted definition of self-directed learning, often self-regulated learning is synonymously used with self-directed learning. Self-regulation is a more narrowly defined concept, representing one dimension of self-directed learning, according to Brandt (2020). Self-directed learning emphasizes learner autonomy in selecting learning goals and learning tasks, while self-regulated learning focuses on learner active management and regulating their learning process (Mulyawati, 2020).

In traditional teaching-learning environments, where teachers are not adequately trained on their roles and responsibilities as facilitators of student learning, shifting from teacher-driven learning to self-directed learning is challenging, and can lead to uncertainty in methods of implementation (Shpeizer, 2019). In their paper, Cintang et al. (2018), explained that most teachers, even experienced teachers, may face difficulties in implementing pedagogical strategies toward SDL goals. One of the primary sources of the difficulties as discussed by the authors is the fact that teachers are not trained to implement SDL strategies. Pablos et al. (2020), highlighted that due to lack of guidance, high school students did not know what they had to do during the project implementation. Literature reports that the students face difficulties in engaging with the traditional pedagogies toward SDL goals (Cintang et al., 2018). For example, Hussein et al. (2021), reported the lack of collaboration ability among tertiary students with their peers that leads to priority conflicts while implementing a project-based learning strategy.

Many of the existing pedagogies enable students with the autonomy to follow the phases of self-directed learning, however, most of them do not provide adequate support to the students to transition toward the specific goals of SDL (Emily et al., 2019). Wu (2020), found differences between high- and low-creativity tertiary students to execute the SDL strategies. Providing absolute autonomy can be devastating if students don't know what they should learn, where they should learn from, and how to learn. Ideally, a pedagogical strategy should enable students to execute all the processes of SDL. This is a challenge for the pedagogy designers.

In Table 1, we have synthesized student difficulties in achieving SDL outcomes in different pedagogical strategies from the literature. The second column of Table 1 represents different pedagogical strategies that the researcher followed to achieve SDL. The third column represents student difficulties in following the corresponding SDL process. The Last column in the table emphasizes the hindered SDL process(es) in any given pedagogical strategy. The SDL processes are encoded as (A) recognizing learning needs, (B) formulating learning goals, (C) selecting learning resources, (D) employing appropriate learning strategies, and (E) assessing learning outcomes respectively. Most of the pedagogies worked on the process D.

Reference	Pedagogical strategies	Student difficulties	Obstructed SDL Process
Hussein, 2021	Project-based learning	Lack of collaboration ability with peers that leads to priority conflicts.	B, C, D
Pabloset al., 2020	Project-based learning	Lack of guidance during the project, at times when they did not know what they had to do.	B, D
Chin et al., 2010	Collaborative learning strategy	Identify peer misconceptions. Construct the student's own explanations.	A, B,D
Mamun et al., 2022	Inquiry-based learning	Difficulty in engaging with content, cognitive difficulty in processing the instruction or information.	C, D
Pott et al., 2021	Open-ended approach to learning	Difficult for students to be actively involved with the learning content in order to have a meaningful learning experience.	D
Rasheed et al., 2021	Blended Learning	Students' inability to properly self- regulate their peer learning activities in online mode. More specifically to prevent social loafing and refusal to participate in the online peer-learning discussion.	D
Kim et al., 2018	Problem-based learning	Students have difficulty in perception of both task difficulty and their own ability to tackle the task.	A, D

Table 1.	. Student	difficulties	in ac	hieving	SDL	outcomes	across	different	strategies
----------	-----------	--------------	-------	---------	-----	----------	--------	-----------	------------

Generally, in classroom settings, students need clear guidance on what they should be doing in a set of learning activities (Kokotsaki et al., 2016). Merely providing broad instructions like "find the learning goal", "find the resources", etc. should not suffice, if students don't know "how to find the goals" or "What resources are needed". SDL Strategies must elicit the need for learning among the students and should enable them to identify what they need and want to learn.

#### 2. SDL and Question Posing

Watkins (2017), proposed a set of meta questions that a student should ask to himself/herself to follow a student-directed learning process. These self-questions may ensure the successful achievement of SDLs before, during, and after a learning process. Question posing encourages learners to take ownership of their learning students by engaging them in metacognitive strategies (Mishra et al., 2015). When learners pose questions, they formulate their own questions, taking responsibility for their own knowledge acquisition and driving their own learning process. Question posing is a strategy that can help students think at a deep level, construct knowledge, and conduct high-order thinking (Rosenshine et al., 1996). Question posing involves new question generation based on life experience or learning situations, that can lead to new topics related to prior subject knowledge in a particular learning context.

Student questioning has been regarded as a tool to better address students' misconceptions and has been acknowledged as a type of high-level cognitive strategy that plays an important role in facilitating students' learning process (Hwang et al., 2020). Han et al. (2006), suggested that questioning has the potential to be used as an epistemic probe and heuristic tool initiating argumentation in inquiry-based science. Students' questions are the outcome of individual brainstorming and then expressing the questions, by which students become aware of both what they understood and what they did not comprehend in a topic. Therefore, students' own questions help them to identify what they need and want to learn, in the context of any given topic, and consequently can lead to a deeper understanding of the topic.

Student questions can help resolve conflicts during collaboration with peers. Students' questions serve as a cognitive tool for the students to foster critical conversation while resolving conflicts (Chin et al., 2010). Questions act as catalysts for argumentative and epistemic moves like concessions, challenges, and counter-challenges, which are expressed through explanations and justifications in the conceptions of the topic. Articulating their conflicts and beliefs with peer ideas, helps students to take meta-linguistic moves and formulate their concepts. Overall, student questions seem to be a promising pedagogical idea to help students realize various SDL processes, especially recognition of learning needs, formulation of learning goals, and self-monitoring and self-assessment of learning outcomes.

#### 3. Research Questions

We are interested in understanding how student questions can help them in overcoming the SDL challenges and achieving SDL processes. We restrict our exploration to (online and offline) classroom learning contexts where a learner has a number of peers and a facilitator in the teaching-learning environment.

#### **Broad Research Question**

Can student-question-posing-based pedagogical strategy support SDL processes? **Specific Research Questions (RQ)** 

- **RQ1:** What are the ways in which student QP has been used to develop pedagogical strategies?
- **RQ2:** What are the challenges that are addressed using QP-based pedagogical strategies?
- RQ3: What are the outcomes achieved using QP-based pedagogical strategies?
- RQ4: Which of the QP-based pedagogical strategies adequately aligns with the SDL processes?

In the next section, we present the methodology used for literature synthesis. In the later sections, we discuss our findings from the literature review against each RQ.

#### 4. Methodology

We have searched online databases with the keywords "question posing" AND "classroom" AND ("self-directed learning" OR "self-directed learning") with papers from 2015 to 2023 and received 150 papers. In the below Figure.1, we reported PRISMA techniques to select papers from the online publicly accessible databases. Prisma stands for preferred reporting items for systematic reviews and meta-analyses (PRISMA) (Page et al., 2021). It helps to identify, select, appraise, and synthesize studies from a pool of papers collected from databases. Among these 67 papers are either a thesis or article or not published papers. Among them, 4 papers are not in the English language and 31 papers are not on questioning. After the screening process, we received 48 papers. Among them, 11 papers are not based on empirical study, 10 papers are not based on class or lab or online, and 5 papers. Among them, we included 21 papers in our analysis as 5 papers do not refer to any student difficulty. All papers are read and based on the criteria followed in Table 2, we analysed the papers. Here we have considered only papers that focus on classroom, lab, and online.





#### 5. Result and Discussion

While analyzing the result, we found that most of the student difficulties in Table 1 are reported in Table 2 also. In below Table 2, we highlighted learner difficulties that are reported by the corresponding papers. In column 4, we represented how these papers used QP-based strategies. In column 3, students' difficulties addressed by these papers are represented. In column 5, SDL-related outcomes achieved by the corresponding paper are highlighted.

Reference	Demography / Context /Mode	Challenges Addressed	How student-QP is incorporated into the strategy	SDL - Related Outcome achieved
Lin et al., 2019	Graduate student. Nursing Course. Online in-class activity	The difficulty of students in in- depth understanding of the content.	ASQ (Annotation to extract key points, summarizing the whole concept and questioning in pre- class) based online flipped learning framework is used.	Achieved academic performance, self-efficacy, and critical thinking tendency as compared to the control group.
Pursitasari et al., 2020	Environmental pollution. In the class laboratory experiment.	In a guided inquiry method, students may face difficulty in getting involved in the experiment.	Students are allowed to ask questions as a part of a science context-based inquiry learning (SCOIL) model. Its open- ended problem and guided phases of observation, investigation, representation, conclusion, and communication help students to enhance critical thinking skills.	The critical thinking skills of students with the proposed model were greater than the guided inquiry learning model.
Sason et al., 2020	Ninth-grade student. Science text. In class intervention	Students face difficulty in connecting different pieces of information within the context, which may lead to inefficient and decentralized reading.	Self-generating questions are helpful in monitoring, regulating, and evaluating the learning process. This study shows that rather than self- generated questions connecting to prior knowledge (outside text), self-generating questions connecting between the text (within text) is helpful in having a long-term effect.	Achieved In- context coordination and connection, which leads to in-text comprehension achievement.
Chen et al., 2020	High school student. English Language learning course. E-learning method.	Difficulties of the student in engaging language learning.	students are encouraged to pose questions during the learning process. It enables students to	The learning performance of high- engagement participants was better than that of

Table 2: Different applications of question-posing (QP) based intervention to assist in-class student difficulties

			think and solve problems.	low-engagement participants.	
George et al., 2022	University students. Mathematics Online, asynchronous.	Difficulty in self- motivation, monitoring, and regulation during Covid-19 online class.	Students are encouraged to ask questions that can clear their doubts, and misconceptions, and promote active learning.	Students' engagement, understanding, and self- regulation were achieved.	
Looi et al., 2023	Grade 1 class student, Taiwan school.	The difficulty of disengaged and unmotivated students in the traditional examination- driven culture to learn by interest.	Proposed Learning- By-Questioning (LBQ) method. Teachers create an environment that fosters curiosity, deep exploration, and active participation. Students are encouraged to ask questions.	Showed deep interest and progression in reading. Developed reading and writing habits.	
Mishra et al., 2015	Undergraduate CS course. In- class activity.	Difficulty in generating questions relevant to the given domain.	After listening to the lecture students are asked to pose questions and assign priority to own and peer questions.	Students posed questions and unfolded their own prior and future knowledge.	

To answer RQ 1, different strategies designed using student question posing are listed in column #4 of Table 2. We found that most papers encouraged students to pose questions. For example, George et al., (2022), Sason et al., (2020), Pursitasari et al., (2020), and Chen et al., (2020) encouraged students to pose questions. Whereas Looi et al, (2023) created an environment using IDC theory, that fosters students' curiosity, and eventually students ask questions by themselves, and Mishra and Iyer (2015), allowed students to pose their questions and drive the subsequent lecture sessions based on student questions. The answers to RQ2 and RQ3 are listed in columns #3 and #5 of Table 2 respectively. We grouped student challenges from Table 2 and found that among all groups of challenges, engagement, and metacognition were mostly reported challenges. As far as RQ3 is concerned, we found that most papers applied QP as a tool to foster a better understanding of the content and improve engagement in the classroom.

In response to RQ4, we refer to Knowles' (1975) five SDL processes and analyze each of the QP-based pedagogies in this synthesized list (Table 2, Column 4) to determine which SDL processes are supported by each pedagogy. We find that Mishra and Iyer's (2015) question-posing-based strategy aligns directly with three out of five SDL processes, viz., (A) recognizing learning needs, (B) formulating learning goals, and (E) assessing (formative) learning outcomes. Other QP-based strategies synthesized in the table mostly address "(D) employing appropriate learning strategies." Mishra and Iyer (2015) name their strategy as Student Query Driven Learning (SQDL). They have employed student questions to recognize the learning needs and formulate the learning goals.



Figure 2. SQDL structure

In Figure 2, the teacher will deliver the initial lecture in Phase 1. In phases 2, 3, and 4 students are allowed to pose questions and assign tagging and priority to their questions. The questions are treated as a representation of what students need to learn. Moreover, SQDL allows students to assign priority to their peer questions. This allows students to reflect on their learning priorities, considering diverse aspects of what they need and want to learn next. The teacher uses the priorities of the questions set by the students to determine the topic and order of the next lecture/ discussion in the classroom in phases 5, 6, and 7.

#### 6. Future Work

The main motivation of self-directed learning is the learner's ability to become autonomous and the primary driver of SDL. Learners are surrounded by teacher-oriented learning in the contemporary teaching-learning situation. Additionally, teachers are not taught to apply SDL techniques. Implementing any SDL method in these situations is a challenging endeavour for both students and teachers. Here, we've examined how learners struggled to follow the SDL process facilitated by different pedagogies. The majority of the problems are engagement and metacognition-related. Each SDL process can benefit from the usage of the instrument of question posing. We discover how QP can be utilized to get over student challenges through the literature review. QP-based tactics largely succeeded in achieving engagement and content comprehension. We have also discovered that the majority of techniques concentrate on using various strategies to enhance comprehension of the information, engagement with the content, or peer interaction. Whereas Mishra and Iyer (2015), employed a QP-based technique to understand students' learning needs and future learning objectives. We will investigate the SQDL method more in the future.

#### References

- Al Mamun, M. A., Lawrie, G., & Wright, T. (2022). Exploration of learner-content interactions and learning approaches: The role of guided inquiry in the self-directed online environments. *Computers & Education*, 178, 104398.
- Almulla, M. A. (2020). The effectiveness of the project-based learning (PBL) approach as a way to engage students in learning. *Sage Open*, *10*(3), 2158244020938702.
- Angeli, C. (1998). An Instructional Model for Teaching Literacy: Implications for Instructional Theory.
- Basilotta Gómez-Pablos, V., Torrecilla Sánchez, E. M., García-Valcárcel Muñoz-Repiso, A., & Hernández Martín, A. (2020). Project-based learning in the classroom: a case study at High school level. Aula abierta, 49(4), 429-438.
- Bell, L., & Lygo-Baker, S. (2019). Student-centred learning: a small-scale study of a peer-learning experience in undergraduate translation classes. *The Language Learning Journal*, 47(3), 299-312.

- Biggs, J. B. (1989). Approaches to the enhancement of tertiary teaching. *Higher education research* and development, 8(1), 7-25.
- Brandt, W. C. (2020). Measuring Student Success Skills: A Review of the Literature on Self-Directed Learning. 21st Century Success Skills. *National Center for the Improvement of Educational Assessment*.
- Brooks-Gunn, J., Guo, G., & Furstenberg Jr, F. F. (1993). Who drops out of and who continues beyond high school? A 20-year follow-up of black urban youth. *Journal of Research on Adolescence*, *3*(3), 271-294.
- Chen, A. M. R., Majumdar, R., Hwang, G. J., Lin, Y. H., Akcapinarapinar, G., Flanagan, B., & Ogata, H. (2020, November). Improving EFL students' learning achievements and behaviors using a learning analytics-based e-book system. In 28th International Conference on Computers in Education Conference Proceedings (Vol. 1, pp. 474-483). Asia-Pacific Society for Computers in Education (APSCE).
- Chin, C., & Osborne, J. (2010). Students' questions and discursive interaction: Their impact on argumentation during collaborative group discussions in science. *Journal of research in Science Teaching*, *47*(7), 883-908.
- Cintang, N., Setyowati, D. L., & Handayani, S. S. D. (2018). The obstacles and strategy of project based learning implementation in elementary school. *Journal of Education and Learning (EduLearn)*, *12*(1), 7-15.
- Miller, E. C., & Krajcik, J. S. (2019). Promoting deep learning through project-based learning: A design problem. *Disciplinary and Interdisciplinary Science Education Research*, *1*(1), 1-10.
- Mulyawati, D., & Tarsidi, I. (2020, February). Self-Regulated Learning Constructions for First Middle School Students. In International Conference on Educational Psychology and Pedagogy-" Diversity in Education" (ICEPP 2019) (pp. 40-42). Atlantis Press.
- George, W., & Gallagher, K. (2022). What in the world just happened? Students' self-regulation in asynchronous PreCalculus during the COVID-19 pandemic. *Journal of Humanistic Mathematics*, *12*(2), 315-345.
- Gultom, S., Hutauruk, A. F., & Ginting, A. M. (2020). Teaching skills of teacher in increasing student learning interest. *Budapest International Research and Critics Institute (BIRCI-Journal): Humanities and Social Sciences*, *3*(3), 1564-1569.
- Han, K. S., Song, Y. I., Kim, S. B., & Rim, H. C. (2006). A definitional question answering system based on phrase extraction using syntactic patterns. *IEICE transactions on information and* systems, 89(4), 1601-1605.
- Hussein, B. (2021). Addressing collaboration challenges in project-based learning: The student's perspective. *Education Sciences*, *11*(8), 434.
- Hwang, G. J., Zou, D., & Lin, J. (2020). Effects of a multi-level concept mapping-based questionposing approach on students' ubiquitous learning performance and perceptions. *Computers & Education*, 149, 103815.
- Lin, H. C., Hwang, G. J., & Hsu, Y. D. (2019). Effects of ASQ-based flipped learning on nurse practitioner learners' nursing skills, learning achievement and learning perceptions. *Computers* & *Education*, 139, 207-221.
- Lorås, M., Sindre, G., Trætteberg, H., & Aalberg, T. (2021). Study behavior in computing education a systematic literature review. *ACM Transactions on Computing Education (TOCE)*, 22(1), 1-40.
- Looi, C. K., Wong, S. L., Kong, S. C., Chan, T. W., Shih, J. L., Chang, B., ... & Liao, C. C. (2023). Interest-driven creator theory: Case study of embodiment in an experimental school in Taiwan. *Research and Practice in Technology Enhanced Learning*, *18*, 023-023.
- Kim, L., & Zitzer, M. (1999). Powerful Learning Framework for Teachers.
- Knowles, M. S. (1975). Self-directed learning: A guide for learners and teachers.
- Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: A review of the literature. *Improving schools*, *19*(3), 267-277.
- Martinetti, A. (2020). Optimizing student-driven learning (SdL) through a framework designed for tailoring personal student paths. *Education Sciences*, *10*(9), 249.
- Mishra, S., & Iyer, S. (2015). An exploration of problem posing-based activities as an assessment tool and as an instructional strategy. *Research and practice in technology enhanced learning*, *10*(1), 1-19.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International journal of surgery*, 88, 105906.
- Pott, R. W., & Nortjé, S. (2021). The use of student question-posing in reactor design to encourage an open-ended approach to learning. *European Journal of Engineering Education*, 46(6), 874-888.

- Pursitasari, I. D., Suhardi, E., Putra, A. P., & Rachman, I. (2020). Enhancement of student' s critical thinking skill through science context-based inquiry learning. *Jurnal Pendidikan IPA Indonesia*, *9*(1), 97-105.
- Rasheed, R. A., Kamsin, A., & Abdullah, N. A. (2021). An approach for scaffolding students peerlearning self-regulation strategy in the online component of blended learning. *IEEE Access*, 9, 30721-30738.
- Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of educational research*, 66(2), 181-221.
- Sason, H., Michalsky, T., & Mevarech, Z. (2020). Promoting middle school students' science text comprehension via two self-generated "linking" questioning methods. *Frontiers in Psychology*, *11*, 595745.
- Sasson, I., Yehuda, I., & Malkinson, N. (2018). Fostering the skills of critical thinking and questionposing in a project-based learning environment. *Thinking Skills and Creativity*, *29*, 203-212.
- Shpeizer, R. (2019). Towards a successful integration of project-based learning in higher education: Challenges, technologies and methods of implementation. *Universal Journal of Educational Research*, 7(8), 1765-1771.
- Steeb, D. R., Zeeman, J. M., Bush, A. A., Dascanio, S. A., & Persky, A. M. (2021). Exploring career development through a student-directed practicum to provide individualized learning experiences. *Currents in Pharmacy Teaching and Learning*, 13(5), 500-505.
- Vyas, A., Leung, C. W., & Wong, W. O. Developing Student Driven Learning: Impact on Knowledge and Attitude.
- Watkins, C. (2017). Developing student-driven learning: the patterns, the context, and the process. In *Student-driven learning strategies for the 21st century classroom* (pp. 1-9). IGI Global.
- Wu, T. T., & Wu, Y. T. (2020). Applying project-based learning and SCAMPER teaching strategies in engineering education to explore the influence of creativity on cognition, personal motivation, and personality traits. *Thinking Skills and Creativity*, 35, 100631.

# W07

## The 3rd Workshop on Innovative technologies for enhancing interactions and learning motivation

## The Role of VR-Supported Co-Learning Environments in University Students' Self-Efficacy and Learning Engagement

Chiung-Fang CHANG<sup>a1</sup>, & Jerry Chih-Yuan SUN<sup>a\*</sup>

<sup>a</sup>Institute of Education, National Yang Ming Chiao Tung University, Taiwan \*jerrysun@nycu.edu.tw

(Note 1: The authors' names should follow the format of First Name LAST NAME (e.g., Tak-Wai CHAN) – i.e., first (given) name before last (family) name, and the last name be fully capitalized.)

(Note 2: The submitted manuscript need not be an onymized anymore, as the single blind review approach will be practiced, i.e., authors do not know who the referees are, while the referees know who the authors are. Authors may provide their names, affiliations and emails in the manuscript submitted for review.)

(Note 3: Please do not change the above header in the first page.)

Abstract: The purpose of this study was to investigate the role and impact of technology-assisted co-learning environments on student learning. While virtual reality technology can create realistic learning situations in limited space, peer-supported regulation tasks can promote learning engagement and self-regulated learning. Due to abstract texts often being used in university textbooks and an increase in self-paced distance learning, keeping focused on reading might lead to learning burnout and anxiety. This mode of learning is too monolithic and relies on a high degree of selfdiscipline while lacking interactivity and affective support. Therefore, this study examined the potential outcomes of VR-supported co-learning environments, including self-efficacy and three types of learning engagement. A total of 46 undergraduate and graduate students were invited to participate in a learning activity where the control group was asked to read the material about collaborative learning concepts on a PDF file, while the VR group was asked to practice regulation skills with a peer via a 2D VR platform, Gather.town, before and after reading the assigned material. Data were collected using a pretest-posttest design and included demographics and self-reports on self-efficacy and learning engagement. Results showed that there was no group difference concerning self-efficacy via the ANCOVA test. However, the VR group was found to report greater affective engagement than the control group with the selfefficacy pre-test held constant. Implications concerning practical contributions to instructional design and the integration of virtual technology into learning were discussed. Future studies are encouraged to create rich and diverse learning opportunities to enhance student's learning self-efficacy and engagement.

Keywords: virtual reality, collaborative learning, learning engagement, peer regulation

#### 1. Introduction

The shift to online and distance learning has gained significance in recent years, especially due to the global pandemic. This has led to the need for effective and engaging virtual learning environments that promote student learning and achievement. Students often struggle to stay focused and engaged when dealing with abstract or resource-demanding texts found in university textbooks, corresponding to subsequent burnout and stress (e.g., Rai et al. 2015). Regulating learning processes, including metacognitive skills like time management (Walczyk et al., 1999) and strategy use (Rai et al., 2011), is crucial in such situations. While research has focused on technology-supported personalized learning (Xie et al., 2019), there is still limited understanding of the effectiveness of social-enabled novel-

tech environments that combine instructional strategies and technological tools to achieve teaching goals (Parmaxi, 2020).

Technology has increasingly been used to create virtual learning environments that mimic real-world scenarios, providing interactive and collaborative experiences for students (Wang & Sun, 2021). VR-supported co-learning environments offer immersive and interactive learning experiences, while peer-supported regulation tasks encourage students to take ownership of their learning and develop self-regulated learning skills. By working together in an immersive environment, students can learn from each other, share ideas, and provide support, resulting in a more engaging and regulated learning state.

In addition, given the modern learning paradigm emphasizing sociocultural learning (Jonassen et al., 1995; Murphy et al., 2012), it is worth exploring the potential benefits of incorporating co-learning and co-regulated tasks in a virtual reality (VR) setting. While previous studies have investigated the impact of co-learning and co-regulation on student learning (Hadwin et al., 2018; Quackenbush & Bol, 2020), few have examined the effectiveness of using VR technology to enhance these collaborative learning strategies (e.g., Wang & Sun, 2021, 2022). The use of VR technology in co-learning environments provides immersive peer-interaction experiences that promote active student engagement and foster deeper levels of collaboration between peers. Therefore, this study aimed to investigate the impact of technology-assisted co-learning environments on student learning, particularly to address needs in the context of distance learning. The use of virtual reality platforms such as Gather.town (https://www.gather.town) provides a promising avenue for creating interactive learning experiences that can engage students and promote learning. This study sought to understand the potential outcomes of VR-supported co-learning environments, including self-efficacy and learning engagement, which are critical factors for student success in online and distance learning contexts. The research framework is shown in Figure 1 with the research question outlined as: Are there significant differences in students' self-efficacy and learning engagement between traditional learning and VRsupported co-learning environments?



Figure 1. Research Framework

#### 2. Literature Review

#### 2.1 The Social Constructive Dimension for Online/Distance Learning

Distance education is often challenged by isolation and lack of social interaction. Social interaction is essential for human learning as it facilitates the construction of knowledge through dialogue, feedback, and collaboration (Vygotsky & Cole, 1978). Several studies have investigated the importance of social interaction in online and distance learning. For

example, a study by Weidlich and Bastiaens (2017) found that learners who perceived a sociable environment experienced greater social interaction that, in turn, explained their satisfaction, confidence, and perceived learning in online courses. Furthermore, interaction among peers using a cloud-based technology application for well-structured collaborative tasks enhanced academic writing scores for those with underdeveloped language skills (Li & Mak, 2022).

The social constructive dimension is especially important for online and distance learning because it enables learners to engage in dialogue and collaboration, which facilitates knowledge construction. As stated by Vygotsky's theory of social constructivism, learning occurs when individuals interact with their environment and with each other (Vygotsky, 1978). Given richer and situated learning experience which novel technology affords, such as VR, learners could benefit from the optimization of sociability, social space, and social presence of the virtual environment to work collaboratively.

#### 2.2 VR-Supported Co-Learning and Co-Regulation

In line with the educational purpose of learning facilitation by technology, virtual reality technology has been increasingly used in peer-learning environments which afford more opportunities for co-regulation and awareness of set goals. Immersive peer-learning environments have been found to promote learning motivation (Wang & Sun, 2022) and emotional engagement (Wang & Sun, 2021). Sun et al. (2023) also found that using wearable hybrid AR/VR materials helps create positive effects on students' situated interest, engagement, and learning performance among high school learners, with similar findings for elementary students using VR-based wearable devices (Hung et al., 2023). However, scarce findings are available concerning various forms of VR tools (i.e., Gather.town) suiting colearning tasks with the quality determined by individual and group effort to regulate behaviors toward goal attainment.

The self-regulated learning (SRL) model proposed by Zimmerman (2000) is a cyclical process that involves three phases: forethought, performance or volitional control, and self-reflection. In the forethought phase, learners set goals, activate prior knowledge, and develop plans. During the performance phase, learners use strategies to implement their plans and monitor their progress. Finally, in the self-reflection phase, learners evaluate their performance, attribute their successes or failures, and modify their plans for future learning. The structure of regulated-learning practice usually remains the same when more than one person is involved in the process, but the task ownership might depend on the level of individuals' collaborative beliefs and task setting.

Co-regulated and socially-shared regulated learning complement self-regulated learning theory, as they add a new dimension to the regulation of learning influenced by others (Panadero & Järvelä, 2015). Co-regulated learning and socially-shared regulated learning are two related but distinct concepts in the context of learning. Co-regulated learning involves a relationship between individuals where one person is more knowledgeable or skilled, and both parties self-regulate and share regulation of behavior, emotions, motivation, and cognition for personal goal attainment (Hadwin et al., 2018). Socially-shared regulated learning on the other hand, refers to the joint effort, where the group takes metacognitive control of the task together through negotiation and an iterative process of behavioral, emotional, motivational, and cognitive exchange (Hadwin et al., 2018). As such, regulation of learning in this given context inclines toward the direction addressed by the whole group and away from the members within the group (Järvelä & Hadwin, 2013; Panadero & Järvelä, 2015).

The co-learning environment pinpointed in the present study adopts the notion more related to socially-shared learning, as prompts are provided to assign the leading role of collaborative tasks, involving reading out the instruction and navigating the process of

responding to questions. Therefore, the present study builds on prior research by examining the impact of VR-supported co-learning environments on self-efficacy and learning engagement. While prior studies have investigated the impact of collaborative learning on student learning, few have examined the role of VR technology in promoting positive affective states and increasing student engagement. By exploring the relationship between VR-supported co-learning environments and student self-efficacy and engagement, the present study provides new insights into the potential benefits of using immersive technologies to support online education and distance learning.

#### 3. Methods

#### 3.1 Participants

The present study conducted in northern Taiwan utilized a quasi-experimental design and the convenience sampling method to gather information from 46 university students. Both males and females were equally represented in the study (50%) with an average age of 23.90 years (SD = 2.57). The majority of the students were pursuing majors related to STEM fields, followed by social sciences. The participants' educational levels comprised undergraduate (n = 11) and graduate (n = 35) programs.

#### 3.2 Experimental procedure

Figure 2 illustrates the experimental design, where the participants were randomly divided into two groups. Before the experiment, the students were informed about the research and gave their consent. They also responded to the questionnaire consisting of demographics and a self-efficacy self-report. After that, the control group was asked to read PDF material on the computer, while the VR group participated in co-learning tasks on a 2D map situated on the Gather.town conferencing platform (https://www.gather.town), before and after reading the assigned material. The reading time lasted 15 minutes, with an additional 13 minutes for the VR group to complete their co-learning tasks. Afterwards, the participants were asked to complete a semi-structured questionnaire, which took about 6 minutes.



Figure 2. Experimental Procedure

#### 3.3 Instruments

#### 3.3.1 Self-Report Questionnaires

Self-efficacy and learning engagement scales were used in our self-report questionnaires. The *self-efficacy scale* aimed to assess students' confidence in comprehending the reading topic with the version in this study adapted from (Sun et al., 2015). There were eight items on a 6-point Likert scale, with sufficient reliability ( $\alpha$  = .95; Nunnally, 1967). In addition, the *learning engagement scale* was translated from Whitney et al. (2019) with expert review, and assessed to what extent students revealed affective, behavioral, and cognitive engagement. The affective sub-dimension consisted of eight items, the behavioral sub-dimension comprised nine items, and the cognitive sub-dimension had seven items on a 5-point Likert scale. The Cronbach's  $\alpha$  of each subscale was .89, .78, .80, respectively.

#### 3.3.2 Co-Learning Tasks on the Gather.town Platform

Three co-learning tasks were conducted on the Gather.town conferencing website, which is a virtual platform that provides a customizable and interactive space (see Figure 3) for students to collaborate and learn together. The tasks included evaluation of personal learning strategies, goal setting, and learning reflection, and each task was completed in pairs, with two students working collaboratively. In the evaluation of personal learning strategies task, each pair discussed their learning strategies and identified ways to improve them. In the goal-setting task, the pairs shared their personal goals for their time-limited reading task. Finally, during the learning reflection task, the pairs reflected on their learning progress and identified their goal achievements for the reading and areas for further improvement. The co-learning approach encouraged collaboration between students, and provided an opportunity for them to clarify learning goals with collective effort in a virtual environment. The Gather.town conferencing website facilitated the co-learning process by providing a platform for students to connect and work together, using a variety of customizable features such as virtual avatars, 2D pixel maps, and interactive elements.



Figure 3. Interface on Gather.town

#### 4. Results and Discussion

Several ANCOVA tests were employed to examine between-group differences concerning the outcomes of the self-efficacy ratings and learning engagement ratings (see Table 1). All the analyses were conducted using the IBM SPSS Statistics 24 software. Students' self-

efficacy ratings were found to be equivalent across the two groups after controlling for their prior self-efficacy. As for students' ratings of learning engagement by the control of prior self-efficacy, only the affective engagement sub-dimension showed that the VR group reported higher ratings than the control group. No group differences were observed for the behavioral sub-dimension or cognitive sub-dimension.

Results of the ANCOVA tests for self-efficacy and learning engagement							
	n = 26 Control group	<i>n</i> = 20 VR group	F	р	partial η²		
Outcomes	М (SD) <u>М<sub>Аdj</sub> (SE)</u>	М (SD) <u>М<sub>Аdj</sub> (SE)</u>					
Self-efficacy <sup>a</sup>	4.50 (0.69) <u>4.48 (0.12)</u>	4.26 (0.68) <u>4.29 (0.14)</u>	1.09	.303	.025		
Learning engagement <sup>a</sup>							
Affective	3.59 (0.75) <u>3.57 (0.12)</u>	3.94 (0.48) 3.96 (0.14)	4.49	.040*	.095		
Behavioral	3.49 (0.50) 3.49 (0.10)	3.67 (0.56) 3.68 (0.12)	1.47	.232	.033		
Cognitive	3.69 (0.53) <u>3.68 (0.10)</u>	3.58 (0.47) <u>3.86 (0.11)</u>	1.50	.228	.034		

Table 1

*Note.* \* p < .05, <sup>a</sup> analysis with the covariate of prior self-efficacy

Despite greater enjoyment and desire of learning collaborative learning concepts found in the VR group as corresponded to previous studies (Slater, 2018), there were a few speculations for the minimal benefit of VR-supported co-learning environments. First, it is difficult to change students' learning styles and behavior with a one-time effort (Wang & Sun, 2021), although students were asked to be more aware of using effective learning strategies and clarifying their goals before their tasks. As such, students might still use maladaptive learning strategies that confound with behavioral or cognitive engagement. In addition, experienced affect in the VR-supported co-learning environment as one of the four sources (i.e., affect, past performance, vicarious learning, and social persuasion) of self-efficacy may not be enough to enhance subsequent self-efficacy (Bandura, 1977).

#### 5. Conclusion

This study aimed to investigate the impact of technology-assisted co-learning environments on student learning. The results showed that only the affective engagement sub-dimension of learning engagement ratings showed a significant difference, with the VR group reporting higher ratings than the control group. There were no significant differences in self-efficacy ratings between the two groups after controlling for prior self-efficacy. The limitations of this study include the possibility that students may still use maladaptive learning strategies that confound with cognitive and behavioral engagement, and the potential inadequacy of the experienced affect in the VR-supported co-learning environment as a source of self-efficacy. Moreover, there may be a lack of statistical power given the small sample size to assess VRsupported co-regulation effects. Future studies should consider a longer timeframe and explore richer and more diverse learning opportunities to enhance student learning in technology-assisted co-learning environments. Nonetheless, this study provides practical implications for instructional design and the integration of virtual technology into learning environments.

#### Acknowledgements

This research was supported by the National Science and Technology Council (formerly the Ministry of Science and Technology) in Taiwan through Grant numbers MOST 111-2410-H-A49-018-MY4 and MOST 110-2511-H-A49-009-MY2. We would like to thank NYCU's ILTM (Interactive Learning Technology and Motivation, see: <u>http://ILTM.lab.nycu.edu.tw</u>) lab members and the students for helping conduct the experiment. We would also like to thank all the students for participating in the study, Shih-Jou Yu, and the reviewers who provided valuable comments.

#### References

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215. <u>https://doi.org/10.1037/0033-295X.84.2.191</u>
- Hadwin, A., Järvelä, S., & Miller, M. (2018). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In *Handbook of self-regulation of learning and performance, 2nd ed.* (pp. 83-106). Routledge/Taylor & Francis Group. <u>https://doi.org/10.4324/9781315697048-6</u>
- Hung, C.-Y., Lin, Y.-T., Yu, S.-J., & Sun, J. C.-Y. (2023). Effects of AR- and VR-based wearables in teaching English: The application of an ARCS model-based learning design to improve elementary school students' learning motivation and performance. *Journal of Computer Assisted Learning*. <u>https://doi.org/10.1111/jcal.12814</u>
- Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist*, 48(1), 25-39. https://doi.org/10.1080/00461520.2012.748006
- Jonassen, D., Davidson, M., Collins, M., Campbell, J., & Haag, B. B. (1995). Constructivism and computer-mediated communication in distance education. *American journal of distance education*, *9*(2), 7-26. <u>https://doi.org/10.1080/08923649509526885</u>
- Li, J., & Mak, L. (2022). The effects of using an online collaboration tool on college students' learning of academic writing skills. *System*, *105*. <u>https://doi.org/10.1016/j.system.2021.102712</u>
- Murphy, P. K., Alexander, P. A., & Muis, K. R. (2012). Knowledge and knowing: The journey from philosophy and psychology to human learning. In S. G. K. R. Harris, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Ed.), *APA educational psychology handbook* (Vol. 1, pp. 189-226). American Psychological Association. https://doi.org/https://doi.org/10.1037/13273-008
- Nunnally, J. C. (1967). Psychometric theory. McGraw-Hill.
- Panadero, E., & Järvelä, S. (2015). Socially shared regulation of learning: A review. *European Psychologist*, 20(3), 190-203. <u>https://doi.org/10.1027/1016-9040/a000226</u>
- Parmaxi, A. (2020). Virtual reality in language learning: a systematic review and implications for research and practice. *Interactive Learning Environments*, *31*(1), 172-184. https://doi.org/10.1080/10494820.2020.1765392
- Quackenbush, M., & Bol, L. (2020). Teacher support of co-and socially-shared regulation of learning in middle school mathematics classrooms [Brief Research Report]. *Frontiers in Education*, 5. <u>https://doi.org/10.3389/feduc.2020.580543</u>
- Rai, M. K., Loschky, L. C., & Harris, R. J. (2015). The effects of stress on reading: A comparison of first-language versus intermediate second-language reading comprehension. *Journal of Educational Psychology*, 107(2), 348-363. <u>https://doi.org/10.1037/a0037591</u>
- Rai, M. K., Loschky, L. C., Harris, R. J., Peck, N. R., & Cook, L. G. (2011). Effects of stress and working memory capacity on foreign language readers' inferential processing during comprehension. *Language Learning*, 61(1), 187-218. <u>https://doi.org/10.1111/j.1467-9922.2010.00592.x</u>
- Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *Br J Psychol*, 109(3), 431-433. <u>https://doi.org/10.1111/bjop.12305</u>
- Sun, J. C.-Y., Chang, K.-Y., & Chen, Y.-H. (2015). GPS sensor-based mobile learning for English: An exploratory study on self-efficacy, self-regulation and student achievement. *Research and Practice in Technology Enhanced Learning*, *10*(1), 1-18. <u>https://doi.org/10.1186/s41039-015-0024-y</u>
- Sun, J. C.-Y., Ye, S.-L., Yu, S.-J., & Chiu, T. K. F. (2023). Effects of wearable hybrid AR/VR learning material on high school students' situational interest, engagement, and learning performance: The case of a physics laboratory learning environment. *Journal of Science Education and Technology*, 32, 1-12. <u>https://doi.org/10.1007/s10956-022-10001-4</u>

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

- Walczyk, J. J., Kelly, K. E., Meche, S. D., & Braud, H. (1999). Time limitations enhance reading comprehension. *Contemporary Educational Psychology*, 24(2), 156-165. <u>https://doi.org/10.1006/ceps.1998.0992</u>
- Wang, H.-Y., & Sun, J. C.-Y. (2021). Real-time virtual reality co-creation: Collective intelligence and consciousness for student engagement and focused attention within online communities. *Interactive Learning Environments*. <u>https://doi.org/10.1080/10494820.2021.1928711</u>
- Wang, H.-Y., & Sun, J. C.-Y. (2022). Influences of online synchronous VR co-creation on behavioral patterns and motivation in knowledge co-construction. *Educational Technology & Society*, 25(2), 31-47. <u>https://www.jstor.org/stable/48660122</u>
- Weidlich, J., & Bastiaens, T. J. (2017). Explaining social presence and the quality of online learning with the SIPS model. *Computers in Human Behavior*, 72, 479-487. https://doi.org/10.1016/j.chb.2017.03.016
- Whitney, B. M., Cheng, Y., Brodersen, A. S., & Hong, M. R. (2019). The scale of student engagement in statistics: Development and initial validation. *Journal of Psychoeducational Assessment*, 37(5), 553-565. <u>https://doi.org/10.1177/0734282918769983</u>
- Xie, H., Chu, H.-C., Hwang, G.-J., & Wang, C.-C. (2019). Trends and development in technologyenhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education*, *140*, 103599. https://doi.org/https://doi.org/10.1016/j.compedu.2019.103599
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). Academic Press. https://doi.org/https://doi.org/10.1016/B978-012109890-2/50031-7

## Designing and Evaluating an Attention-Engagement-Error-Reflection (AEER) Approach to Enhance Primary School Students Artificial Intelligence Literacy and Learning-to-Learn Skills: A Pilot Study

#### Siu-Cheung KONG<sup>a\*</sup>, Yin YANG<sup>a</sup>

<sup>a</sup> Artificial Intelligence and Digital Competency Education Centre, The Education University of Hong Kong, Hong Kong \*sckong@eduhk.hk

**Abstract:** This paper reports on a mixed-method pilot study that designed and evaluated an Attention-Engagement-Error-Reflection (AEER) approach to enhance primary school students' artificial intelligence (AI) literacy and learning-to-learn skills. A total of 35 Grade six primary school students in Hong Kong were involved. Both quantitative and qualitative data were collected and analyzed. Data collection included (1) pre-, post-, and delayed post-AI concepts tests, (2) pre- and post-questionnaires on learning-to-learn skills, and (3) six focus group interviews. The results demonstrated the positive impact of the proposed AEER approach on students' AI conceptual understanding and learning-to-learn skills. Students' motivation in the learning process was analyzed. Finally, limitations and future directions were discussed.

**Keywords:** Innovative pedagogical design, artificial intelligence literacy, learning-tolearn skills, motivation

#### 1. Introduction

With the rapid advancement of technology, the importance of artificial intelligence (AI) literary education has emerged as a global topic (Dwivedi et al., 2021). The goal of AI literacy education is to equip individuals with the fundamental AI concepts, empower students to participate effectively in a digital society, raise awareness of the ethical use of AI, and enable them to collaborate with AI systems in their daily lives (Kong et al., 2023; Su et al, 2023). However, the potential of AI literacy education to foster learning-to-learn skills in K-12 educational contexts remains largely unexplored (Azevedo & Wiedbusch, 2023).

To address this gap, the present pilot study was designed to develop and evaluate an innovative pedagogical approach - the Attention-Engagement-Error-Reflection (AEER) approach, implemented using the AlphAI learning robots. This approach aimed to enhance primary students' AI literacy and learning-to-learn skills. Additionally, the study explored students' motivation during the learning process. The following three questions were addressed:

- (1) What is the impact of an AEER approach on students' AI conceptual understanding?
- (2) What is the impact of the AEER approach on students' learning-to-learn skills?
- (3) How does the AEER approach support learners' motivation?

#### 2. Literature Review

#### 2.1 Using Robots for Developing Artificial Intelligence Literacy

Educational robotics applications are increasingly being utilized to enhance students' learning experiences and stimulate their interests in the areas of STEM (science, technology, engineering, and mathematics and AI (Su et al.,2023). Despite the great potential of robots in engaging learners (Zhong & Xia, 2020), most of these applications require an understanding of algorithms and programming skills, posing an accessibility challenge for beginners, particularly in primary education contexts (Noh & Lee, 2020). Empirical research has suggested that the use of educational robotics applications could positively influence academic performance (AI Hakim et al., 2022). Despite the growing interest in AI literacy education, rare studies have been explored in the literature regarding the effective integration of learning robots through innovative pedagogical design.

#### 2.2 Using Robots for Developing Learning-to-Learn Skills

Learning-to-learn skills refer to individuals' ability to monitor, regulate and control their own learning activities (Cornford, 2002) to learn effectively in various contexts. The learning-to-learn skills are not only vital for academic success but also facilitate students to become lifelong learners, helping students become more adaptable and resilient in the face of new challenges and changing learning settings (Thrun & Pratt, 2012).

In this pilot study, learning-to-learn skills were adapted from Flavell's (1979) and Schraw and Moshman's (1995) model of metacognition, including metacognitive knowledge (i.e., individual's views as a learner and beliefs about their ability to learn) and metacognitive regulation (i.e., setting goals, monitoring, and reflecting). Despite the need for nurturing learning-to-learn skills in primary education (e.g., Vainikainen et al., 2015), few studies have explored using robots in AI literacy education to engage young students in an interactive way.

#### 2.3 Understanding Students' Motivation in AI Literacy Education

A number of studies have showed the vital role of motivation in AI literacy education (e.g., Kong et al., 2021). One commonly used motivational design model is the attention, relevance, confidence, and satisfaction (ARCS) model proposed by Keller (1987). Attention focuses on capturing students' curiosity. Relevance involves relating learning to students' goals and interests. Confidence aims to build students' self-efficacy of competence. Satisfaction enables students to apply their new learned knowledge. The ARCS model provides useful guidance for understanding and enhancing student motivation. However, few studies have examined primary school students' motivation using the ARCS model in the context of AI literacy education.

#### 3. Research Design

#### 3.1 The AlphAl Learning Robots

The AlphAl learning robots (<u>https://learningrobots.ai/?lang=en</u>) used in this study aimed to provide learners with hands-on experience and graphical interfaces to visualize Al algorithms in machine learning and deep learning (Martin et. al., 2023). Each robot, as shown in Figure 1(a), was connected to the AlphAl software, illustrated in Figure 1(b). The AlphAl software helps visualize the workings of neural networks and how machines learn and allow students to take control of the robots by using the arrows on the computer keyboard.



neural network on screen (i.e. supervised learning)

Figure 1. The AlphAl Learning Robots.

#### 3.2 Participants

In this pilot study, a convenience sampling approach was adopted (Etikan et al., 2016) due to the close communication of researchers and the teachers from the selected schools. Thirty-five Grade six primary school students (22 boys and 13 girls) with average age of 11.54 in Hong Kong were involved. Regarding the research ethics associated with data collection from the participants, written informed consent was obtained from both the students and their parents prior to commencement of the study.

#### 3.3 Pedagogical Design: Attention-Engagement-Error-Reflection (AEER) Approach

Figure 2 shows the proposed pedagogical design. The AEER approach was adapted from Dahaene's four pillars of learning (2020), including attention, active engagement, error feedback, and consolidation. To be specific, the first component in the AEER approach is to harness learners' attention. This was achieved via the use of the AlphAI learning robots. The use of the AlphAI learning robots in an educational setting could spark curiosity and interest among students, thereby effectively focusing their attention on the learning task.

The next step was to promote active engagement through collaborative learning. In this study, learners were organised into small teams, each consisting of 3-4 students, to facilitate collaborative learning. A "play-learn-play" strategy, emphasising 'learning by doing', was implemented. The learning journey commenced with students interacting with the AlphAl learning robots. This was followed by the delivery of conceptual knowledge, seamlessly integrated with hands-on activities, enabling students to apply the theoretical constructs practically and instantaneously. The culmination of this learning experience was a robot racing competition, providing an opportunity to manifest their newly acquired knowledge and skills in a competitive yet enjoyable setting. This cyclical process of play, learn, and play again aimed to help students remain actively engaged, thereby enhancing their motivation, understanding and retention of the Al concepts, and learning to learn from robots.

The third component in AAER approach is error, encouraging students to learn from their trial-and-error experiences with the AlphAl learning robots. In each robot racing competition, the robots would find themselves in situations not encountered during the initial training (for example, getting stuck on an edge of the arena, or on another robot). If the robot could not move, students were encouraged to observe the mistakes made by the robots and took control of the robot with the on-screen arrows or the keyboard under the guidance of a tutor in each group. This process aimed to foster a growth mindset in students, as they began to perceive errors as learning opportunities rather than failures.

The fourth component is reflection. By reflecting on students' learning process, they can gain a deeper understanding of strengths and weaknesses, identify their learning strategies, and develop plans for improvement.



Figure 2. The Proposed AEER Approach.

#### 3.4 Research Procedure

The pilot study lasted for two weeks, including a two 3-hour workshops. The concepts of machine learning, deep learning, supervised learning, and reinforcement learning were covered. In the first workshop, students were introduced to basic concepts of AI, supervised learning and reinforcement learning by training the robots. In the second workshop, students applied the knowledge learned in the first workshop and reflected on their learning to train the robots to avoid obstacles in a new environment.

Before the first workshop, the students completed a pre-test of AI concepts and a pre-survey on learning-to-learn. After the second workshop, the students completed a post-test of AI concepts and a post-survey. A total of 29 students were invited to participate in the interviews on a voluntary basis. After one week, participants were invited to complete the delayed post-tests to measure the retention of AI concepts.

#### 3.5 Data Collection and Analysis

In this study, data sources included (1) pre-, post-, and delayed post-AI concepts tests, (2) pre- and post-questionnaires on learning-to-learn, and (3) six focus group interviews, involving a total of 29 students. The instruments used in this study were jointly developed and validated through the collaborative efforts of two experts in AI literacy and two researchers specializing in metacognition. The AI concept test was used to assess students' AI conceptual understanding in machine learning and deep learning. The test was developed by the research team, including seven items. The questionnaire on learning-to-learn skills included four dimensions of goal setting, monitoring, reflection, and growth mindset with 11 items. The questions of focus group interviews were guided by ARCS motivation model to obtain more in-depth and comprehensive views regarding using AEER approach with the AlphAI learning robots for AI literacy education and learning-to-learn.

Quantitative data analysis was used to address the two research questions. For the first research question, a one-way repeated measures ANOVA was conducted. For the second research question, paired sample-t-tests were used. For the third research question, content analysis was used to analyze focus group interviews to understand primary students' learning motivation in using AlphAl learning robots via AEER approach.

#### 4. Results

#### 4.1 AI Conceptual Understanding

A one-way repeated measures ANOVA was conducted to determine whether there were statistically significant differences in the scores of conceptual understating of AI in pre-tests, post-tests, and delayed post-tests. The scores of AI tests were statistically significantly different at the different time points, F (2, 68) = 43.23, p < .001, partial  $\eta^2 = .56$ . Post hoc analysis with a Bonferroni adjustment revealed that the scores were statistically significantly significantly increased from pre-tests to post-tests (M<sub>diff</sub> = 2.34, 95% CI [1.66, 3.03], p < .001), and from pre-tests to delayed post-tests after one week (M<sub>diff</sub> = 2.11, 95% CI [1.37, 2.86], p < .001). Although there was a slight decrease from delayed post-tests to post-tests (M<sub>diff</sub> = -.23, 95% CI 1.65, 3.03], p > .05), no significant difference was observed.

#### 4.2 Leaning-to-Learn Skills

Paired sample-t-tests were conducted to determine whether there were statistically significant differences in the scores in the pre-and post-surveys. The results showed that students learning-to-learn skills improved in terms of four dimensions: goal setting, monitoring, reflection, and growth mindset. Overall, there was a statistically significant increase in students' learning-to-learn skills from the pre-survey to the post-survey ( $M_{diff} = 0.11, 95\%$  CI [0.03,0.04], p < .05).

#### 4.3 Interview Results

Beyond the significant improvement in the test and survey, the focus group interviews indicated that using the AlphAl learning robots via the AEER approach improved students' learning motivation. The focus group interviews showed that the majority of interviewees agreed that the AlphAl learning robots made Al learning interesting and engaging, thereby enhancing their learning motivation.

Regarding "attention", all of the students felt that the AlphAl learning robots aroused their attention. The interactive nature of the robots, which could be controlled through keyboard inputs and provided visual feedback, sparked students' curiosity. Some interviewees (pseudonyms were used) further explained:

- "Controlling the robots was like a game, and I was also learning at the same time." (Charles)
- "Using the robots to learn AI was novel. I have never experienced it before." (Cindy)

In terms of "relevance", the interviews indicated that controlling robots was relevant to their learning. For example, one student said, "I feel that supervised learning is similar to us being supervised by teachers or parents. They accompany us as we study. On the other hand, reinforcement learning requires time to learn. That is, if we rush to finish learning something and then cram for tests, we will not be able to achieve good results because we are not yet familiar with the materials".

Regarding "confidence", the majority of students stated that the learning led to increased confidence in their ability to learn and understand AI. Claire shared, "At first, AI seemed daunting, but after using the robot and reflecting on what I learned, I feel more confident. I believe I can master AI on one day."

Finally, in relation to "satisfaction", the overall positive feedback suggest that students found the AEER approach and the use of the AlphAl learning robots satisfying. All of the students reported they would like to join the workshop to learn Al if given opportunity. A number of students said they were motived to learn more about Al after attending the study.

#### 5. Conclusion

The pilot study demonstrated promising results regarding the impact of the AEER approach on enhancing primary school students' AI conceptual understanding, learning-to-learn skills, and motivation. The findings of the study show that students' overall learning outcomes and learning-to-learn skills increased significantly. The findings suggest that the AEER approach holds considerable potential for the future of education, providing a model for how we can equip students with the knowledge and skills they need to prepare for an AI-infused future.

A number of avenues for future research are suggested. First, the sample size of this pilot study was small. Thus, it would be beneficial to replicate the study with a larger and more diverse sample in order to generalize the findings. Second, the intervention duration of this study was short. Therefore, longitudinal studies will be conducted to examine the long-term effects of the AEER approach on students' AI literacy and learning-to-learn skills. Lastly, research could explore the integration of the AEER approach across various age groups (i.e., secondary school students, university students), and how the approach can be tailored to meet the specific needs of different learners.

#### References

- Al Hakim, V. G., Yang, S. H., Liyanawatta, M., Wang, J. H., & Chen, G. D. (2022). Robots in situated learning classrooms with immediate feedback mechanisms to improve students' learning performance. *Computers & Education*, 182, 104483.
- Azevedo, R., & Wiedbusch, M. (2023). Theories of metacognition and pedagogy applied in AIED systems. *Handbook of Artificial Intelligence in Education*, 45-67.
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., & Eirug, A. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994.
- Dehaene, S. (2020). How we learn: The new science of education and the brain. Penguin UK.
- Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1-4.
- Flavell J. H., (1979), Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry, *American Psychologist*, 34, 906.
- Martin, M., Chevalier, M., Burton, S., Bonvin, G., Besançon, M., & Deneux, T. (2023). Effects of introducing a learning robot on the metacognitive knowledge of students aged 8-11, *Proceedings Robotics in Education*, Cyprus, April 2023.
- Noh, J., & Lee, J. (2020). Effects of robotics programming on the computational thinking and creativity of elementary school students. *Educational Technology Research and Development*, 68, 463-484.
- Kong, S. C., Cheung, W. M. Y., & Zhang, G. (2021). Evaluation of an artificial intelligence literacy course for university students with diverse study backgrounds. *Computers and Education: Artificial Intelligence*, 2, 100026.
- Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7, 351-371.
- Su, J., Ng, D. T. K., & Chu, S. K. W. (2023). Artificial Intelligence (AI) Literacy in Early Childhood Education: The Challenges and Opportunities. *Computers and Education: Artificial Intelligence*, 4, 100124.
- Thrun, S., & Pratt, L. (Eds.). (2012). *Learning to learn*. Springer Science & Business Media.
- Vainikainen, M. P., Wüstenberg, S., Kupiainen, S., Hotulainen, R., & Hautamäki, J. (2015). Development of learning to learn skills in primary school. *International Journal of Lifelong Education*, 34(4), 376-392.
- Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of Instructional Development*, *10*(3), 2-10.
- Zhong, B., & Xia, L. (2020). A systematic review on exploring the potential of educational robotics in mathematics education. *International Journal of Science and Mathematics Education*, *18*, 79-101.

## Exploring the Benefits of Chatbots Gamebased Learning in Science Learning Outcomes and Motivation

#### **Ching-Huei CHEN**

Department of Industrial Education and Technology, National Changhua University of Education, Taiwan

chhchen@cc.ncue.edu.tw

Abstract: In this study, we aimed to assess the effectiveness of using Chatbot-assisted game-based learning to enhance science learning outcomes and motivation. A total of 202 seventh-grade students participated in the study and were randomly assigned to three groups: (1) Game only, (2) ChatbotGame, and (3) ChatbotGame with examples. The experimental groups received game-based learning with the assistance of Chatbot, either with or without examples, while the control group only received gamebased learning without any Chatbot support. The results indicated that students in the ChatbotGame with examples group achieved significantly better learning outcomes compared to those in the Game only group. Moreover, students in both the ChatbotGame and ChatbotGame with examples groups reported significantly higher levels of perceived competence compared to the Game only group. Additionally, the data collected from learning analytics and interviews suggested that Chatbot-assisted game-based learning can effectively enhance students' intrinsic motivation and promote positive learning behaviors in the context of science education. These findings hold important implications for the design and implementation of Chatbots in gamebased learning environments, as they have the potential to significantly improve students' learning outcomes and motivation.

Keywords: Chatbot, game-based learning, science learning, motivation, ChatGPT

#### 1. Introduction

The rise of digital game-based learning (DGBL) in educational settings has been populated by its ability to boost student engagement, motivation, and learning performance, as evidenced by studies conducted by Clark et al. (2016). Nonetheless, creating effective DGBL environments poses challenges, as it necessitates careful consideration of factors like game mechanics, instructional design, and learning objectives, as highlighted by Cai et al. (2022). Moreover, without appropriate scaffolding during DGBL, students may become overly focused on game mechanics or engage in various non-learning behaviors, as warned by Chu et al. (2022).

As the product of artificial intelligent, virtual agents such as Chatbots have the capability to generate human-like responses in natural language conversations. Numerous studies have examined the effects of Chatbots on learning outcome, yet scant studies exploring its potentials in game-based learning (GBL). To fully unlock the advantages of Chatbots-assisted GBL, further research is necessary to gain a deeper understanding of the most effective design and implementation of such environments, while also addressing challenges related to overreliance on Chatbots and the importance of effective design and implementation.

#### 2. Review of relevant literature

#### 2.1 Digital game-based learning

Digital Game-Based Learning (DGBL) is an educational approach that employs games and digital technologies to enhance learning outcomes. It integrates game design principles like problem-solving and interactive engagement to make learning more enjoyable and increase motivation and knowledge retention among learners. Existing research has shown that DGBL can boost engagement, motivation, and knowledge transfer among learners (Westera, 2015). However, there are challenges associated with designing effective educational games that align with learning objectives (Ke, 2016). Some learners might not be interested in gaming or find it difficult to learn through this approach (Hao et al., 2019).

To overcome and address challenges faced in DGBL, artificial intelligence (AI) technology may offer alternated and optimal solutions in DGBL. For example, AI can analyze student gameplay behaviors to predict early knowledge or assess learned knowledge more accurately (Alonso-Fernández et al., 2020). Unlike traditional intelligent tutoring systems, AI can offer personalized feedback, pinpointing areas of mistakes, and suggesting alternative approaches (Alam, 2021). This advances the potential of DGBL in creating more effective and tailored learning experiences.

#### 2.2 Chatbots

Chatbots utilized natural language processing techniques is becoming increasingly important in many fields because of its ability to process and analyze large amounts of data and perform complex tasks quickly and accurately. A recent noteworthy application of chatbots is ChatGPT, a language model developed by OpenAI, which has garnered significant attention and found use in diverse areas like customer service, education, and entertainment. By leveraging machine learning algorithms, ChatGPT generates responses based on patterns and relationships learned from extensive training data. As it continues to learn and improve from additional data and user feedback, this tool has the potential to revolutionize technology interactions, offering more personalized and engaging experiences for users (Hassani & Silva, 2023).

Although some studies have already demonstrated ChatGPT's effectiveness as a valuable reference and self-learning tool, further rigorous and empirical research is necessary to comprehensively assess its impact on student performance and behavior (Lo, 2023). The purpose of this study is to explore the effects of integrating ChatGPT in game-based learning on students' science learning outcomes and intrinsic motivation.

#### 3. Methodology

#### 3.1 Design and participants

A quasi-experimental study design was employed to answer the proposed research questions. A total of 202 seventh graders (103 male and 99 female) from nine classes in two different middle schools were recruited and randomly assigned to three treatment groups: (1) Game only, (2) ChatbotGame, and (3) ChatbotGame with examples.

#### 3.2 ChatGPT-assisted game-based learning

The GPT coding application programming interface (API) with access to OpenAI's GPT language (gpt-3.5-turbo) model was used in this study. With this API, we can integrate the GPT language model into the game, allowing it to generate human-like language and respond to user input in a natural and coherent manner. The GPT coding API is an essential tool used to incorporate NLP capabilities into *Summon of Magicrystal*. GPT-assisted *Summon of Magicrystal* represents an exciting new approach to digital game–based learning that utilizes cutting-edge language processing technology to provide a more engaging, personalized, and effective learning experience for students. As shown in Figure 1, students

can click on Ask-AI for a conversation forum to pop-up so that they can ask questions. In addition to students initiating the questioning, the system also provide examples for them to ask questions.



Figure 1. ChatGPT with examples in Summon of Magicrystal

#### 3.3 Instrument

#### 3.3.1 Science learning outcomes

In order to evaluate the students' conceptual understanding in response to this study, a conceptual test was created by two experienced science teachers with 20 years of expertise. This test served as both a pretest and a posttest and comprised 16 multiple-choice questions centered on the concepts and principles addressed during the gameplay. The maximum score achievable on the test was 100. To ensure its reliability and validity, the questions underwent rigorous examination, revision, and validation by two other seventh-grade science teachers. Additionally, pilot testing with seventh-grade students was conducted. The Cronbach's alphas for the pretest and posttest were found to be 0.82 and 0.80, respectively, indicating good internal consistency and reliability of the test.

#### 3.3.2 Learning motivation

The intrinsic motivation questionnaire utilized in this study was adapted from Deci and Ryan's (2000) instrument and consisted of 14 items, distributed across four subscales. These subscales were interest (four items, e.g., "I enjoyed playing this game very much"; Cronbach's alpha = .73), perceived competence (three items, e.g., "I think I am pretty good at this game"; Cronbach's alpha = .91), tension (three items, e.g., "I felt very tense while playing this game "; Cronbach's alpha = .87), and perceived value (four items, e.g., "I believe this game could be of some value to me"; Cronbach's alpha = .94). Using a seven-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree), the students were asked to rate these items.

#### 3.4 Procedure and data analyses

Participants were recruited from their respective classes according to the regulations of the university's institutional review board. The study lasted 2 weeks, with one 50-min session each day. The participants began by taking a multiple-choice conceptual pretest. Following that, the researchers informed the students about the study purposes, explained the game environment, and assigned each student a unique username and password. The students in the ChatbotGame and ChatbotGame with examples groups were informed of the capabilities of such functions in the gameplay. Toward the end of the gameplay, the participants completed a posttest and a web-based survey that included intrinsic motivation questionnaire. Analysis of covariance (ANCOVA) was performed to evaluate science

learning outcomes with a significance level of p < 0.05, and a t-test was used to assess intrinsic motivation. All statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS 11.0 for Windows).

#### 4. Results

#### Effects on science learning outcomes and learning motivation

One-way ANCOVA was adopted to examine the three groups' science learning outcomes. The ANCOVA results revealed that after excluding the effects of the pretest scores, there was a significant difference between groups (F = 5.94, p = 0.003), indicating that the students had significantly different learning achievements as a result of using the different learning models. The posttest score of the ChatbotGame with examples group (M = 49.06) was significantly higher than the game only group (M = 44.46). This indicates that the students who used ChatbotGame with examples outperformed those who used the game alone.

One-way analysis of variance (ANOVA) was employed to explore the effects of the different gameplay modes on students' motivation. There was a significant difference in perceived competence among the three groups (F = 3.69; p = 0.03). The average perceived competence scores of the ChatbotGame with examples and ChatbotGame groups and the Game only group were 2.63, 2.54, and 2.21, respectively. While there was no significant difference among the three groups in terms of interest, tension, and value, the students in the ChatbotGame with examples and ChatbotGame groups reported higher perceived competency than those in the Game only group.

#### 5. Discussion and conclusions

The purpose of this study was to examine the potential of Al-assisted GBL in terms of students' science learning outcomes and intrinsic motivation. In terms of the students' science learning outcomes, it was found that there was a significant difference between the ChatbotGame with examples and Game only groups. Those students learning with the assistance of examples provided by ChatbotGame outperformed those learning only within the game. The students learning without examples also did better than those who learned with only the game. To investigate the students' perceptions of the two different ChatbotGame modes further, some students in both groups were selected for in-depth interviews. Several students who had learned with examples provided by ChatbotGame pointed out that those examples helped direct their attention to relevant concepts and facilitated their gaming strategies. However without the examples, the students found that it was difficult to initiate questions because they had limited knowledge of the subject matter and were not sufficiently well-versed in the topic for them to come up with meaningful and relevant questions.

Overall, the use of Chatbots-assisted GBL has the potential to revolutionize education by creating personalized, adaptive, and engaging learning environments that meet the needs of every student. However, more research is needed to fully understand the potential of these technologies and to ensure that they are implemented in an ethical and responsible manner (Lo, 2023). Our study yields theoretical and practical implications for Chatbots-assisted DGBL. It theoretically incorporates ChatGPT into DGBL, providing insights into how learners' interactions with conversational agents help them strengthen their knowledge and revealing the capabilities of NLP to enable more conversational and interactive learning experiences. Practically, ChatGPT can be utilized as an intelligent assistant within DGBL, offering examples or explanations to help learners navigate challenging tasks or overcome obstacles. The examples can foster deeper understanding of complex concepts as well as enhance students' problem-solving abilities.

This study has several limitations that future research should address. The first concerns the transfer effect of Chatbots-assisted GBL. Future research could investigate, for

example, how ChatGPT can support the transfer of learning from game-based environments to real-world contexts. This includes exploring techniques to bridge the gap between virtual experiences and real-life applications, ensuring that the skills and knowledge acquired through DGBL can be effectively transferred and applied. A second area of future research concerns the effectiveness of ChatGPT in DGBL varies across different subject areas and age groups. Continued exploration in these areas can contribute to the development of more effective and engaging learning experiences for students.

#### References

- Alam, A. (2021, 3-4 Dec. 2021). Should robots replace teachers? Mobilisation of AI and learning analytics in education. International Conference on Advances in Computing, Communication, and Control (ICAC3), Mumbai, India.
- Alonso-Fernández, C., Martínez-Ortiz, I., Caballero, R., Freire, M., & Fernández-Manjón, B. (2020). Predicting students' knowledge after playing a serious game based on learning analytics data: A case study. *Journal of Computer Assisted Learning*, *36*(3), 350-358.
- Cai, Z., Mao, P., Wang, D., He, J., Chen, X., & Fan, X. (2022). Effects of scaffolding in digital gamebased learning on student's achievement: A three-level meta-analysis. *Educational Psychology Review*, *34*(2), 537-574. <u>https://doi.org/10.1007/s10648-021-09655-0</u>
- Chu, S. T., Hwang, G. J., Chien, S. Y., & Chang, S. C. (2022). Incorporating teacher intelligence into digital games: An expert system-guided self-regulated learning approach to promoting EFL students' performance in digital gaming contexts. *British Journal of Educational Technology*.
- Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, 86(1), 79-122.
- Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the selfdetermination of behavior. *Psychological Inquiry*, *11*(4), 227-268. https://doi.org/10.1207/S15327965PLI1104\_01
- Emerson, A., Min, W., Azevedo, R., & Lester, J. (2023). Early prediction of student knowledge in game-based learning with distributed representations of assessment questions. *British Journal of Educational Technology*, *54*(1), 40-57. https://doi.org/https://doi.org/10.1111/bjet.13281
- Hao, Y., Lee, K. S., Chen, S.-T., & Sim, S. C. (2019). An evaluative study of a mobile application for middle school students struggling with English vocabulary learning. *Computers in Human Behavior*, 95, 208-216. <u>https://doi.org/10.1016/j.chb.2018.10.013</u>
- Hassani, H., & Silva, E. S. (2023). The role of ChatGPT in data science: How AI-assisted conversational interfaces are revolutionizing the field. *Big Data and Cognitive Computing*, 7(2), 62. <u>https://www.mdpi.com/2504-2289/7/2/62</u>
- Ke, F. (2016). Designing and integrating purposeful learning in game play: A systematic review. *Educational Technology Research & Development*, 64(2), 219-244. <u>https://doi.org/10.1007/s11423-015-9418-1</u>
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*(4), 410.
- Westera, W. (2015). Games are motivating, aren't they? Disputing the arguments for digital gamebased learning. *International Journal of Serious Games*, 2(2), 3-17. <u>https://doi.org/doi.org/10.17083/ijsg.v2i2.58</u>
# Enhancing Trigonometry Learning through a Mobile App

### Maria Alva Q. ABERIN<sup>a,\*</sup>, Ma. Louise Antonette N. DE LAS PEÑAS<sup>a</sup>, Agnes D. GARCIANO<sup>a</sup>, Juan Carlo F. MALLARI<sup>a</sup>, Jumela F. SARMIENTO<sup>a</sup>, Mark Anthony C. TOLENTINO<sup>a</sup>, & Debbie Marie B. VERZOSA<sup>b</sup>

<sup>a</sup>Department of Mathematics, Ateneo de Manila University, Philippines <sup>b</sup>Department of Mathematics and Statistics, University of Southern Mindanao, Philippines \*maberin@ateneo.edu

**Abstract:** A mobile app called *Trigomatch* is designed for students to gain an in-depth understanding of trigonometric concepts in an engaging and dynamic learning environment. A description of the app as well as the pedagogical framework and game design principles that form the backbone of the app are discussed.

Keywords: Mobile app, trigonometry, trigonometric ratios

# 1. Introduction

Trigonometry is one of the topics in high school mathematics that involves developing skills in algebra and geometry, which is essential in studying pre-calculus and calculus. Moreover, knowledge of trigonometric functions is needed for understanding topics in Newtonian physics, architecture, surveying, and many branches of engineering (Weber, 2005).

Oftentimes, students have difficulties in problem solving in trigonometry. One reason is due to lack of motivation (Nurmeidina & Rafidiyah, 2019). Difficulty is also due to understanding the concepts of trigonometry and its applications (Nanmumpuni & Retnawati, 2021). Understanding multiple representations of trigonometric functions can also be challenging to students (Marchi, 2012). Students have difficulty in the presentations of the topic: the unit circle and the right triangle approach (Jenkins, 2022).

Studies in the literature have shown the benefits of technology in studying trigonometry. Bedada and Machaba (2022) showed that teaching trigonometric functions with the aid of GeoGebra (https://www.geogebra.org/) has significant effects on the achievement of students. Similarly, Rahman et al. (2016), using an interactive GeoGebra Learning Module, showed that students developed positive motivation in studying trigonometry. Emerging technologies such as *TrigReps* (Bornstein, 2020) and *Tangible User Interface (TUI)* (Zamorano Urrutia et al., 2019) proved to have the potential to facilitate learning of trigonometry. The representations offered by *TrigReps* (Bornstein, 2020) assisted the students in connecting their modifications to the algebraic representation with the subsequent modifications to the graphical representation. Moreover, *Tangible User Interface (TUI)* (Zamorano Urrutia et al., 2019) provides a pedagogical experience that prioritizes physical manipulation-based exploration and encourages intuitive and group learning.

Taking advantage of the benefits of technological tools (Haleem et al., 2022) to build knowledge, understanding, and skills in trigonometry and to increase engagement and learning motivation in students, we designed a mobile app called *Trigomatch*. Designed through an interactive game with interactive feedback, the app focuses on the understanding of fundamental concepts in trigonometry such as reference angles and trigonometric ratios.

# 2. Pedagogical Basis

Drawing on the work of Hans Freudenthal, Martín-Fernández et al. (2022) distinguished between trigonometry from the perspective of elementary geometry and goniometry,

corresponding to the right triangle and unit circle approaches presented by Jenkins (2022). In the former, angles are absolute and static, and are determined by the sides of a right triangle in a non-oriented plane. In the latter, angles are dynamic, and determined by a rotation of halflines in an oriented plane. They found that transitions between these two systems of thought are not straightforward–even pre-service teachers cling to their conceptions of a static angle and barely know how to draw oriented angles in a circle.

A possible intervention for facilitating understanding of trigonometry is through the presentation of trigonometric representations through interactive computer-aided instruction (CAI). Jenkins (2022) argued that when visual representations of trigonometry are presented using interactive technology, they are more accurate and allow the learners to proceed at their own pace. Further, because of the accuracy of these presentations, they can be used to reinforce images in a person's memory, which is not the case for hand-drawn images which tend to be distorted. While she found that students can learn as well from CAI as from traditional instruction, the use of CAI allows the teachers to differentiate instruction and personalize tasks for their students.

In common with computer-aided technology described by Jenkins (2022), the *Trigomatch* app presents visual representations of basic trigonometric concepts to allow students to extend their understanding of trigonometry from the right triangle to the unit circle approach. The app is interactive in the sense that it presents the concept with feedback mechanisms in a game-like environment. More details on the app are provided in the next section.

# 3. The Trigomatch App

#### QUADRANTS Match the radians or degrees to the correct quadrant! CONVERSION Match the degree to its radian! $\leftarrow$ $\leftarrow$ × × Timer Timer $\frac{6\pi}{5}$ $\frac{9\pi}{10}$ $\frac{5\pi}{18}$ 120 162 162° OII 01 $\frac{\pi}{2}$ 216 126 50 QII QIII QIV 240 $\frac{7\pi}{10}$ $\frac{2\pi}{3}$ 90 162 QII 315 40° 135 (b) (a) COTERMINAL ANGLES TRIGONOMETRIC RATIO 4 $\leftarrow$ Timer Timer 1 2 370 <u>√3</u> 510 $\frac{2\sqrt{3}}{2}$ $\frac{2\sqrt{3}}{3}$ -60 $\sqrt{2}$ 240 100 1 30 (d) (c) TRIGONOMETRIC FUNCTION Match the trigonometric functions! IDENTITIES Motch the identities! - $\leftarrow$ Timer Timer $2\cos^2 A$ Ð $\frac{1}{\csc A}$ $\frac{2\sqrt{3}}{3}$ $(\mathcal{P})$ $\sin 2A$ 1 $2 \sin A$ $\cos A$ $\sin A$ $\Delta^{-}$ -2 -1 $\cos 2A$ sec A sec A cos A 09 $\sqrt{3}$ tan A $\sin A$ $\frac{1}{\cot A}$ cos A (e) (f) Figure 1. Screenshots showing the six options in Trigomatch.

### 3.1 App Description

In the *Trigomatch* app, six (6) different options can be played: *Conversion* (Figure 1(a)), *Quadrants* (Figure 1(b)), *Coterminal angles* (Figure 1(c)), *Trigonometric Ratio* (Figure 1(d)), *Trigonometric Function* (Figure 1(e)), and *Identities* (Figure 1(f)). Moreover, learners may choose to display 12, 20, or 36 cards at a time and then match two equivalent representations or trigonometric expressions shown by the cards. The objective is to locate every pair that matches within the allotted time. The game's time aspect feature has the potential to motivate and thrill the learners while they play.



Figure 2. Screenshots showing the different levels, with or without pictures.

# 3.2 Game-design Factors

One of the challenges that students face in learning trigonometry is the presence of several trigonometry facts, conversions, and identities that they need to memorize. These include, among others, conversions between radians and degrees, the trigonometric function values for special angles, and the sum, product, half-angle, and double-angle identities. While these can be learned through rote memorization or drill-and-practice, students usually tend to find such tasks tedious and uninteresting. Thus, a game-based learning approach has the potential to support students' learning of these trigonometry topics. With game-based learning, we aim to increase students' motivation and engagement by situating the learning of trigonometry within a game that can create high situational interest (Hidi & Renninger, 2006).

In the design and development of *Trigomatch*, we took into consideration 11 crucial game-design factors identified in the Game-based Learning (GBL) Design Model of Shi and Shih (2015). Figure 3 presents a summary of how these 11 game-design factors have been applied for *Trigomatch*.



Figure 3. Game-design factors (Shi & Shih, 2015) for Trigomatch

The first game-design factor is the game goal, which significantly influences all the other game-design factors. In the case of *Trigomatch*, the game goal is a learning objective focused on trigonometry topics. The game mechanism is then decided towards achieving the game goal. As previously described, the mechanics of *Trigomatch* involve matching pairs of cards (akin to some memory games) that contain equal or equivalent trigonometric expressions. Learners interact with the game through simple click or tap actions on game buttons and cards. An important feature of the game mechanism is adaptivity (Plass et al., 2015) in the sense that learners have the freedom to select from different topics, modes, and options. This allows learners to align their playing experiences with their current level or interest with respect to the learning content. The availability of these different topics, modes, and options also contributes to the game's mystery, which in turn can keep learners interested in progressing in the game. The main challenge in *Trigomatch* is to match all the cards within the time limit. As part of adaptivity, learners can disable the timer if they are still beginning to learn the topics.

*Trigomatch* is envisioned to be a simple arcade-like game with relatively short rounds that can be played repeatedly. Thus, the game fantasy and narrative are kept simple and are delivered mainly through the game area and game objects such as cards, timer, and buttons. In terms of sensation, *Trigomatch* has display and sound prompts during gameplay. Moreover, as the game is also intended for independent use, *Trigomatch* has been designed to be a single-player game. However, teachers can use the app as part of group activities in the classroom, face-to-face or virtual. All the features of *Trigomatch* contribute to its game value, which is oriented towards creating situational interest among learners.

# 4. Integration and Use of *Trigomatch*

This section describes how the *Trigomatch* can be used as a pedagogical tool based on the RAT (Replacement, Amplification, Transformation) framework (Hughes et al., 2006). Using this framework, the app can replace the traditional method of providing examples and exercises for students to solve without changing the learning goals of a lesson. The printed worksheets or exercises from a textbook have been replaced by a set of cards shown in a more appealing game-like setting. This function can help teachers offer exercises in an engaging and enjoyable manner.

Playing the app can enrich or amplify a student's experience in solving problems involving concepts in trigonometry. The design of the app (*with or without pictures*) gives students the option to work at just the correct degree of difficulty and includes components that can excite and push them to work on various levels.

Finally, the *Trigomatch* app may encourage teachers to shift from teacher-centered to student-centered instruction. This is achievable because the app was made to be simple enough for students to use and comprehend without much help from a teacher. Students who are learning asynchronously or remotely can thus utilize the app. Task redefinition is another way that transformation might take place. For instance, utilizing the app, one may convert individual work into group activities. This allows the students a chance to interact and instruct one another on the problem-solving strategies needed to match two trigonometric representations displayed on the cards.

### 5. Conclusion and Future Outlook

Advanced mathematics and other allied disciplines, such as engineering and physics, depend heavily on trigonometry. However, many students struggle to understand and apply trigonometric principles and often lack the motivation to learn these. The *Trigomatch* app presented in this paper is a tool that aims to help students gain a better understanding of trigonometric concepts and reinforce the skill of recognizing equivalent trigonometric representations. Game design principles were employed in the app's design to keep students motivated and engaged. By playing the app, students can recall trigonometric ratios, conversions, identities and coterminal angles in an enjoyable and productive way. The option to display pictures as one plays the app is a useful feature especially for learners who may benefit from the visual representations of trigonometric ideas.

This study is work in progress. Opportunities for further research on how *Trigomatch* is used by students and how effective it can be to improve their trigonometric skills is currently being pursued. An experimental study with pretest and posttest research design is going to be used to accomplish this. Furthermore, the app may be expanded by adding other options such as trigonometric functions and their graphs or using the laws of sines and cosines. This will help build on the skills that students have learned and allow them to advance in other topics.

### Acknowledgements

This paper is one of the outputs of the Ateneo Mathplus Resources Laboratory housed at the Department of Mathematics, School of Science and Engineering, Ateneo de Manila University. The authors thank the Department of Science and Technology-Philippine Council for Industry, Energy, and Emerging Technology Research and Development (DOST-PCIEERD) and the University Research Council (URC), Ateneo de Manila University for the support of the development of the mathematical resources for Grades 1 to 11 Mathematics. Acknowledgement also goes to Dr. Ma. Mercedes T. Rodrigo and the Ateneo Laboratory for the Learning Sciences (ALLS) for providing technical assistance. We thank our app developers Victor Antonio M. Ortega, Mr. Nigel Benedict Cargo, Mr. Jose Teodoro Lacson, and Mr. Amiel Damian F. Justiniani for the apps and video art assets.

### References

- Bedada, T. B., & Machaba, F. (2022). The effect of GeoGebra on STEM students learning trigonometric functions. *Cogent Education*, *9*(1), 2034240.
- Bornstein, N. (2020). Teaching transformations of trigonometric functions with technology. *Journal of Interactive Media in Education*, 2020(1), 1-9.
- Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, *3*, 275-285.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111-127.
- Hughes, J. E., Thomas, R., & Scharber, C. (2006). Assessing technology integration: The RAT Replacement, Amplification, and Transformation – Framework. In C. M. Crawford, R. Carlsen, K. McFerrin, J. Price, R. Weber, & D. A. Willis (Eds.), *Proceedings of SITE 2006--Society for*

*Information Technology & Teacher Education international conference* (pp. 1616-1620). Orlando, Florida, USA: Association for the Advancement of Computing in Education (AACE).

- Jenkins, R. (2022). Impact of interactive computer-aided instruction in learning trigonometry in a high school precalculus course. [Doctoral dissertation, Franklin University]. https://fuse.franklin.edu/ docpub/71
- Marchi, D. (2012). A study of student understandings of the sine function through representations and the process and object perspectives [Doctoral dissertation, The Ohio State University]. http://rave.ohiolink.edu/etdc/view?acc\_num=osu1343253667
- Martín-Fernández, E., Rico, L., & Ruiz-Hidalgo, J. F. (2022). Conversions between trigonometric representations by pre-service secondary school teachers. *PNA (Revista de Investigacion en Didactica de la Matematica), 16*(3), 237-263.
- Nanmumpuni, H. P., & Retnawati, H. (2021). Analysis of senior high school student's difficulty in resolving trigonometry conceptual problems. *Journal of Physics: Conference Series*, 1776, 012012.
- Nurmeidina, R., & Rafidiyah, D. (2019). Analysis of students' difficulties in solving trigonometry problems. In N. Ishartono, M. S. Adhantoro, Y. Sidiq, & Y. Sulistyono (Eds.), *Proceedings of the 4th progressive and fun education international conference*, *Profunedu 2019*. EAI.
- Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, *50*(4), 258-283.
- Rahman, M. H. A., & Puteh, M. (2016). Learning trigonometry using GeoGebra learning module: Are under achieve pupils motivated? *AIP Conference Proceedings*, *1750*, 040001.
- Shi, Y.-R., & Shih, J.-L. (2015). Game factors and game-based learning design model. *International Journal of Computer Games Technology*, 2015, 1-11.
- Weber, K. (2005). Students' understanding of trigonometric functions. *Mathematics Education Research Journal*, 17(3), 91–112.
- Zamorano Urrutia, F. J., Cortés Loyola, C., & Herrera Marín, M. (2019). A tangible user interface to facilitate learning of trigonometry. *International Journal of Emerging Technologies in Learning*, *14*(23), 152-164.

# What's Going on Behind the Scenes? Peer Assessment as a Strategy for Enhancing Students' Commitment and Learning Motivation in Working Collaboratively

### Nur Aira ABD RAHIM<sup>a\*</sup>, Norliza GHAZALI<sup>a</sup>, Zakiah PONRAHONO<sup>b</sup>, Sharifah Intan Sharina SYED ABDULLAH<sup>a</sup>, Fadzilah MOHAMAD<sup>c</sup>, & Syahidatul Khafizah MOHD HAJARAIH<sup>d</sup>

 <sup>a</sup>Faculty of Educational Studies, Universiti Putra Malaysia (UPM), Malaysia
 <sup>b</sup>Faculty of Forestry & Environment, Universiti Putra Malaysia (UPM), Malaysia
 <sup>c</sup>Faculty of Medical & Health Sciences, Universito Putra Malaysia (UPM), Malaysia
 <sup>d</sup>Dept. of Marketing, Student Admission, and Corporate Communication, Universiti Tenaga Nasional (UNITEN), Malaysia

\*nuraira@upm.edu.my

Abstract: Peer assessment (PA) refers to any arrangement in which learners consider the amount, level, value, worth, quality or success of the learning outcomes of their peers. In simpler words, peer assessment is a form of assessment which allows the learners to provide grades/marks and/or feedback to their peers by assessing their peers' work and contribution. This paper conceptually explains about peer assessment as a strategy to increase students' commitment and motivation when working collaboratively. The first part of this paper discusses about the nature of collaborative in higher education setting. Essentially, peer assessment can be the key to uncovering what happened behind the scenes during the collaborative learning process, i.e. during group assignment or group project. In this research, an exploratory case study design was conducted among higher education instructors at a public university in Malaysia. This research seeks to explore the strategies and tools used by higher education instructors to conduct peer assessment in their teaching, with a focus on innovative approaches and digital tools used by these instructors in multidisciplinary educational contexts.

**Keywords:** peer assessment, alternative assessment, collaborative assessment, learning motivation

# 1. Introduction

Assessment is an important component in teaching and learning as it informs the instructors about their student's performance and progress during learning. One of the basic principles in conducting assessment in learning is to ensure that the assessment conducted is equitable and appropriate for the learning tasks that the students are assigned to do. The design of assessment influences the way in which the students will approach and engage in their learning task (Astin & Antonio, 2012). An assessment that focuses on repetition of facts will encourage the students to memorize, whereas assessment that evaluates class participation will prompt the students to be more active in class. Essentially, the type of assessment used can be a source of learning motivation for the students to learn and exhibit behavior or learning outcomes intended by the course instructor.

For learning tasks that requires the students to work together, such in collaborative learning setting, the appropriate assessment should be able to assess not just the outcomes of the group work, but to gauge the joint collective efforts and individual contribution towards the completion of the group work. Johnson et al (1990) discuss five basic elements in collaborative learning, involving: 1) positive interdependence among the group members; 2) considerable interaction in which the members help and encourage each other to learn; 3) individual accountability and personal responsibility for doing their share in the collective efforts; 4) engaging in social skills with team members; and 5) group self-evaluating for periodically assessing team's progress towards the common goals (as cited in Laal & Laal, 2012).

In this paper, firstly we would like to conceptually explain about peer assessment as a strategy for assessing collaborative work among students. So, what is peer assessment? Peer assessment (PA) refers to any arrangement in which learners consider the amount, level, value, worth, quality or success of the learning outcomes of their peers (Topping, 1998). In simpler words, peer assessment is a form of assessment which allows the learners to provide grades/marks and/or feedback to their peers by assessing their peers' work and contribution. There are wide range of PA activities and format existed, as peer assessment have been found to be conducted in many different curriculum areas and subjects (Topping, 1998, 2021). For example, Topping (2021) outlines 44 different variables that can explain the variations of peer assessment in the literature across different subject/fields. Array of terms are also used in the literature to describe a similar process, such as peer grading, peer feedback, peer evaluation, and peer review. Methods of peer assessment can also include peer nominations, peer ratings, and peer rankings (Kane & Lawler, 1978). The next part of this paper will discuss about the nature of collaborative learning in higher education practices, leading to the exploration of innovative strategies and digital tools used in support of peer assessment practices by higher education instructors.

# 2. The Implications of Collaborative Learning in Higher Education Practices

There are myriad of literature available reporting the impacts and implications of collaborative work in teaching and learning practices in higher education. Collaborative learning, or the experience of being able to work collaboratively, is often regarded as a valuable way of learning. It is also viewed as an important interpersonal skill, where students can learn to be a team player and develop their teamwork skills, which is a skill that is being linked to employability skills (Winterbotham et al., 2018). Notably, the merits of collaborative learning are well recognized in practice as well as in the literature. From the students' experiences and perspectives, working collaboratively through group work enables them greater value of learning via the development of a wider breath of knowledge through discussion, clarification of ideas, as well as evaluation of others' ideas (Hassanien, 2006). From faculty's standpoint, group work assessments allow them the opportunity to develop and assess various skills that are regarded as the 21<sup>st</sup> century skills, such as communication, team-building, problem-solving, collaboration, creativity and technical skills (Tumpa, Skaik, Ham, & Chaudhry, 2022).

At the same time, literature has also discussed some demerits associated with collaborative work, group work, or group assignments among students. A primary concern raised in perceptions of group work among faculty and students was often about the allocation of grades (LaBeouf et al, 2016). Group work is often perceived as more time consuming and therefore usually are assigned with higher grade allocation. However, as raised by Goold et al (2008), students typically do not favor being assessed as a group, as they feel penalized if and when some group members do not perform as expected. Ensuring participation from all members are often cited as one of the difficulties in working collaboratively (Piezon & Ferre, 2008). At the same time, from a faculty standpoint, it could also be difficult to assess the individual contributions within a group work (Quinn & Hughes, 2007). Therefore, when implementing group work/assignment, it is important to have clear guidelines not just on the assessment of the group task itself, but also transparency on how the contribution of each group member can be recognized and assessed fairly (Murray, 2003).

# 3. The Merits of Peer Assessment in Collaborative Learning

Based on the recollection of merits and demerits of collaborative learning presented above, there seems to be a dynamic tension between the collective nature of collaborative learning versus how they are being assessed. The nature of collaborative learning, due to the interdependent nature of their learning that happened within the team, presumable through collective efforts of the members, may requires assessment that goes beyond behind their team's finished work. Essentially, peer assessment can be the key to uncovering what happened behind the scenes during the collaborative learning process, i.e. during the conduct the group assignment or group project. For example, do the students really work together as a team? More importantly, did all the students contribute to the group project equally that merit all of them to receive the same marks for the completed project? Or should a team member who completed his or her task, be penalized for an incomplete project because another team member did not deliver his or her task accordingly?

With considerations of the above, for faculty wanting to implement collaborative work or group assignments in their courses, what is the effective way forward in practice? As a key takeaway to this inquiry, a comprehensive semi-systematic literature review was conducted by Fittipaldi (2020) to discover the best practices on managing the dynamics of group projects in higher education based on reviews of empirical research. More importantly, of a particular interest of this paper, Fittipaldi's (2020) review revealed a significant support for implementing a formative assessment for evaluating collaborative learning, without negating the importance of summative assessment (assessment of the finished task/product). Herein, the use of peer assessment, peer review, and/or peer rating, as a form of evaluation, have been shown to reduce issues of social loafing or free riding in group work. It is also perceived as the fairest and best method for assessing individual contributions within a group work. Essentially, according to Fittipaldi (2020), the data clearly indicate the value of peer assessment in successfully managing the dynamics of group projects in higher education.

# 4. Research Purpose and Design

The literature above has unfolded the body of scholarship that supports the use of peer assessment in higher education teaching practices. Peer assessment has been identified as a versatile approach that can be used in multidisciplinary settings (Topping, 2021). According to Fittipaldi (2020), a range of peer assessment instruments and strategies exists in the literature. With the rapid use of digital tools being integrated in teaching and learning these days, there seems to be limited research that examines the use of digital tools and applications in the context of peer assessment practices. This investigation aims to explore the strategies and tools used by higher education instructors to conduct peer assessment in their teaching, with a focus on innovative approaches and digital tools used in multidisciplinary educational contexts. Two primary research questions guided this study: 1) How do higher education instructors conduct their peer assessment in their teaching? 2) What are the digital tools used in support of the peer assessment practices among higher education instructors?

This study employs an exploratory case study design. This study is being conducted at one of the public higher education institutions in Malaysia and will be carried out in two phases (currently in phase one). In the first phase, we are conducting data collection for determining the peer assessment practices among the lecturers across different faculties at our university. In order to ensure representation of data collected from multidisciplinary practices, a maximum variation sampling strategy was used in targeting sample with at least one representative from all faculties at our university (from a total of 14 faculties). The goal of this sampling strategy is to ensure inclusivity of peer assessment practices from diverse learning environments.

Data are gathered using interviews and observation (where applicable). During the data collection, each instructor will be asked to explain how they conducted peer assessment in

the context of subjects that they taught/are teaching, as well as the digital tools that they used to support their peer assessment practices. From the interview discussion, it will be followed by observation in their classes or demonstration of the digital tools used, where applicable. If made available by the instructors, the data collection also includes documents related to peer assessment, such as evaluation form, scoring rubric, written instructions dan others.

The second phase of data collection will take place after the completion of the interviews. Two primary indicators will be used to determine sufficiency of interviews conducted, based on i) obtaining representation from all faculties of different fields in the data collection, and ii) point of data saturation. At current, this study has completed about 8 interviews from 7 faculties representation. The second phase of this research will focus on developing a framework on peer assessment practices across multidisciplinary settings. Data from the interviews, observation, and document analysis will then be triangulated and synthesized to identify emerging patterns for building the framework, which include integration of digital tools and innovative approaches used to support peer assessment, as reported by the data.

# 5. Preliminary Findings

This research is still currently ongoing with 8 interviews completed from 7 different faculties representing different field of studies (Faculty of Educational Studies, Faculty of Medicine and Health Sciences, Faculty of Science Computer and Information Technologies, Faculty of Biotechnology and Biomolecular, Faculty of Agriculture, Faculty of Forestry and Environment, and Faculty of Humanities, Management and Science.

Our preliminary findings so far point towards instructors practicing peer assessment as both a form of formative and summative assessment, and that it is very commonly practiced by instructors from different fields to enrich the learning process and improve the assessment practice in their teaching. Generally, these instructors identified with two ways of conducting their peer assessments:

Assessing contribution. In this PA approach, group members evaluate each other's performance in regards to their contributions to the group assignment/collaborative project. The PA is usually conducted after the completion of the project. This form of PA is typically done as part of summative assessment for the group project (please refer Figure 1)



Figure 1: Peer assessment process for assessing contribution

Assessing coursework. In this PA approach, the assessment of the coursework involves the evaluation of the product of the project / presentation evaluation by their peers. Usually, this approach entails feedback that allows for improvement of the coursework. This form of PA can be done both as a formative and summative assessment (please refer Figure 2)



Figure 2: Peer assessment process for assessing coursework

As reported by the participants, the most common tools used to support the conduct the peer assessment was by developing rubric and preparing peer evaluation form to guide the peer assessment. While these components were the standard structure used, few instructors also incorporated some innovative approaches to conduct their peer assessment, as follow:

i. Gamification strategy in peer assessment

One instructor shared an approach where he used gamification approach to get students involvement in the peer assessment. Incorporating gamification principles into peer assessment allow the instructor to incorporate rewards and incentives to motivate students and creating a sense of achievement as they progress through levels of understanding acquired from the lecture.

The instructor begins the class with a current real-life case that is related to the class's topic, and proceed to teach the subject matter that could potentially help resolve the case presented. Next, the class with proceed with a game activity, with groupings of class members assign to various roles within the game. The peer assessment was incorporated during the gamification process where the students were asked to reward their team members' contribution using Star Reward system based on their contribution during the game, as well as feedback after the game has completed.

In the literature, another similar approach was conducted by Simionescu, Sukinova, and Kubincova (2017). Using gamified peer assessment, these instructors design the reward system through the use of badges generated from the learning management system.



Note: These are examples of the badges type used for gamified peer assessment. From, "Gamification of Peer Assessment in Learning Management System" by S. Simionescu, D. Sukinova, and Z. Kubincova, 18th International Carpathian Control Conference (ICCC), 2017, pp. 571 – 575, doi: 10.1109/CarpathianCC.2017.7970465

These gamification elements offer immediate feedback, fostering engagement and driving a continuous improvement cycle in the assessment process, ultimately enhancing the quality of feedback provided by peers.

ii. Role play strategy in peer assessment

Another instructor reported a peer assessment approach that was incorporated through role playing in his assessment. Role play in peer assessment involves creating scenarios that mimic real-life situations. This allows students to apply their knowledge and skills in a context similar to what they might encounter in their future careers or daily lives. In role playing, students are assigned different roles within the scenario, each with specific responsibilities, objectives, or viewpoints. These roles can be based on characters, professions, or situations relevant to the subject matter.

For this instructor, the created scenario was based on "tour packages" created by the groups of students for their assignment. The execution of the tour package was then being assessed by their peers through the viewpoints of tourists (their peers role playing as tourists). The marks used for the peer assessment took place in a form of currency, which then converted to a total of 100 marks, paid by the tourists to the operators of the tour packages. This role-play approach encourages students to assess their peers objectively based on predefined criteria (defined by the instructor using a rubric), often linked to the roles they are portraying. Engaging in role play as assessors encourages interactive learning, as students actively discuss and reflect on their own and their peers' performances. It can also make the assessment process more engaging and enjoyable for students, as they are actively involved in scenarios that mimic real-life experiences.

Integrating role play with the use of peer assessment was shown to be an effective pedagogical strategy (Ferrari, Spoletini, Bano & Zowghi). Through their quasi experimental study, these instructors found that the students made less mistakes from their 1<sup>st</sup> to 2<sup>nd</sup> interview conducted through the use of role-playing and mediated through both self-assessment and peer review during the assessment process.

Incorporating role play into peer assessment can be a powerful method for evaluating a wide range of skills and competencies. It encourages students to think critically during

the moment, communicate effectively, and apply their knowledge in practical situations as mimic by the role and situations created for their learning.

iii. Digital tools used in peer assessment

At the current stage of data collection, the reports on digital tools used by course instructors to support the peer assessment have been quite limited. From the interviews, many instructors reported preferences for manual handling of the peer assessments or supported with very basic digital tool such as Google Form to convert the evaluation from paper-based form to electronic form (see one example shared via Google Form link: https://tinyurl.com/UPMICCE2023). This preference for manual methods can, at times, be attributed to factors such as familiarity with established practices, concerns about technological complexities, in availability or inaccessibility of the digital tools, or a preference for a more hands-on, personalized approach to assessment. The use of other digital tools in support of peer assessment will be further explored in the remaining data collection process of this research.

# 7. Conclusion

This preliminary paper served as a precursor for our ongoing research on developing a peer assessment framework in multidisciplinary setting. Our team consist of members from different disciplinary, as we are currently conducting research on various instructors' practices of peer assessment in taught courses across different field of studies. From our literature search, we discovered peer assessment can be conducted in combination of 44 different variation across different subjects/fields (Topping, 2021). From our preliminary findings, we also discovered that peer assessment can be used together with other pedagogical method for a more engaging form of assessment. Our future goal for this research is to lay an empirical supported ground for designing for enhancing applications (apps) for wider use of peer assessment in multidisciplinary settings.

### Acknowledgements

We would like to thank CADE-Lead (Centre for Academic Development and Leadership Excellence) and Universiti Putra Malaysia for providing support and funding for our work through GIPP (Teaching and Learning Initiative Grant) grant.

# References

- Astin, A. W., & Antonio, A. L. (2012). Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education (2nd. ed.). American Council on Education.
- Bruffee K. A. (1993). Collaborative learning: Higher education, interdependence, and the authority of knowledge. John Hopkins University Press.
- Ferrari, A., Spoletini, P., Bano, M., & Zowghi, D. (2019). Learning Requirements Elicitation Interviews with Role-Playing, Self-Assessment, and Peer-Review. In 2019 IEEE 27th International Requirements Engineering Conference (RE) (pp. 28-39). doi: 10.1109/RE.2019.00015.
- Fittipaldi, D. (2020). Managing the dynamics of group projects in higher education: Best practices suggested by empirical research. *Universal Journal of Educational Research*, *8*(5), 1778 1796.
- Goold, A., Craig, A., Coldwell, J., Goold, A., Craig, A. & Coldwell, J. (2008). The student experience of working in teams online. In *Proceedings of ASCILITE - Australian Society for Computers in Learning in Tertiary Education Annual Conference 2008* (pp. 343-352). Australasian Society for Computers in Learning in Tertiary Education.

Hassanien, A. (2006). Student experience of group work and group assessment in higher education. *Journal of Teaching in Travel & Tourism, 6*(1), 17 – 39.

Ingram, A.L. & Hathorn, L.G. (2004). Methods for analyzing collaboration in online communications. In T. Roberts (Ed.). *Online collaborative learning: Theory and practice* (pp. 215-241). Information Science Publishing.

- Johnson, D.W., Johnson, R.T., Stanne, M.B. & Garibaldi, A. (1990). Impact of group processing on achievement in cooperative groups. *Journal of Social Psychology*, *130*(4), 507-516. https://doi.org/10.1080/00224545.1990.9924613
- Kane, J. S., & Lawler, E. E. (1978). Methods of peer assessment. Psychological Bulletin, 85(3), 555 586. https://doi.org/10.1037/0033-2909.85.3.555
- LaBeouf, J. Griffin, J. & Roberts, D. (2016). Faculty and student with group work: What is problematic with college group assignments and why? *Journal of Education and Human Development*, 5(1), 13-23.
- Murray, J. (2003). *New Teacher Educators' Needs: Perspectives from Research and Practice.* Teacher Training Agency.
- Panitz, T. (1996). A definition of collaborative vs cooperative learning. Retrieved September 2, 2005, from London Guildhall University, Deliberations on Learning and Teaching in Higher Education: Collaborative Learning Web site: http://www.lgu.ac.uk/deliberations/collab.learning/panitz2.html
- Piezon, S., & Ferree, W. (2008). Perceptions of social loafing in online learning groups: A study of Public University and U.S. Naval War College students. *The International Review of Research in Open and Distance Learning*, *9*(2). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/484 1034
- Quinn, M. and Hughes, J. (2007) *Quinn's Principles and Practice of Nurse Education* (5<sup>th</sup>. ed.). Nelson Thornes.
- Simionescu, S., Šuníková, D., & Kubincová, Z. (2017). Gamification of peer assessment in learning management system. In 2017 18th International Carpathian Control Conference (ICCC) (pp. 571-575). Sinaia, Romania. doi: 10.1109/CarpathianCC.2017.7970465.
- Topping, K. (1998). Peer assessment between students in colleges and universities. Review of Educational Research, 68(3), 249 276. https://doi.org/10.2307/1170598
- Topping, K. (2021). Peer assessment: Channels of operation. *Education Sciences, 11*(91), 1 20. https://doi.org/ 10.3390/educsci11030091.
- Tumpa, R. J., Skaik, S., Ham, M., & Chaudhry, G. (2022). A holistic overview of studies to improve group-based assessments in higher education: A systematic literature review. *Sustainability*, 14(9638), 1 – 23. https://doi.org/10.3390/su14159638
- Winterbotham et al (2018) Department for Education Employer Skills Survey Report 2017. https:// assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\_data/file/7464 93/ESS\_2017\_UK\_Report\_Contr olled\_v06.00. pdf
- Zach, L. & Agosto, D. (2009). Using the online learning environment to develop real-life collaboration and knowledge-sharing skills: A theoretical discussion and framework for online course design. *Journal of Online Learning and Teaching, 5*(4), 590-599.

# The Effect of Enhancing Students' Learning Motivation on Bilingual Class with Integrating Augmented Reality in a Multimodality Learning Environment

# Yi-Fang LO<sup>a</sup>, & Jerry Chih-Yuan SUN<sup>a\*</sup>

<sup>a</sup> Institute of Education, National Yang Ming Chiao Tung University, Taiwan \*jerrysun@nycu.edu.tw

**Abstract:** The purpose of this study was to explore the effect of enhancing students' learning motivation on bilingual integrative activities classes with augmented reality (AR) creative video. One hundred and six 5<sup>th</sup> grade elementary students filled in the motivation questionnaire with Likert 5-point scale after the bilingual classes were conducted, those assigned to the non-augmented reality (AR) group, and an individual AR group. For the non-AR group, students discussed the outdoor activities orally with the traditional teaching method, and the individual AR group used group work with the "AR Makr" APP to make outdoor activities videos which were created on iPad. The results showed that integrating AR enhances students' learning motivation and self-efficacy in bilingual classes in a multimodality learning environment. It is suggested that teachers can incorporate multimodalities instruction in bilingual and language teaching in order to enhance students' learning motivation on language.

Keywords: Bilingual class, augmented reality, learning motivation

### 1. Introduction

### 1.1 Bilingual Policy and Teaching in Taiwan

Recently, Taiwan has made plans to implement a bilingual language policy by incorporating English as the main foreign language. The National Development Council's "Developing Taiwan into a Bilingual Nation by 2030" serves as a blueprint for the entire education system in Taiwan. Therefore, elementary schools are attempting to integrate different subjects and teach them in English (Alessandra & Lin, 2021). The pressure for students to learn subjects in a foreign language has not decreased. A multimodality enhanced pedagogical approach contributes to the development of foreign language learners' knowledge. Song and Ng (2023) suggest that instructional multimodality design can enable students to participate in teaching activities and develop linguistic competence, which contributes to the connection between subject content and bilingual language support. Instructors can create multimodality tasks to activate students' engagement in bilingual classes and encourage them to construct their linguistic expressions. Naujok (2023) also indicated that multimodalities are used in different strategies and provide opportunities to engage students in translanguaging interaction, which is helpful for content meaning-making and language acquisition.

#### 1.2 "Internet access in every class, and tablets for students." Policy

The Ministry of Education (MOE) in Taiwan has implemented a digital plan for elementary and junior high school students, which includes the provision of internet access in every class and tablets for students. This initiative aims to enhance digital learning and multimodalities in language teaching (Zhang & Wu, 2023). Many instructors aim to cultivate and integrate augmented reality (AR) techniques in their teaching designs based on the multimodality teaching approach (Wu, 2023).

Instructors integrate different kinds of applications on mobile devices to incorporate diverse resources into the multimodalities approach. In the present study, the "AR Makr" app, an iPad application, was used to create stories through a variety of templates and different tools (Dotson-Shupe & Mullins, 2022). ElSayary's (2022) research found that students can collaboratively create short movies with AR Makr as a positive engagement. The research also showed that teachers have higher motivation to highlight digital competency with course preparation. Asim et al. (2022) identified three domains to properly integrate technology into subject teaching: supporting the process of learning, catalyzing the acquisition of information, and communicating acquired knowledge. These domains are beneficial for implementing the findings of educational technology usage.

Therefore, e-learning has become an important teaching approach in elementary schools in Taiwan. E-learning provides students with autonomy and a multimodal environment, allowing teachers to act as mentors in their learning (Liaw et al., 2007). Harandi's (2015) study found that e-learning can positively affect higher education students' motivation. The researcher used a questionnaire to apply the results.

#### 1.3 English Learning through Augmented Reality in a Multimodality Learning Environment

The use of augmented reality (AR) technology in English learning has increased as a language learning tool, which is important for enhancing the learning experiences of the COVID-19 generation. Previous studies have shown that AR not only increases emotions and the immersive feeling in the learning environment, but it also enhances learners' motivation and cultivates the effects of English acquisition (Yu et al., 2019). In this study, the researcher conducted students' learning through AR experiences and English-speaking outputs during outdoor activities. By completing the AR task assigned by the researcher, students were able to enhance their self-regulation strategies and cooperate with team members. Yu et al. (2023) found that integrating AR tools into learning could enhance students' self-efficacy in subject learning.

In the multimodal learning types that Moreno and Mayer (2007) had integrated in their review, students could manipulate the objects through the process in AR environment. Rather than typical learning type which is non-interactive, in an multimodality learning environment, students can build their own representation through augmented reality (Mayer, 2001; Moreno & Mayer, 2007). Different tools to enhance multimodality learning environment are supposed to build up students' knowledge and linguistic acquisition.

#### 1.4 Aim of the Study

Building on Harandi's (2015) research, this study aims to investigate how integrating elearning through the AR Makr APP can enhance motivation among elementary students who are taking bilingual courses for the first semester.

### The main research question of this study is: **Does integrating AR technology** enhance motivation among students in bilingual classes?

### 1.5 Objective of the present study

The research aims to determine whether teaching bilingual classes through e-learning and AR integration in elementary education affects students' learning motivation. The purpose of the present study is to focus on the outcomes related to the motivation of students who participate in bilingual classes taught through e-learning and AR integration in elementary education.

### 2. Methods

### 2.1 Participants

The present study included 106 fifth-grade students from an elementary school in a remote county in the middle of Taiwan. The participants were selected using a sampling method with quasi-experimental design, which randomly assigned four classes to two groups: a non-AR group (n= 53) and an individual AR group (n= 53). The sample consisted of 56 male students (52.83%) and 50 female students (47.17%), as shown in Table 1.

Table 1.	Ratio of	male and	female	research	particip	bants
10010 1.	1 (01)0 01	maio ana	10111010	100001011	pairion	/4//100

Group	Non-AF	R group	Individual AR group		
	Number	%	Number	%	
Male	29	54.72	27	50.94	
Female	24	45.28	26	49.06	

### 2.2 Experimental Procedure

The research, which is conducted as a quasi-experimental design, consisted of two groups: a control group, non-AR group; and an experimental group, individual AR group. Each group attended bilingual classes for one semester consisting of 80 periods. The non-AR group was taught using traditional lecture and discussion methods, while the individual AR group was taught using e-learning through the AR Makr APP, which integrated designing and creative products. The individual AR group had 8 periods dedicated to this method out of the 80 periods.

Finally, the participants completed a questionnaire that included motivation scales for both groups. Additionally, the experimental group answered written open-ended questions to collect their motivational thoughts after completing the experiment.



*Figure 1.* Experimental Procedure

### 2.3 Instruments

In the experimental group, individual AR group, each participant was guided to learn about preparing for and executing outdoor activities, and then create a video of their outdoor activities using the AR Makr app, which showed the function and operation of augmented reality on Figure 2 and 3. Moreover, Figures 4 and 5 show two examples of creative camping preparation videos created using the AR Makr app. In contrast, the non-AR group did not receive instruction on outdoor activities using the AR Makr app.



Figure 2. The function on AR Makr App (1)



Figure 3. The function on AR Makr App (2)



Figure 4. The example of students' creative video on AR Makr APP (1)



Figure 5. The example of students' creative video on AR Makr APP (2)

The motivation scale questionnaire, which is a self-edited questionnaire by the researcher, was used to test the hypotheses for e-learning with bilingual classes and intrinsic motivation. It utilized a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The questionnaire's reliability was measured using Cronbach's alpha (.732). The data for the present study was collected and analyzed using the statistical analysis program, SPSS Statistic 22.

The instruments used in the study included two open-ended questions about the motivation for continuing to learn bilingual classes through e-learning. The questions were: "If you would like to use the AR Makr app to learn for the next semester, why?" and "If you want the instructor to design more iPad activities, what is the reason?" Individual AR group students were allowed to list as many responses as they wanted to make their answers clear. However, non-AR groups were not asked to answer the open-ended questions.

# 3. Results

### 3.1 Motivation with bilingual classes through e-learning

The findings on Table 2 indicate positive motivation with e-learning through AR integration in the individual AR group. The results highlighted that students had higher motivation to continue with e-learning activities for the new stage of learning. Additionally, the efficacy of the activities in the experiment was also satisfactory for learning.

 Table 2. Motivation of e-learning through AR Makr APP (individual AR group)

Items in Motivation Scale	Ν	М	SD	SD Error Mean
Would you like to integrate iPad	53	4.8870	.4234	.0582
activities for bilingual classes				
Do you like the AR Makr activities?	53	4.2698	.6381	.0876
Would you like to have iPad learning	53	4.9250	.3310	.0455
activities for next semester?				

### 3.2 Motivation with bilingual classes between non-AR group and individual AR group

According to Table 3, the mean motivation scale of the experimental group, which received individual AR, was higher than that of the controlled group, which did not receive AR. This suggests that learning bilingual classes with e-learning is more effective in enhancing motivation compared to traditional teaching methods. However, there was no significant difference between the two groups in response to the second question *"Would you like traditional teaching rather than bilingual classes?"* (*M*=2.283, *SD*= .1690), which was the reverse question.

Table 3. Results of Motivation	Questionnaire in bilingual	classes through	and throughout
integrating augmented reality			

Dimensions	Number of Items	Ν	М	SD	SD Error Mean
Motivation with bilingual classes in individual AR group	5	53	4.0063	.58100	.07981
Motivation with bilingual classes in non-AR group	5	53	3.5000	.05119	.37268

# 3.3 Students' thoughts with learning bilingual classes through e-learning and AR Makr App

The questionnaire for the experimental group contains two open-ended questions. The purpose of this unit is to analyze the students' comments.

• Why do you like to use AR Makr APP to learn for the next semester?

S16: "There will be multimodal resources available for learning, which will make it more interesting"

S22: "We can collaborate with our classmates and learn from their diverse talents."

S26: I can interact with the teacher more frequently, and it creates a more joyful learning environment in our class."

S35: "Learning with content is enjoyable."

S36: "AR learning is different from traditional classes that are stressful and pressurizing, and it motivates me to learn."

S51: "I always look forward to new bilingual classes."

S52: "AR allows us to explore different learning methods beyond traditional written work, and we can now learn in more active and creative ways."

In conclusion, using the AR Makr app to interact with classmates and teachers is a beneficial way to enhance and engage students in a positive manner, without the pressures of a traditional classroom.

Why do you want the instructor to design more AR Makr activities?

S1: "The combination of AR activities with my English learning this semester has made me excited for new learning opportunities in the upcoming semester."

S3: "The AR activities provide me with more opportunities to use my creativity in creating English content. Despite my limited English proficiency, the teacher challenges and encourages me to improve."

S19: "I find these types of classes to be more engaging and enjoyable. If given the opportunity, I would attend more of these classes."

S23: "I find the use of multimodal class and content to be very engaging in a bilingual learning environment."

S24: "I believe the teacher will create more exciting e-learning activities for us to engage in."

S44 "It is important that the teacher does not force us to learn, but instead motivates us in an energetic manner."

S49: "I love this type of learning because the classroom atmosphere is enjoyable and meaningful."

E-learning facilitates students to learn in different, energetic, and meaningful ways. Moreover, students have more opportunities to collaborate with classmates and utilize their diverse talents and intelligence while learning with bilingual content.

With e-learning, teachers act as promoters to facilitate students' autonomous learning, which has been found in research to provide more effective paths for student engagement in the class.

### 4. Discussion

The aim of this study is to investigate the effects of integrating augmented reality (AR) into bilingual classes and its impact on student motivation. To answer the research question, "Does integrating AR technology enhance motivation among students in bilingual classes?", a self-edited questionnaire was used to gather students' thoughts on bilingual classes and AR integration teaching. Harandi's (2015) research found that the AR Makr APP enhances students' language skills through augmented creation. Similarly, the present study showed that students had more opportunities to create an English environment meaningfully and energetically, enhancing their motivation for e-learning integration with bilingual classes.

Additionally, Yu, Liu, and Sun's (2023) research investigated the effects of integrating AR experiences and English speaking for college students. Similarly, this study found that elementary students' motivation towards English learning was raised through AR integration, with many students expressing positive thoughts on it.

### 5. Conclusion

The present study highlights the significant relationship between learning motivation and AR integration in the classroom through e-learning. The results specifically indicate that e-learning and AR integration can positively enhance students' motivation and learning engagement. Harandi's (2015) research on the relationship between e-learning and motivation in higher education suggests that the findings of this study can be applied to elementary education.

In summary, integrating AR Makr APP into teaching design can enhance students' different creative intelligence and talent, and even promote cooperation between classmates. The positive effects found in the qualitative data are reflected in students' responses. There is a stronger relationship between students' motivation, efficacy, and elearning compared to traditional classroom teaching. In summary, e-learning can have an appropriate educational effectiveness on students' bilingual learning and motivation.

However, the present study only investigated bilingual teaching through AR integration as one form of e-learning. Future studies can further develop these approaches. Since the researcher mainly focused on elementary education, we suggest future studies be conducted with more participants of different ages.

### References

- Alessandra, F. & Lin, T.-B. (2021). Official bilingualism in a multilingual nation: a study of the 2030 bilingual nation policy in Taiwan, *Journal of Multilingual and Multicultural Development*, 245-257. https://doi.org/10.1080/01434632.2021.1909054
- Asim, S., Ellis, J., Slykhuis, D. & Trumble, J. (2022). Three Domains for Technology Integration in Science Teacher Education. *Contemporary Issues in Technology and Teacher Education*, 22(4), 664-684. <u>https://doi.org/10.4018/978-1-61520-897-5.ch014</u>
- Dotson-Shupe, E., & Mullins, R. (2022). Moving a Step Beyond Technology Enhanced Traditionalism: Examining Practical Technology Applications in the College Classroom. 50 Ways to Use Technology Enhanced Learning in the Classroom: Practical strategies for teaching, 10-13. <u>https://doi.org/10.4135/9781529793550.n3</u>
- ElSayary, A., Mohebi, L., & Meda, L. (2022). The Impact of the Relationship of Social/Emotional, Cognitive, and Behavioral Engagements on Developing Preservice Teachers' Digital Competencies. *Journal of Information Technology Education: Research, 21,* 269-295. <u>https://doi.org/10.28945/4982</u>
- Harandi, S. R. (2015). Effects of e-learning on Students' Motivation. *Procedia-Social and Behavioral Sciences, 181*, 423-430. <u>https://doi.org/10.1016/j.sbspro.2015.04.905</u>
- Liaw, S. S., Huang, H. M., & Chen, G. D. (2007). An activity-theoretical approach to investigate learners' factors toward e-learning systems. *Computers in Human Behavior*, 23(4), 1906-1920.

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

- Moreno, R., Mayer, R. (2007). Interactive Multimodal Learning Environments. *Educ Psychol Rev, 19*, 309-326. <u>https://doi-org.ezproxy.lib.nctu.edu.tw/10.1007/s10648-007-9047-2</u>
- National Development Council. (2018). *Blueprint for Developing Taiwan into a Bilingual Nation by 2030.*

https://ws.ndc.gov.tw/Download.ashx?u=LzAwMS9hZG1pbmlzdHJhdG9yLzEwL3JlbGZ pbGUvMC8xMjE2Ny9hNGM4YWMwMS0zNDMyLTRhMDAtOGYwNy02NDExOWVjNW Q2ODgucGRm&n=MjAzMOmbmeiqnuWci%2BWutuaUv%2BetlueZvOWxleiXjeWcli5wZ GY%3D&icon=..pdf

Naujok, N. (2023). Participation in Storytelling Settings. Multimodality in Multilingual Contexts. *Zeitschrift für Interkulturellen Fremdsprachenunterricht, 28*(1).

- Song, H., & Ng, J. (2023, November 8-10). *Multimodality-enhanced teaching: Fostering global citizenship and intercultural competence in ELT.* TESL Ontario's 51st Annual Conference, Online. <u>http://contact.teslontario.org/wp-content/uploads/2023/03/Heejinand-Jacqueline.pdf</u>
- Yu, S.-J., Sun, J. C.-Y., & Chen, O. T.-C. (2019). Effect of AR-based online wearable guides on university students' situational interest and learning performance. *Universal Access in the Information Society, 18*(2), 287-299. <u>https://doi.org/10.1007/s10209-017-0591-3</u>
- Yu, S.-J, Liu, Y.-W, & Sun, J. C.-Y. (2023). Effects of Integrating Chatbots and Augmented Reality on University Students' Science Learning Self-Efficacy and Engagement: The Case of a Scientific Simulation Experiment. *Contemporary Journal of Science Education*, 31(1), 53-79. <u>https://doi.org/10.6173/CJSE.202303\_31(1).0003</u>
- Wu, J. G. (2023). Mobile Assisted Language Learning Across Educational Contexts. *Applied Linguistics*. <u>https://doi.org/10.1093/applin/amac071</u>
- Zhang, D., & Wu, Y. (2023). Becoming smart "digital natives": cultivating Chinese English majors' new media literacy via Journalism English Reading and Listening. *Journal of China Computer-Assisted Language Learning*. <u>https://doi.org/10.1515/jccall-2022-0029</u>

# The Course Learning Experience as a Forerunner in IT Internship: Assessment from the Student and the Host Training Establishment

Arlene Mae CELESTIAL-VALDERAMA<sup>a\*</sup> <sup>a</sup> Jose Rizal University, Mandaluyong City, Philippines \*arlene.valderama@jru.edu

Abstract: On commencing internship, undergraduates are required to perform core and professional skills under the supervision of external host training establishments. The variability of knowledge and skills experience can provide and identify placement in the Information Technology industry following graduation. This study presents intern assessments on how well the internship program supports development and demonstration of essential listed attributes. It explores perspectives on how well their course learning experience acts as a forerunner to develop skills/competencies. It seeks to validate how course categories from degree programs prepared them for training and which of the categories have been utilized and applied. This article includes the assessment of the interns from their immediate supervisors. The internship enabled the interns to develop or demonstrate essential attributes where 52% implied that they gained skills to apply in their future career, 35% appreciated the importance of continuing personal and professional development, and 10% inferred that the internship program was a substantial contribution in preparing them for professional career upon graduation. On the learning experience, more than 75% of the responses reflected seven listed skills and competencies. More so, 85% of the students have chosen the Internship course has prepared them substantially. Moreover, the course was highly utilized among other course categories as responded by 83% of the interns. The host training establishment response to the intern's performance has a high of 77%-92% in the range of 1-1.87 under the Excellent top box score. These results all contribute to the overwhelmingly positive perception of student intern performance.

Keywords: Internship, Course Learning Experience, Performance Assessment

# 1. Introduction

The internship program for every bachelor's degree is the most crucial supplemental training for students in practical skills, knowledge, and attitude necessary to adapt to the changing industrial world (Vo, Le, & Lam, 2021). The degree programs under the College of Computer Science and Engineering (CCSE) in Jose Rizal University prepares students through the reinstatement of its program educational objectives as 1) excel in their profession/career utilizing the knowledge acquired in the Information Technology courses, 2) become effective collaborators and innovators in information technology, applying professional/technical skills and competencies to make a positive impact on society; and 3) be engaged in life-long learning and professional development through self-study, continuing education, or graduate and professional studies. These objectives clearly state and aspire that a graduate must own the complete preparedness package of useful skills and competencies following program completion.

Every JRU student is entitled to the university's rich course learning experience as the it continues to reach for quality standards confirmations for its more than 100 years of

educational excellence, 3-star rating of excellence in QS Stars Rating System, CHED Autonomous Status, UKAS ISO Certified, and PACUCOA Level 4 accredited in some programs, among others. This paper intends to chronicle the perspectives of the students and the hose training establishments on how the course learning experience of an intern has become a forerunner in their on-the-job training experience. JRU decided on blended learning in 2007. For a few disciplines in the general education curriculum, the university has implemented the Course Redesign Program (CRP). A collaborative set of learning activities was prepared for the registered students from several general education courses, and the CRP eliminated face-to-face encounters in the classroom. Students like having the freedom to complete their online coursework and free-time tasks on or off campus. Blended learning has long been employed successfully by JRU students known as Rizalians, and as a result, scientists' learning outcomes were enhanced. At present, students from all levels relish the employment of the Canvas Learning Management System which was entirely beneficial during the pandemic.

The CCSE internship program is designed specifically to meet the students' training needs within the context of their work for IT-related responsibilities. An internship based on a strong model of intern recruitment and supervision, structured inclusion of interns in a supportive work culture, a detailed orientation, and a monitoring program through the Internship course was prepared for CCSE interns.

# 1.1 Research Objectives

This paper aims the following: 1) examine the students' perspectives on how their internship program has enabled them to develop and demonstrate essential skills and attributes, 2) propound how the university has prepared the students through their learning experiences in a set of defined skills or competencies, 3) present how the students defined course categories prepared them for the internship program, 4) depict students standpoint on what extent has their course types utilized or applied in the internship program, and lastly, 5) in response, convey how the students were evaluated by their host training establishments.

# 2. Research Methods

The setting of this study was in Jose Rizal University located in Mandaluyong City, Philippines. The target study population was students under the facilitation of one internship adviser through purposive sampling. The student interns are under the College of Computer Studies and Engineering (CCSE) who are pursuing the degrees of Bachelor of Science in Information Technology (BSIT) with some majors, Bachelor of Science in Business Administration Major in Computer Science (BSBA-CSc), Bachelor of Science in Computer Engineering (BSCpE), and Bachelor of Science in Entertainment and Multimedia Computing (BS EMC) with some majors. The internship course was taken in the months of January to May during the second semester in the senior year (4th year), following graduation. During the 2020-2021 academic year, under the supervision of the internship adviser, there are 14 students enrolled in the internship course. The 2020-2021 academic year holds 49 students and the academic year 2022-2023 has 48 students. Table 1 shows the demographics of the students and their courses.

SY & Internship	Degree Program	# of Students
Course Code		
2020/21 – ITC 53	BSIT	11
	BSCpE	2
	BSBA-CSc	1
2021/22 – EMC C402	BSIT	29
	EMC-DAT	19
	BSIT-AGD	1
2020/21 – ITC C402	BSIT	46
	BSIT-BA	2
	TOTAL	111

Table 1. Students in the CCSE and Respondents

All courses from the university employ the utilization of the Canvas Learning Management System. More so this has been realized since the pandemic as it has assisted all faculty members across all divisions of the institution to hold their classes synchronously and asynchronously. This is especially realized in the online environment and while this modality supports learning continuity and was evident that online classes have become a place for students to socially engage with others to meet various psychological needs to buffer pandemic stress (Ebardo and Suarez, 2023). The Canvas LMS course for the Internship course covered all submission assignment links per module on a weekly basis. Some are categorized submissions for pre-, upon, and post-deployment documents. Since academic year 2020/21, an Outcome-Based Education (OBE) Reflection Survey was required to all student interns through the Canvas LMS as an assignment type under the post-Deployment documents category. This was to be filled up by the students after their 486-hour internship (for BSIT program) from the host training establishment (HTE).

For academic year 2020/21, the interns were working in various IT industries nationwide and as far as Singapore as all were done through the work-from-home setup until the academic year 2021/22. As for 2022/23, the internship was done through hybrid mode. The virtual internship approach provided opportunities for large-scale distance learning institutions to routinely offer work experience to their students, supported by the further development of effective remote working practices (Reid, Butler, Comfort & Potter, 2023).

The OBE Reflection survey was created by the representatives from the HEI's academic council with internship facilitators and other experts. The questionnaire is divided into 4 parts with a total of 11 questions. The 4 parts of the questionnaire are shown in Table 2.

Parts	# of Questions	Туре
A. Course Matters	6	5 Qualitative
		1 Quantitative
B. Teacher Matters	1	Qualitative
C. Personal	7	Qualitative
D. Course & Program Evaluati	ion 4	Quantitative

 Table 2. OBE Reflection Survey Questionnaire

On this study, the first 4 research questions focus on the last part of the questionnaire, the Course & Program Evaluation. Jose Rizal University uses surveys because an effective Customer Experience Management (CEM) program necessitates the compendium, synthesis, exploration, and dissemination of customer metrics. When summarizing the responses to the survey, the indicators are calculated using the Top Box Score. This indicates the proportion of

the student responses who gave the highest rate on the scale (Vinluan, Mangaba, & Valderama, 2021).

The author used descriptive statistics to analyze the data. To simplify reporting of results, the questionnaire dichotomized the responses as follows: "excellent" for the Course and Program Evaluation part, question #8, #9, and #10, and "to a large extent for question #11".

# 3. Results and Discussion

# 3.1 Course & Program Evaluation – Development and Demonstration of Essential Attributes

The interns assessed how successfully the following vital qualities were developed or demonstrated during their internship program. Using the Top Box Score, 84.6% of the interns for the academic year 2020/21 interns responded "excellent" in the following 2 attributes: "Gain skills that I can apply in my future career" and "My Internship Program was a substantial contribution in preparing me for professional career upon graduation." For academic year 2021/22, 82.2% of the interns responded "excellent" attributing to "Appreciate the importance of continuing personal and professional development." And for the academic year 2022/23, 91.7% of the interns chose the attribute "Gain skills that I can apply in my future career" as their top score. A summary is shown in Table 3.

Table	3.	Internship	Program	Enablement	in	Developing	and	Demonstrating	Essential
Attribu	tes	Top Box So	core						

SY and Course Code	Attributes	# of Excellent Responses	Total Responses	% of Top Box Score Response	Total # E nrolled
SY 2020/21 ITC 53 OJT/Practicum	H. Gain skills that I can apply in my future career. N. My Internship Program was a substantial contribution in preparing me for professional career upon eraduation.	11	13	\$4.6	14
SY 2021/22 EMC C402 Internship	G. Appreciate the importance of continuing personal and professional development	37	45	82.2	49
SY 2022/23 ITC C402 Internship	H. Gain skills that I can apply in my future career.	44	48	91.7	48

# 3.2 JRU Learning Experience – Skills and Competencies

JRU uses surveys because a successful Customer Experience Management (CEM) program requires the collection, synthesis, investigation, and distribution of customer metrics. The JRU learning experience is one of the hallmarks as the institution values good education. It pays attention how the Rizalian learning experience is valuable according to a Rizalian himself. Another voice of the customer process is done through JRU Classroom Learning Experience (CLE) Survey and the JRU Customer Satisfaction Survey (CSS) (Vinluan, Mangaba, Valderama, 2021). The interns are no exception here. An item in the OBE Reflection Survey (question #9) entails responses to the top box score, which received the grade "excellent" due to its subsequent returns which are: a) the analysis is simplified, only 1 item is taken into account instead of 5 or more, b) contrasts are made quickly and easily, compare the results by variables. Top 2 Box scores permit for more effective comparisons of scores and c) trends become more visible as the survey is steered at the end of the semester, tracking the measurement during this time. Top 2 Box scores benefit in recognizing trends in the data. Question #9 asks about how well the JRU learning experience mitigated the interns develop the listed skills/competencies. About 76.9% of the interns in academic year 2020/21 believed that working well with others, applying problem solving strategies, acting responsibly with the

interests of the larger community in mind, connecting to others sensing and stimulating reactions and interactions, and proficiency at thinking beyond what is rote or rule-based are their excellent competencies acquired. About 34 of 45 or 75.6% of the interns presume that working well with others irrespective of age, gender, race, religion, or political persuasion is their key excellent amassed skill. Lastly, 36 of 48 or 75% of the interns reckon that working well with others, and the ability to work productively, drive engagement, and demonstrate presence as a member of a team are 2 of their obtained competencies. Table 4 shows a summary of the Top Box score for the JRU learning experience as a developmental skill and competency for the interns.

SY and Course Code	Skills / Competencies	# of Excellent Responses	Total Responses	% of Top Box Score Response
SY 2020/21 ITC 53 OJT/Practicum	<ul> <li>D. Work well with others irrespective of age, gender, race, religion or political persuasion</li> <li>G. Applying a range of strategies to problem solving</li> <li>N. Act responsibly with the interacts of the larger</li> </ul>	10	13 of 14	76.9
	interests       of       the       harger         community in mind       O.       Ability to connect to others         in a deep and direct way, to       sense and stimulate reactions         and desired interactions       P. Proficiency at thinking         and coming up with solutions         and responses beyond that         which is rote or rule-based			
SY 2021/22 EMC C402 Internship	D. Work well with others irrespective of age, gender, race, religion or political persuasion	34	45 of 49	75.6
SY 2022/23 ITC C402 Internship	<ul> <li>D. Work well with others irrespective of age, gender, race, religion or political persuasion</li> <li>T. Ability to work productively, drive engagement, and demonstrate presence as a member of a team</li> </ul>	36	48 of 48	75.0

Table 4. JRU Learning Experience as a Developmental Skill and Competency Top Box

# 3.3 Course Preparedness

An internship can be considered as one of the positive strategies for the schools to compete for a larger intake of students by promoting a comprehensive curriculum with an attractive internship program (Lam & Ching, 2007). The internship is defined as work placements, industry work experience, practicum, work integrated education. Davies (1990) states that internship is a kind of experiential learning where students take the opportunity to apply learned theories from schools in real-world situations, and it provides an opportunity for students to integrate and consolidate thinking and action. A well-rounded curriculum is however critical for the preparedness of the student to immersion in the internship. Question #10 in the OBE Reflection Survey focuses on how well the student was equipped in the course

categories finished in the past 3 years as training groundwork of concepts and applications for the Internship Program.

The course categories are listed as Languages, Mathematics, History & Social Sciences, Professional courses, Capstone, Elective, and the Internship (OJT). Table 5 exhibits that the Internship course, 76.9%, 80%, and 91.7% respectively for the three (3) school years has the most answered for the Top Box among the interns for the three (3) academic school years. This is an average of 83.8% or 90 of 106 interns. The second Top Box score is displayed due to this. Courses related to Languages (53.8%) and Professional courses (66.7% or 62 of 93 interns) are second most chosen as second Top Box score ("Excellent") for the three (3) academic years.

SY and Course	Courses	# of	Total	% of Top
Code		Responses	Responses	Response
SY 2020/21	(1) Internship	10	13 of 14	76.9
ITC 53	(2) Language	7		53.8
OJT/Practicum	S			
SY 2021/22	(1) Internship	36	45 of 49	80.0
EMC C402	(2) Professio	29		64.4
Internship	nal			
	courses			
SY 2022/23	(1) Internship	44	48 of 48	91.7
ITC C402	(2) Professio	33		68.8
Internship	nal			
	courses			

### 3.4 Utilization and Application of Courses in the Internship

A broad and well-balanced degree program increases the employability of students. It requires inclusivity if courses that hone and provide skills that can be applied in the future workplace, and that internships strengthen these skills and broaden the scope of non-academic studies (Stamati & Wilmott, 2023).

Zopiatis (2007) defined that internships provide students with the opportunity to apply classroom theory to practical problems in a significant business environment and, most importantly, ensure that their career choices are consistent with their interests and personality. Internships are especially helpful for resourceful students who want to gain some practical experience about the work by working in a bona fide business environment.

It is equally vital to cross-check and validate curriculum courses during scheduled curriculum reviews whether the general, core, and professional courses adapt well to a disciplinal competencies and values of a Rizalian graduate. Internships also provide the students with the necessary skills to be effective in the work environment (Chinomona & Surujlal, 2012). It gives students the opportunity to put their theoretical knowledge into practice and further develop their skills and experience in the industry (Bogdana, Mihaela, & Ancalona, 2012). Allowing the intern to self-evaluate whether the courses in their degree program is the focal point of Question #11. They were asked about the extent to which their courses were utilized and applied to their internship. All intern-respondents from the 3-year period chose the Internship course as the Top Box score, the highest Likert scale indicated that this course has helped them utilize and apply their courses "to a large extent." The internship course poses the following course learning objectives (CLO): 1) Discuss the importance of finding the right company for their on-the-job training or practicum to learn from a real workplace setting and remind them of the duties and responsibilities of an on-the-job trainee, 2) Demonstrate logical and analytical thinking skills, communications and social skills, and IT/EMC related skills needed to perform the assigned tasks in their on-the-job training/practicum, and 3) Develop the proper work ethics of an Information Technology professional and continue to improve

themselves for their professional development. Through teaching and learning activities, the intern adviser/facilitator was keen in the proper dissemination of every procedure relating to the CLOs.

83% or 88 of 106 of the interns presumed that the internship course sustained them "to a large extent" throughout their internship in the Host Training Establishment. In this regard the second Top Box score was retrieved among the course types. 61.5%, 64.4%, and 62.5% responded that the courses on Languages, Professional courses, and Capstone course respectively have been utilized in their internship "to a large extent" as well. Table 6 shows a summary.

SY and Course Code	Courses	# of Large Extent	Total Responses	% of Top Box Score	Total # Enrolled
		Responses		Response	
SY 2020/21	<ol><li>Internship</li></ol>	11	13	84.6	14
ITC 53	(2) Languages	8		61.5	
OJT/Practicum					
SY 2021/22	<ol><li>Internship</li></ol>	36	45	80.0	49
EMC C402	(2) Professional	29		64.4	
Internship	courses				
SY 2022/23	<ol><li>Internship</li></ol>	41	48	85.4	48
ITC C402	(2) Capstone	30		62.5	
Internship					

 Table 6. Course Utilization in Internship

# 3.5 Performance Evaluation from Host Training Establishment

The number of hours in the different degrees vary. In the completion of the 486 hours of internship, at least for the BSIT program, the interns were evaluated by their host training establishments (HTE), specifically their immediate supervisors using the University's ISO labeled form Internship Performance Evaluation. This contains an 8-item indicator list with a 5-point Likert Scale of excellent, very good, good, poor, needs improvement. The interns' immediate supervisor shall rate them through relations with others, judgment, ability to learn, behavior, dependability, quality of work, attendance & punctuality, and over-all performance as compared to other interns. Part of the form were 3 qualitative questions and 1 comment area asking the strengths and weaknesses of the intern and given the sole decision to hire, would they hire them in the organization. To realize the results of this, the researchers ought to show the evaluations of the interns with less than 2 in a 5-point Likert scale, or those with grades in lines of 1.0.

Averagely, using the highest score of 1.0 as Excellent, the evaluator has scored each item and Table 7 presents the total average score of each item of every intern from grades 1.00 to 1.875. Each set of interns from every school year has gained 77%, 92%, and 83% respectively. This depicts a high turnout of the performance of the interns as evaluated by the immediate supervisor from the host training establishments.

The data has shown high impact that the interns have exhibited their work tasks properly. To note, the lowest grades for the interns from all school years are 2.75, 2.5, and 3.1 respectively, illustrating passing grades on 3 as Good in the 5-Point Likert scale. The internships of the 106 students have been overwhelmingly successful for the interns and supervisors as well.

SY and Course Code	# of Interns with Overall Grade from 1.00-1.88	%	# of Interns with perfect 1.00
SY 2020/21 ITC 53 OJT/Practicum	10 of 13	76.9	5
SY 2021/22 EMC C402 Internship	41 of 45	91.9	13
SY 2022/23 ITC C402 Internship	40 of 48	83.3	7

Table 7. Total Students Evaluation Grade Range

### 4. Conclusions

The original aims of this paper is to investigate the students' viewpoints on how their internship program has helped them acquire and exhibit essential skills, validate how their university learning experience has prepared them for the internship through a list of defined skills or competencies, check how the degree courses taken have prepared them and its utilization in the internship program, and lastly, present how they were evaluated by the host training establishments.

From the foregoing results and discussion, the following conclusions are drawn. The interns' response on the area of Course and Program Evaluation, specifically on how well the Internship Program has enabled them to develop or demonstrate essential attributes reflects that 52% implied that they gained skills to apply in their future career, the other 35% inferred they appreciated the importance of continuing personal and professional development, the last 10% denoted that the Internship Program was a substantial contribution in preparing them for professional career upon graduation.

Educating and maintaining potential qualified employees in the workplace is becoming one of the crucial factors for the sustainable development strategy of institutions and organizations (Lam & Ching, 2007). The interns are potential industry practitioners and employees. Their sustainable learning experience may never be repeated in their course of age, which is why their courses, and the internship program must strategically prepare them for future employment. Across the stream of the JRU learning experience, more than 75% of the responses reflected 7 listed skills and competencies shown in Table 4.

The course categories for each degree program only varies on their professional courses, common general education and core courses in the JRU degree programs have identifiably contain uniformity. The respondents in this study have chosen the Internship course has prepared them subsequently. More so it was highly utilized among other course categories.

The host training establishment response to the intern's performance is overridingly positive. A high of 77%-92% of them has rated the interns in the range of 1-1.87 under the Excellent Likert scale vis a vis the performance indicators. This result all contributes to the overwhelmingly positive perception of intern performance.

### References

Anderson, R. E. (1992). Social impacts of computing: Codes of professional ethics. *Social Science Computing Review*, *10*(2), 453-469.

- Nga Thi Vo, Linh Hoang Phuong Le & Van Thi Thanh Lam (2022). Challenges for Student Satisfaction of Internship Program in Hospitality and Tourism Industry in Vietnam, Journal of Quality Assurance in Hospitality & Tourism, 23:5, 1298-1324, DOI: 10.1080/1528008X.2021.1964414
- Ebardo, R. and Suarez, M. T. (2023). "Do Cognitive, Affective and Social Needs Influence Mobile Learning Adoption in Emergency Remote Teaching?" Res. Pract. Technol. Enhanc. Learn., vol. 18.
- Reid, K., Butler, D. L., Comfort, C. & Potter, A. (2023). Virtual internships in open and distance learning contexts: Improving access, participation, and success for underrepresented students, Distance Education, 44:2, 267-283, DOI: 10.1080/01587919.2023.2209029

- Vinluan, A.A., Mangaba, J.B., Celestial-Valderama, A.M.: Prelude to Full Online Learning: Educational Interventions from the Voice of the Customers. (2021) In: Rodrigo, M. M. T. et al. (Eds.) Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education, vol. 2, pp. 387–396
- Lam, T., & Ching, L. (2007). An exploratory study of an internship program The case of Hongkong students. International of Hospitality Management, 26(2), 336–351. https://doi.org/10.1016/j.ijhm.2006.01.001
- Davies, L. (1990). Experience-based learning within the curriculum: A synthesis study. Community Nurse Audio Metrists Association Press.
- Stamati, K. & Willmott, L. (2023). Preparing UK PhD Students Towards Employability: A Social Science Internship Programme to Enhance Workplace Skills, Journal of Further and Higher Education, 47:2, 151-166, DOI: 10.1080/0309877X.2022.2102411
- Zopiatis, A. (2007), Hospitality Internships in Cyprus: A Genuine Academic Experience or A Continuing Frustration? International Journal of Contemporary Hospitality Management, 19(1), 65-77. //doi.org/10.https://doi.org/10.1080/15313220.2012.72945909596110710724170/095961107107 24170
- Chinomona, R., & Surujlal, J. (2012). The Influence of Student Internship Work Experience on their Self-Improvement and Professionalism in Sport Management. African Journal for Physical, Health Education, Recreation and Dance (AJPHERD), 18(2), 885–899.
- Bogdana, A. B., Mihaela, L., & Anca-Iona, M. (2012). Internship Roles in Training and Professional Development of Student. Annals of Faculty of Economics, 1(1), 986–991.

# Examining Different Affective Factors in Learning with Virtual Reality

### Hsing-Ying TU<sup>a</sup>, Silvia Wen-Yu LEE<sup>a\*</sup> & Ting-Yueh HSU<sup>a</sup>

<sup>a</sup>Graduate Institute of Information and Computer Education, National Taiwan Normal University, Taiwan

### \* swylee@ntnu.edu.tw

**Abstract:** This study aims to examine how prior knowledge and affective factors of virtual reality environments predict science learning achievement through the mediation of learning engagement. Ninety-two sixth-grade students in Taiwan were recruited in this study. Data were analyzed through partial least squares structural equation modeling (PLS-SEM). The results showed that prior knowledge negatively predicted presence and control and active learning. Presence, control and active learning positively predicted learning engagement (behavioral engagement, cognitive engagement, emotional engagement). Cognitive fatigue was found to negatively predict emotional engagement and science learning achievement. Implications and suggestions for future research were addressed in the study.

**Keywords:** presence, cognitive fatigue, control and active learning, learning engagement, virtual reality

# 1. Introduction

In the modern education landscape, digital learning platforms and advanced technological tools have completely transformed the nature of learning. With the continuous evolution of instructional technology, virtual reality (VR) has become a part of innovative educational approaches, the sense of realism in the learning environment has continuously improved and offers students unprecedented levels of engagement and interaction. By exploring, manipulating, and experimenting within a virtual environment, learners can obtain immersive learning experiences (Radianti et al., 2020).

In the VR learning environment, students not only have the ability to freely explore and experience in a VR environment but also have the freedom to control the pace and content of their learning. Previous research has shown that if students have a higher degree of control in learning, they also achieve better learning outcomes and satisfaction (Jang et al., 2017; Lee et al., 2010). Nevertheless, prolonged exposure to digital learning environments, especially when engaging in learning within virtual contexts, may lead to an overwhelming cognitive load, potentially negatively impacting learning outcomes (Parong & Mayer, 2018). Studies have also shown that users have reported experiencing cognitive fatigue after prolonged use (Cummings & Bailenson, 2016; Munafo et al., 2017). However, there is still a lack of research regarding the effect of cognitive fatigue and its relationship with other factors in the context of VR learning environments.

Makransky and Petersen (2021) have developed the cognitive affective model of immersive learning (CAMIL) illustrating the relationships among VR affordances, affective and cognitive factors, and learning outcomes. Among the different research directions, it was suggested that future study can investigate the impact of external factors on personal traits or dispositions. Therefore, the research purpose of this study is to explore the effects of learners' prior knowledge and other affective factors of VR environments on their learning engagement and learning outcomes in the VR learning environment. A model was developed describing the hypothesized relationship among the variables used in this study based on the literature mentioned above (see Figure 1). Prior knowledge is hypothesized to positively predict students' affective factors of VR environments (Kim et al., 2021). Among

the affective factors, presence and control and active learning are hypothesized to be positively predict learning engagement and learning outcomes (Lee et al., 2010; Purarjomandlangrudi & Chen, 2020). On the contrary, cognitive fatigue is hypothesized to be negatively predict learning engagement and learning outcomes (Hwang et al., 2019). Learning engagement is hypothesized to be positively predict learning outcomes (Liu et al., 2022).



*Figure 1*. The hypothesized model regarding relationships among prior knowledge, affective factors of VR environments, learning engagement, and learning outcomes.

# 1.1 Affective Factors of VR environments

In the following, the effects of the affective factors in VR environments were discussed, including presence, cognitive fatigue, and control and active learning.

# 1.1.1 Presence

Presence is the sense of being in one place, which is a psychological state or subjective perception in which even though part of or all of an individual's current experience is generated by the system (Lee et al., 2010). In VR, presence acts as an affordance, leading to deeply immersive experiences (Makransky & Petersen, 2021).

Empirical research has highlighted the relationship between a user's prior knowledge and their sense of presence in virtual environments. For instance, participants with higher levels of prior knowledge related to the content of a virtual environment reported a deeper sense of immersion and spatial presence (Kim et al., 2021). Purarjomandlangrudi and Chen (2020) examined students' sense of presence in a virtual classroom and its subsequent influence on their engagement levels. The findings demonstrated that students who reported a heightened sense of presence were significantly more engaged in their learning tasks than those who felt detached. Studies have also shown that students who perceived a higher presence in the virtual environment might have higher learning outcomes (Lee et al., 2010). However, studies have also shown that higher immersion leads to higher cognitive load (Parong & Mayer, 2018). Researchers have also concluded that students in high-immersion VR environments would have a higher sense of presence but less learning (Makransky et al., 2019).

# 1.1.2 Cognitive fatigue

Cognitive fatigue, also known as mental fatigue, refers to the decline in cognitive performance and efficiency after prolonged periods of cognitive activity. It manifests as reduced attention, slower reaction times, and increased errors (Boksem et al., 2005). Studies have shown that an increase of cognitive fatigue decreased cognitive control (Lorist et al., 2005) and high-level information processing (Tanaka et al., 2014). Researchers indicated that when the learning material required students' cognitive effort in a limited amount of time might lead to increasing cognitive fatigue, thereby affecting performance (Hwang et al., 2019).

### 1.1.3 Control and active learning

Control and active learning refer to leveraging VR technology could empower learners to actively engage with and shape their own learning experiences within a virtual environment

(Lee et al., 2010). Studies have shown that students with higher level of control and active learning might enhance their learning engagement, achievement, and retention of knowledge (Deslauriers et al., 2019; Freeman et al., 2014). Previous studies demonstrated that the level of control and active learning could lead to increased engagement, motivation, and learning outcomes among students (Makransky & Lilleholt, 2018).

# 1.2 Learning engagement

Learning engagement refers to the dynamic and multifaceted involvement of students in the learning process (Furrer & Skinner, 2003). Cognitive, emotional, and behavioral dimensions are typically regarded as being a part of learning engagement (Fredricks et al., 2004). *Cognitive engagement* refers to the mental effort, concentration, and active participation that students invest in learning tasks and academic activities. The observable actions and behaviors that students exhibit in the learning process, such as participation in classroom activities or putting effort in completing assignments, is referred to as *behavioral engagement*. *Emotional engagement* refers to the affective or emotional reaction of a student's involvement in learning, including different emotions, such as enjoyment or boredom.

Previous studies discovered that engagement in school-related tasks, including cognitive and behavioral engagement, predicts academic achievement and educational attainment over time (Fredricks et al., 2004; Wang & Eccles, 2013). In the context of VR learning environment, students were found to demonstrated higher learning achievement and learning engagement (Akman & Çakır, 2023).

# 2. Method

# 2.1 The VR learning materials

This learning material was designed by our research team for elementary school students between 5th to 6th grades to learn concepts related to water, including water in the atmosphere, the structure of plants, the lives of animals, and supplementary knowledge of science. With the guidance of a virtual agent and the visualization of microscopic objects and natural phenomena, students could learn water-related knowledge by exploring the VR learning material.

The VR learning material consisted of five scenes, including pre-training and four different seasons. Students become familiar with the operation and have a preliminary understanding of the content of this learning material by interacting with content in the pre-training scene. A self-evaluation system was also embedded in the VR learning material. After students finished learning science knowledge of each scene, there would be questions for them to practice. Students could check their learning status with immediate feedback and comments after they answered each question.

# 2.2 Participants and procedure

Data were collected from 92 6th-grade elementary students with ages between 11 to 12 in Taiwan. Among these students, there were 35 males (38%) and 57 females (62%). Before the experiment, students were acknowledged the purpose of the experiment and only the students who volunteered to participate were recruited. Students first completed a test of science prior knowledge to assess their prior knowledge of the water-related concepts before experiencing the VR learning activity for a maximum of 30 minutes. Subsequently, students were asked to complete a science post-test and post-survey to measure their presence, cognitive fatigue, control and active learning, and learning engagement.

# 2.3 Instruments

Four questionnaires were designed to measure students' presence, cognitive fatigue, control and active learning, and learning engagement. A test of prior knowledge and a science learning achievement test were developed to test students' understanding of the concepts related to water.

# 2.3.1 Presence questionnaire

Presence questionnaire, adopted from Schubert et al. (2001), consisted of *spatial presence* (*four items*), *involvement (three items*), and *experienced realness (one item*). Spatial presence, defined as the perception that the user's body is actually located in the virtual space, includes items such as "I felt present in the virtual space." The items in involvement measured how much attention the VR learning activity draws to the user, and how much the user still pays attention to the real world, sample item includes "I concentrated only on the virtual space." Experienced realness was defined as the realness the user felt between VR learning activity and reality, including items such as "My experience in the virtual environment is the same as my experience in the real world." Participants responded on a Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

# 2.3.2 Cognitive fatigue scale

Cognitive fatigue was defined as decreased cognitive resources and cognitive function over time due to sustained cognitive demands (Trejo et al., 2005). Cognitive fatigue scale, adapted from Hwang et al. (2019), consisted of four items in this study, and the sample item includes "My concentration would disappear very quickly when I experienced the VR learning activity." Participants responded on a Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Past study has reported good reliability of the scale ( $\alpha = 0.94$ ) (Hwang et al., 2019).

# 2.3.3 Control and active learning scale

Control and active learning was defined as the level of autonomy learners were allowed in the learning environment (Lee et al., 2022). Control and active learning scale, adopted from Lee et al. (2022), consisted of five items. Participants responded on a Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Sample item of the scale includes "This type of VR learning activity allows me to have more control over my own learning." Past study has reported good reliability of the scale ( $\alpha = 0.88$ ) (Lee et al., 2010).

# 2.3.4 Learning engagement scale

Learning engagement scale, adopted from Lee et al. (2021), comprises three subscales: *behavioral engagement* (five items), *cognitive engagement* (four items), *and emotional engagement* (four items). Although the original scale encompassed social engagement, the sub-scale was not included since there was no peer interaction in this study. Behavioral engagement, defined as the behavior related to academic achievement, includes items such as "I keep trying even if the learning activity is hard." Cognitive engagement was defined as the effort, including meta-cognition and self-regulation, to understand learning content. A sample item of cognitive engagement includes "I think about different ways to solve a problem." Emotional engagement was defined as the emotional responses to learning activities, including interest, enjoyment, and perceived value of learning. Sample item of emotional engagement includes "I feel good when I am doing this learning activity." Participants responded on a Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Past study has reported good reliabilities of the factors (behavioral engagement:  $\alpha = 0.85$ ; cognitive engagement:  $\alpha = 0.88$ ; emotional engagement:  $\alpha = 0.90$ ) (Lee et al., 2021).
#### 2.3.5 Science achievement test

The multiple-choice assessments were designed based on the content in the VR learning material to evaluate students' understanding of the concepts related to water, such as the different states of water, the formation of frost and fog, etc. The science prior knowledge test (8 items) and learning achievement test (13 items) were used to measure students' prior knowledge and their learning outcomes after learning through the VR learning material.

#### 2.4 Data analysis

In this study, partial least squares structural equation modeling (PLS-SEM) was used to analyze the data. PLS-SEM, considered to be the second generation of multivariate analysis for verifying a relationship between variables, is suitable for analyzing small sample sizes of data and does not need normal data distribution (Hair et al., 2021).

In this study, presence was treated as a formative construct due to the independent contribution of the items, while other variables were considered as reflective constructs. In a reflective construct, the latent variable causes the observed indicators, which means that any change in the construct would result in changes in its indicators. On the contrary, the observed indicators cause or form the latent variable is a formative construct, which means that the construct is determined by its indicators. Unlike reflective constructs, the indicators in formative construct do not necessarily have to be correlated, and they each contribute uniquely to the formative construct. PLS-SEM also has the advantage to contain both reflective and formative constructs in a model (Hair et al., 2021).

The evaluation of PLS-SEM begins with the measurement model where each indicator's factor loading of reflective constructs should ideally surpass 0.7 for reliability. Additionally, the composite reliability (CR) of constructs should exceed 0.7, while their average variance extracted (AVE) should be above 0.5, ensuring convergent validity. For discriminant validity, the square root of a construct's AVE should be greater than its highest correlation with any other construct, with methods like the Fornell-Larcker criterion often used for further validation. The formative constructs were evaluated by assessing variance inflation factor values (VIF), and outer weights.

Transitioning to the structural model, the coefficient of determination ( $R^2$ ) is scrutinized, where values above 0.75 indicate strong explanatory power. Path coefficients are crucial, with significance denoting impactful relationships among constructs. Effect sizes, represented as  $f^2$ , should ideally exhibit values of 0.02, 0.15, or 0.35 for small, medium, or large effects, respectively. Predictive relevance, measured by  $Q^2$ , becomes vital, with values greater than zero signifying model relevance.

# 3. Results

# 3.1 Measurement model

The quality of the measurement model was examined by construct reliabilities and construct validities. The Cronbach's alpha values and the composite reliability (CR) values were tested to verify the internal consistency of the indicators of each construct (Hair et al., 2021). As shown in Table 1, Cronbach's alpha values of the constructs were between 0.86 to 0.98, which were above the suggested value of 0.70. Additionally, the CR values of the constructs were between 0.91 to 0.98, which also met the requirement of being greater than 0.70. These results show that the measurement model had sufficient reliability and the internal consistency of the indicators for each construct was good.

Convergent validity and discriminant validity were assessed to verify whether the measurements effectively reflected the corresponding measured constructs. Factor loadings of indicators and the average variance extracted (AVE) of constructs were used to validate the convergent validity of the measurements. The factor loadings of the individual items and the AVE values of the constructs were all above 0.7, which is higher than the suggested

value (Hair et al., 2021). BEng 1 was deleted since the value of its factor loadings was below 0.7. These results showed adequate convergent validity.

	Mean	SD	Factor loadings	Cronbach's alpha	CR	AVE
Cognitive Fatigue (CF)				0.97	0.97	0.91
CF 1	2.34	1.23	0.95			
CF 2	2.47	1.24	0.93			
CF 3	2.43	1.17	0.97			
CF 4	2.42	1.28	0.95			
Control and Active Learning (CAL)				0.98	0.98	0.91
CAL 1	3.95	0.86	0.96			
CAL 2	3.90	0.87	0.95			
CAL 3	3.93	0.82	0.97			
CAL 4	3.88	0.87	0.95			
CAL 5	4.02	0.86	0.93			
Behavioral Engagement (BEng)				0.86	0.91	0.71
BEng 2	3.84	0.86	0.82			
BEng 3	3.97	0.89	0.90			
BEng 4	3.85	0.94	0.90			
BEng 5	4.11	0.81	0.74			
Cognitive Engagement (CEng)				0.89	0.92	0.74
CEng 1	4.02	0.85	0.88			
CEng 2	3.91	0.84	0.88			
CEng 3	3.91	0.87	0.81			
CEng 4	3.86	0.88	0.88			
Emotional Engagement (EEng)				0.90	0.93	0.78
EEng 1	4.17	0.90	0.90			
EEng 2	3.76	0.97	0.81			
EEng 3	3.95	0.90	0.91			
EEng 4	4.20	0.86	0.90			

 Table 1. Confirmatory Factor Analyses and Reliabilities of Cognitive Fatigue, Control and

 Active Learning, and Learning Engagement.

The Fornell-Larcker criterion and cross loadings were measured to verify discriminant validity, which indicates the degree to which each construct in the resulting model is distinct from the others (Hair et al., 2021). The cross loadings of measurement variables are suggested to be higher than the related latent variable. The square root of the AVE value of each variable should also be higher than 0.5 and larger than the Pearson's correlation coefficient between the two variables. All cross loadings of the items were higher than each related latent variable. The AVE value of each variable (0.84 - 0.95) achieved the standard as well. In accordance with Hair et al. (2021), the results indicated that the discriminant validity of the variables was verified.

In this study, presence was considered as formative construct. By using a global single item for redundancy analysis, the convergent validity of formative constructs was evaluated by examining its correlation with an alternative measure of the construct. The result of redundancy analysis for convergent validity of presence was 0.85, which was above the suggested value of 0.8. Variance inflation factor (VIF), outer weights, and outer loadings were assessed to verify the collinearity, significance, and relevance of formative items. The acceptable collinearity and adequate construct validity were signified by VIF values less than 5, which indicated that an item's contribution to the primary latent construct was unique (Hair et al., 2021). The outer weight of an item determines its relative importance in formative construct (Hair et al., 2021). An item was kept in the measurement model if it had a significant outer weight (p < .05), or if its outer loadings was higher than 0.5. Items that did

not meet these criteria were further evaluated based on the significance of their outer loadings. An item was ultimately removed from the model if its outer loading was lower than 0.5 and not significant.

### 3.2 Structural model

PLS-SEM was used to test the hypotheses proposed in this study, which included the relationships among prior knowledge, affective factors of VR environments (presence, cognitive fatigue, control and active learning), learning engagement (behavioral engagement, cognitive engagement, emotional engagement), and learning outcomes (science learning achievement). The paths with statistical significance (p < .05) are shown in Figure 2. The results indicated that prior knowledge negatively predicted presence ( $\beta = -0.24$ , p < .05), control and active learning ( $\beta = -0.21$ , p < .05). Presence positively predicted behavioral engagement ( $\beta = 0.44$ , p < .001), cognitive engagement ( $\beta = 0.40$ , p < .001), and emotional engagement ( $\beta = -0.22$ , p < .001) and science learning achievement ( $\beta = -0.25$ , p < .05). Control and active learning positively predicted behavioral engagement ( $\beta = 0.44$ , p < .001), and emotional engagement ( $\beta = 0.44$ , p < .001) and science learning achievement ( $\beta = -0.25$ , p < .05). Control and active learning positively predicted behavioral engagement ( $\beta = 0.44$ , p < .001), and emotional engagement ( $\beta = 0.44$ , p < .001), and science learning achievement ( $\beta = 0.41$ , p < .001), cognitive engagement ( $\beta = 0.44$ , p < .001).



*Figure 2.* Structural Model Results of Prior Knowledge, Affective Factors of VR Environments, Learning Engagement, and Learning Outcomes (only significant paths are shown). (CF=Cognitive Fatigue, CAL=Control and Active Learning, BEng=Behavioral Engagement, CEng=Cognitive Engagement, EEng=Emotional Engagement)

# 3.3 Mediation

As shown in Table 2, control and active learning played significant mediating role between presence and learning engagement. Control and active learning mediated the relationship between presence and behavioral engagement ( $\beta$  = -0.09, *p* < .05), cognitive engagement ( $\beta$  = -0.09, *p* < .05), and emotional engagement ( $\beta$  = -0.09, *p* < .05).

Table 2. Mediation Analyses Results of the Hypothesized Model (only significant paths are shown).

Path	β	t	р
Prior knowledge $\rightarrow$ CAL $\rightarrow$ BEng	-0.09	2.02	.044*
Prior knowledge $\rightarrow$ CAL $\rightarrow$ CEng	-0.09	1.99	.046*
Prior knowledge $\rightarrow$ CAL $\rightarrow$ EEng	-0.09	2.09	.037*

\**p*<.05; CAL=Control and Active Learning; BEng=Behavioral Engagement; CEng=Cognitive Engagement; EEng=Emotional Engagement.

# 4. Discussion

According to the model analysis, it was found that science learning achievement was negatively predicted by cognitive fatigue. In other words, an increase in students' cognitive fatigue decreased their learning performances. Additionally, cognitive fatigue was found to negatively predict emotional engagement, which referred to the increase of students' cognitive fatigue would decrease their emotional response to learning activities. The information given in the learning material and the questions after exploring every scene, which required students' sustained cognitive demands, might gradually increase their cognitive fatigue. Moreover, students had to finish the post-test assessment right after experiencing the learning material. Students might not find the learning activities interesting and not willing to engage in it due to the reasons mentioned above. It is suggested that researchers could divide experiment into several learning stages, and give students time to rest between these stages in order to reduce their perception of cognitive fatigue, improve their learning performance and willingness of engaging in the learning activities.

On the contrary, learning engagement was found to be positively predicted by presence and control and active learning. These results showed similarities with the results of Purarjomandlangrudi and Chen (2020), which indicated that students with higher sense of presence might be more engaged in their learning activities than those who felt detached. However, the findings showed that students' prior knowledge could negatively predict their sense of presence and control and active learning, which showed differences from the results shown in Kim et al. (2021). It is suggested future studies could deepen the exploration of the relationships between students' prior knowledge and their sense of presence with different ages of students.

Furthermore, control and active learning was found to mediate the relationships between prior knowledge and learning engagement. In other words, learners with lower prior knowledge might not only perceive higher level of autonomy over their learning, but also increase their interactions with the learning material. It is recommended that in the future, such learning materials be provided to students with lower prior knowledge, as they can potentially derive greater benefits.

Finally, presence was treated as a formative construct in this study, while most of the studies treated presence as a reflective construct (e.g., Makransky et al., 2019). In future research, we suggest investigating presence as a formative construct, considering its potential to offer deeper insights into the dynamic interplay of its components. Additionally, contextual variations and qualitative inquiries could provide a comprehensive understanding of how learners experience presence in diverse educational settings, ultimately enriching the theoretical frameworks and pedagogical strategies.

# 5. Conclusion

A model was proposed in this study suggesting the relationships among prior knowledge, affective factors of VR environments (presence, cognitive fatigue, control and active learning), learning engagement (behavioral engagement, cognitive engagement, emotional engagement), and learning outcomes (science learning achievement). The effects of prior knowledge and cognitive fatigue were highlighted due to the differences from previous studies and the lack of empirical evidence. Finally, future studies are suggested to consider treating presence as a formative rather than a reflective construct.

# Acknowledgement

This research was supported by National Science and Technology Council, Taiwan, under grant number of MOST 108-2628-H-003-007- MY4.

#### References

- Akman, E., & Çakır, R. (2023). The effect of educational virtual reality game on primary school students' achievement and engagement in mathematics. *Interactive Learning Environments*, *31*(3), 1467-1484.
- Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: an ERP study. *Cognitive brain research*, *25*(1), 107-116.
- Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, *19*(2), 272-309.
- Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences*, *116*(39), 19251-19257.
- Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. *Learning and instruction*, *43*, 1-4.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the national academy of sciences*, *111*(23), 8410-8415.
- Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of educational psychology*, *95*(1), 148.
- Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). A primer on partial least squares structural equation modeling (*PLS-SEM*). Sage publications.
- Hwang, M. Y., Hong, J. C., Ye, J. H., Wu, Y. F., Tai, K. H., & Kiu, M. C. (2019). Practicing abductive reasoning: The correlations between cognitive factors and learning effects. *Computers & Education*, *138*, 33-45.
- Jang, S., Vitale, J. M., Jyung, R. W., & Black, J. B. (2017). Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. *Computers & Education*, *106*, 150-165.
- Kim, S. J., Laine, T. H., & Suk, H. J. (2021). Presence effects in virtual reality based on user characteristics: Attention, enjoyment, and memory. *Electronics*, 10(9), 1051.
- Lee, E. A. L., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education*, *55*(4), 1424-1442.
- Lee, S. W. Y., Hsu, Y. T., & Cheng, K. H. (2022). Do curious students learn more science in an immersive virtual reality environment? Exploring the impact of advance organizers and epistemic curiosity. *Computers & Education*, *182*, 104456.
- Lee, S. W. Y., Shih, M., Liang, J. C., & Tseng, Y. C. (2021). Investigating learners' engagement and science learning outcomes in different designs of participatory simulated games. *British Journal* of *Educational Technology*, 52(3), 1197-1214.
- Liu, S., Liu, S., Liu, Z., Peng, X., & Yang, Z. (2022). Automated detection of emotional and cognitive engagement in MOOC discussions to predict learning achievement. *Computers & Education*, *181*, 104461.
- Lorist, M. M., Boksem, M. A., & Ridderinkhof, K. R. (2005). Impaired cognitive control and reduced cingulate activity during mental fatigue. *Cognitive Brain Research*, *24*(2), 199-205.
- Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, 66(5), 1141-1164.
- Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, 1-22.
- Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and instruction*, *60*, 225-236.
- Munafo, J., Diedrick, M., & Stoffregen, T. A. (2017). The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. *Experimental brain research*, 235(3), 889-901.
- Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785.
- Purarjomandlangrudi, A., & Chen, D. (2020). Exploring the influence of learners' personal traits and perceived course characteristics on online interaction and engagement. *Educational Technology Research and Development*, *68*(5), 2635-2657.
- Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. Presence: Teleoperators & Virtual Environments, 10(3), 266-281.

Tanaka, M., Ishii, A., & Watanabe, Y. (2014). Neural effects of mental fatigue caused by continuous

- attention load: a magnetoencephalography study. *Brain research*, *1561*, 60-66. Trejo, L. J., Kochavi, R., Kubitz, K., Montgomery, L. D., Rosipal, R., & Matthews, B. (2005). Measures and models for predicting cognitive fatigue. In Biomonitoring for Physiological and Cognitive Performance during Military Operations, 5797, 105-115.
- Wang, M. T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. Learning and Instruction, 28, 12-23.



# The 11th Workshop on Technology-Enhanced STEM Education

# Review Process to Investigate Trends of Using Arduino to Enhance Al Study

Pornchai KITCHAROEN<sup>a\*</sup>, Suppachai HOWIMANPORN<sup>a</sup> & Sasithorn CHOOKAEW<sup>a</sup> <sup>a</sup>Department of Teacher Training in Mechanical Engineering, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Thailand \*p.kitcharoen@email.kmutnb.ac.th

**Abstract:** In the industrial sector, Artificial Intelligence (AI) technology has been extensively utilized and integrated into educational curricula and teaching methodologies to enhance educational effectiveness and respond to the needs of entrepreneurs. However, due to its costliness, more learning tools and technology are needed to ensure AI education. The Arduino board offers a user-friendly and cost-effective solution that enables the study of complex applications involving the integration of artificial intelligence with Arduino programming. This study presents a review process to investigate trends of using Arduino to enhance AI study. This study highlights the increasing complexity of using Arduino boards with artificial intelligence algorithms. Specifically, it explores their applications in domains such as multiple linear regression (MLR), particle swarm optimization (PSO), adaptive neuro-fuzzy inference systems (ANFIS), and Node-RED. The findings of this study will serve as a valuable reference for scholars interested in this domain and as a guide for AI education in the future.

Keywords: Al education, Machine Learning, Arduino, Neuro-Fuzzy, Node-RED

# 1. Introduction

Industry in the era of Industry 4.0 has undergone rapid changes and advancements driven by the widespread adoption of various technologies (Pajpach et al., 2022). These technologies have been extensively applied across agricultural, manufacturing, transportation, service, and other industrial sectors. By incorporating technology into industrial processes, efficiency has been increased, production costs have been reduced, time has been saved, and convenience has been enhanced, utilizing the benefits of 21st-century technology into education and teaching methods has been a critical aspect of leveraging the advantages of Industry 4.0 (Görgülü et al., 2021). However, a significant issue in education is the high cost of technology-related devices. Consequently, the Arduino board has been chosen as an affordable solution (Bullis et al., 2021; Guzmán-Fernández et al., 2021). This board offers professional-grade usability with the required precision and reliability (Lee, 2020; Makan et al., 2021).

Arduino microcontrollers have gained popularity in smart home systems, security applications, and STEM education, encompassing science, technology, engineering, and mathematics. They facilitate problem-solving skills, practical training, and activity-based learning (Sarhan, 2020; Sari et al., 2022). Arduino microcontrollers have also been employed in various dimensions of education, ranging from primary and secondary education to vocational and higher education levels. Their utilization aims to foster positive educational outcomes, increased student engagement, and a more enthusiastic approach to learning (Budi et al., 2021). These microcontrollers promote students' systematic thinking, creative ideation, and time management skills by possessing technological literacy and learning collaboratively in the classroom, and they can develop programming skills and apply them practically (Medeiros et al., 2018; Roumen et al., 2021). Additionally, Arduino facilitates the creation of simple learning devices in laboratories (Ueyama et al., 2022; Vidal et al., 2022) and enables the design of increasingly complex systems by incorporating innovative expertise or artificial

intelligence into Arduino programming, exploring the usage of advanced technologies (Tupac-Yupanqui et al., 2022).

Therefore, integrating technology, mainly through using Arduino microcontrollers, has significantly enhanced education, fostering positive learning outcomes and the development of programming skills. Arduino provides an affordable and reliable solution and empowers students to actively engage in learning, cultivate problem-solving abilities, and explore the vast possibilities of advanced technologies.

Artificial Intelligence (AI) involves designing and developing algorithms and models that enable machines to acquire and process information, extract meaningful insights, and make informed decisions or take appropriate actions based on the available data. Advanced technologies and many techniques are used in artificial intelligence, such as Multiple Linear Regression (MLR), Particle Swarm Optimization (PSO), and Adaptive Neuro-Fuzzy Inference Systems (ANFIS), especially in engineering education. In addition, Node-RED is an intelligent and agile system based on innovative communication for monitoring and control with industrial communication (Chookaew & Howimanporn, 2022).

In this study, a literature review was conducted to investigate trends in this field encompassing the following research questions:

RQ1: What research studies employ artificial intelligence with the Arduino board?

RQ2: What types of Arduino boards are utilized in the research study?

RQ3: What is the method to assess the accuracy of artificial intelligence research with the Arduino board?

# 2. Research Methodology

#### 2.1 Process of data searching and collection

A systematic search was conducted on the Scopus database using a predefined set of search keywords: ("artificial-intelligence" OR "machine-intelligence" OR "machine-learning" OR "multiple-linear-regression" OR "particle-swarm-optimization" OR "adaptive-neuro-fuzzy-inference-system" OR "Node-RED" AND "Arduino"). A total of 162 articles were found in selected education-related fields (i.e., "engineering", "computer science", "mathematics", "social sciences", "environmental science", "decision sciences"). Only articles with at least two keywords related to "MLR", "PSO", "ANFIS", "Node-RED", and "Arduino" in their titles, abstracts, or keywords were included in the study. After applying these criteria, the number of articles was reduced to 130 for further consideration. Subsequently, after further reviewing the complete text, ten articles were unrelated to the study, and four needed to be written in English. Finally, 18 articles remained that specifically addressed topics related to MLR, PSO, ANFIS, Node-RED, and Arduino, as shown in Figure 1.

#### 2.2 Data Analysis

Conducting a systematic literature review and analysis in this study is based on the principles and guidelines proposed by Chu et al. (2022). It encompasses data collection, inclusion and exclusion criteria, data coding, and analysis techniques. The publication of selected studies was analyzed, including the research article published in an international journal between 2018-2022. The criteria were used when choosing the valid article to make sure that there is consistency in the studies for data analysis that follow as:

- 1) The article must have been published in the last five years.
- 2) The article must have studied using the Arduino board for operating artificial intelligence with MLR, PSO, ANFIS, and Node-RED topics.
- 3) Each study assessed the accuracy applied to an Arduino board related to machine learning and artificial intelligence.



*Figure 1.* Research literature review process

# 3. Results

# 3.1 RQ1: What research studies employ artificial intelligence with the Arduino board?

In the field of MLR, several research studies were identified that integrated the use of MLR, PSO, ANFIS, and Node-RED with Arduino board in various contexts. For instance, Setyawan et al. (2022) analyzed MLR to determine the hydroponic nutrient solution from the water quantity and pump operation time. Songara and Patel (2022) compared three types of soil sensors to find the most suitable value for soil moisture. Bhoi et al. (2022) employed IoT-EMS to analyze the MLR values. Suresh et al. (2022) utilized an IoT-enabled deep-learning innovative irrigation system (IoTDL-SIS) to control an intelligent canal system. Deberneh and Kim (2018) focused on harvesting energy from a Wave Energy Converter (WEC) and predicting shore-side power by specifying the characteristics of sea waves using floating buoys.

In the area of PSO, Sheikh Ahmadi et al. (2022) aimed to maximize the energy yield of solar cells (PV) by combining the advantages of the incremental conductance (IC) method and PSO for Maximum Power Point Tracking (MPPT) control. Delavari and Naderian (2020) sought to enhance the robustness, flexibility, and level of independence of a microgrid inverter with parameter bounds of disturbances using PSO. De Moura Oliveira et al. (2022) designed a PSO-based Proportional Integral Derivative (PID) controller with the lowest absolute error and the slightest overall control signal variation. Sutyasadi and Wicaksono (2020) applied PSO to low-cost robot arm design and PID parameter tuning for small or medium-sized embedded

controllers. Daouadi et al. (2020) utilized the Takagi-Sugeno (T-S) fuzzy system combined with PSO to improve the stability of DC-DC boost converters.

Regarding ANFIS, Songara and Patel (2022) used ANFIS to determine suitable values for soil moisture. Sanjaya et al. (2019) designed a four-sensor robotic arm with ANFIS integration. Pandey et al. (2019) developed and employed a Multi-Adaptive Networked Fuzzy Inference System (MANFIS) for mobile robot navigation in different two-dimensional environments. Mujiarto et al. (2019) designed and implemented a five-sensor robotic arm with ANFIS integration.

Regarding Node-RED, Mohd Shaari Azyze et al. (2021) showcased the display of IoT data using the Node-RED platform for controlling air pollution and updating data every 5 minutes. Cgseong et al. (2020) designed and utilized a real-time object control system using Node-RED. Effendi et al. (2020) developed an intelligent farm and agriculture project that employed various internet-connected devices in the farm, including IoT for monitoring, with data inspection through Node-RED. Tedla et al. (2019) addressed the issue of crop loss during storage by using sensors to monitor environmental conditions in the storage facility and displaying real-time web applications via Node-RED. Huang et al. (2019) implemented a multi-agent system that used Node-RED to display real-time data on a screen.

According to Table 1 in the literature review, four articles utilized MLR, five employed PSO, three utilized ANFIS, five incorporated Node-RED, and 1 article combined MLR and ANFIS, as shown in Figure 2.

ID	Year	Authors	Source	Al
				Topic/Techniques
P1	2022	Setyawan et al.	Journal of Engineering Science	MLR
			and Technology	
P2	2022	Songara & Patel	Measurement	MLR+ANFIS
P3	2022	Bhoi et al.	Intelligent Automation &	MLR
			Soft Computing	
P4	2022	Suresh et al.	Computers, Materials & Continua	MLR
P5	2018	Deberneh & Kim	Applied Sciences	MLR
P6	2022	Sheikh Ahmadi et al.	Iranian Journal of Science	PSO
			and Technology	
P7	2020	Delavari & Naderian	IET Generation	PSO
P8	2020	De Moura Oliveira et al.	Algorithms	PSO
P9	2020	Sutyasadi & Wicaksono	Telecommunication Computing	PSO
			Electronics and Control	
P10	2020	Daouadi et al.	Electrotehnica	PSO
P11	2019	Sanjaya et al.	International Journal of Recent	ANFIS
			Technology and Engineering	
P12	2019	Pandey et al.	World Journal of Engineering	ANFIS
P13	2019	Mujiarto et al.	Indonesian Journal of Electrical	ANFIS
			Engineering and Computer Science	;
P14	2021	Mohd Shaari Azyze et al.	Asian Research Publishing Network	Node-RED
P15	2020	Cgseong et al.	Indonesian Journal of Electrical	Node-RED
			Engineering and Computer Science	)
P16	2020	Effendi et al.	International Journal of Integrated	Node-RED
			Engineering	
P17	2019	Tedla et al.	International journal of scientific &	Node-RED
			technology research	
P18	2019	Huang et al.	Journal of Manufacturing Systems	Node-RED

Table 1 Summary of the 18 included studies.



Figure 2. Research studies utilizing AI topics/techniques

# 3.2 RQ2: What types of Arduino boards are utilized in the research study?

Upon reviewing the literature, it was found that 18 articles utilized several types of Arduino boards for the research study. Most article uses Arduino Uno (6 articles), Arduino Mega 2560 (4 articles), Arduino Nano (2 articles), Arduino ESP-32 (2 articles), Arduino ESP8266 (1 article), Arduino Leonardos (2 articles), and two articles did not specify the Arduino board used, as shown in Figure 3.



Figure 3. Types of Arduino boards used in research studies.

# 3.3 RQ3: What is the method to assess the accuracy of artificial intelligence research with the Arduino board?

In the reviewed research studies that employed Arduino boards, the accuracy was assessed using various metrics, including R-squared (R^2), Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percent Error (MAPE), and Integral Absolute Error (IAE). Among the 18 articles included in the literature review, the evaluation of accuracy revealed the following distribution: R^2 was utilized (2 articles), MSE (4 articles), RMSE (1 article), MAE (2 articles), MAPE (2 articles), IAE (2 articles), and other evaluation methods were employed (10 articles). It is worth noting that some articles conducted multiple evaluations to assess accuracy, as shown in Figure 4.



Figure 4. The methods of accuracy assessment.

# 4. Conclusions

The findings reveal that MLR is utilized in four research articles, PSO in five articles, ANFIS in three articles, Node-RED in five articles, and a combination of MLR and ANFIS in one article. Moreover, the Arduino boards used in these studies consist of Arduino Uno in six articles, Arduino Nano in two articles, Arduino ESP8266 in one article, Arduino ESP-32 in two articles, Arduino Mega 2560 in four articles, Arduino Leonardo in two articles, and unspecified Arduino boards in two articles. The evaluation of research accuracy involves different metrics, including R-squared (r^2) in two articles, Mean Squared Error (MSE) in four articles, Mean Squared Error (MAE) in two articles, Mean Absolute Percentage Error (MAPE) in two articles, Integrated Absolute Error (IAE) in two articles employ multiple evaluation methods to assess accuracy.

This research highlights the increasing complexity of utilizing Arduino boards with artificial intelligence algorithms, particularly in MLR, PSO, ANFIS, and Node-RED domains. The findings contribute valuable insights for researchers interested in utilizing Arduino boards in various disciplines and facilitate the development of Arduino-based microcontroller applications. The study aims to serve as a comprehensive reference for future researchers in this field, providing them with a solid foundation for further investigations and advancements.

#### Acknowledgment

The authors thank the Department of Teacher Training in Mechanical Engineering, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok for support.

#### References

- Bhoi, S. K., Panda, S. K., Jena, K. K., Sahoo, K. S., Jhanjhi, N. Z., Masud, M., & Aljahdali, S. (2022). IoT-EMS: An Internet of Things Based Environment Monitoring System in Volunteer Computing Environment. *Intelligent Automation & Soft Computing*, *32*(3).
- Budi, A. H. S., Juanda, E. A., Fauzi, D. L. N., Henny, H., & Masek, A. (2021). Implementation of Simulation Software on Vocational High School Students in Programming and Arduino Microcontroller Subject. *Journal of Technical Education and Training*, 13(3), 108-114.
- Bullis, R., Coker, J., Belding, J., De Groodt, A., Mitchell, D. W., Velazquez, N., ... & Gunderson, J. E. (2021). The Fluorino: A Low-Cost, Arduino-Controlled Fluorometer. *Journal of Chemical Education*, 98(12), 3892-3897.
- Cgseong, C. G., Kim, J. Y., & Park, D. J. (2020). Real-time object control system using open source platform. *Indonesian Journal of Electrical Engineering and Computer Science*, *20*(1), 313-319.
- Chookaew, S., & Howimanporn, S. (2022). Upskilling and reskilling for engineering workforce: implementing an automated manufacturing 4.0 technology training course. *Global Journal of Engineering Education*, 24(1).
- Chu, S. T., Hwang, G. J., & Tu, Y. F. (2022). Artificial intelligence-based robots in education: A systematic review of selected SSCI publications. *Computers and education: Artificial intelligence*, 100091.
- Deberneh, H. M., & Kim, I. (2018). Predicting output power for nearshore wave energy harvesting. *Applied Sciences*, *8*(4), 566.
- Delavari, H., & Naderian, S. (2020). Design and HIL implementation of a new robust fractional sliding mode control of microgrids. *IET Generation, Transmission & Distribution, 14*(26), 6690-6702.
- De Moura Oliveira, P. B., Hedengren, J. D., & Solteiro Pires, E. J. (2020). Swarm-based design of proportional integral and derivative controllers using a compromise cost function: An Arduino temperature laboratory case study. *Algorithms*, *13*(12), 315.
- Daouadi, R., Ounass, D., Soufi, Y., Metatla, A., & Guiza, D. (2020). Real-time Implementation of a Takagi-Sugeno Fuzzy Controller for a DC-DC Boost Converter. *Electrotehnica, Electronica, Automatica*, 68(4).
- Effendi, M. K. R., Kassim, M., Sulaiman, N. A., & Shahbudin, S. (2020). IoT smart agriculture for aquaponics and maintaining goat stall system. *International Journal of Integrated Engineering*, *12*(8), 240-250.
- Görgülü Arı, A., & Meço, G. (2021). A New Application in Biology Education: Development and Implementation of Arduino-Supported STEM Activities. *Biology*, *10*(6), 506.
- Guzmán-Fernández, M., Zambrano de la Torre, M., Ortega-Sigala, J., Guzmán-Valdivia, C., Galvan-Tejeda, J. I., Crúz-Domínguez, O., ... & Durán-Muñoz, H. A. (2021). Arduino: a Novel Solution to the Problem of High-Cost Experimental Equipment in Higher Education. *Experimental Techniques*, 1-13.
- Huang, Z., Kim, J., Sadri, A., Dowey, S., & Dargusch, M. S. (2019). Industry 4.0: Development of a multi-agent system for dynamic value stream mapping in SMEs. *Journal of Manufacturing Systems*, 52, 1-12.
- Lee, E. (2020). A Meta-Analysis of the Effects of Arduino-Based Education in Korean Primary and Secondary Schools in Engineering Education. *European Journal of Educational Research*, *9*(4), 1503-1512.
- Makan, G., Mingesz, R., & Gingl, Z. (2021). Reliable readings from Arduino voltmeters. *Physics Education*, 56(4), 043001.
- Medeiros, R. P., Ramalho, G. L., & Falcão, T. P. (2018). A systematic literature review on teaching and learning introductory programming in higher education. *IEEE Transactions on Education*, *62*(2), 77-90.
- Mujiarto, A. D., Komaro, M., Mohamed, M. A., & Rahayu, D. S. (2019). Colored object detection using 5 dof robot arm based adaptive neuro-fuzzy method. *Indonesian Journal of Electrical Engineering and Computer Science*, *13*(1), 293-299.
- Mohd Shaari Azyze, N. L. A., Md Isa, I. S., & Kamaruzaman, N. S. (2021). Development Of Real-Time Lora-Based Air Pollution Monitoring System. *Asian Research Publishing Network (ARPN), 16*(9)

- Pandey, A., Kashyap, A. K., Parhi, D. R., & Patle, B. K. (2019). Autonomous mobile robot navigation between static and dynamic obstacles using multiple ANFIS architecture. *World Journal of Engineering*.
- Pajpach, M., Haffner, O., Kučera, E., & Drahoš, P. (2022). Low-cost education kit for teaching basic skills for industry 4.0 using deep-learning in quality control tasks. *Electronics*, *11*(2), 230.
- Roumen, G. J., & Fernaeus, Y. (2021). Envisioning Arduino Action: A collaborative tool for physical computing in educational settings. *International Journal of Child-Computer Interaction*, 29, 100277.
- Sarhan, Q. I. (2020). Systematic survey on smart home safety and security systems using the Arduino platform. *IEEE Access*, *8*, 128362-128384.
- Sari, U., Çelik, H., Pektaş, H. M., & Yalçın, S. (2022). Effects of STEM-focused Arduino practical activities on problem-solving and entrepreneurship skills. *Australasian Journal of Educational Technology*, 38(3), 140-154.
- Sarı, U., Pektaş, H. M., Şen, Ö. F., & Çelik, H. (2022). Algorithmic thinking development through physical computing activities with Arduino in STEM education. *Education and Information Technologies*, 27(5), 6669-6689.
- Sanjaya, W. S. M., Sambas, A., Jannah, M., Rahayu, D. S., Mamat, M., & Mohamed, M. A. (2019). Design of 4 Dof Robot ARM Based on Adaptive Neuro-Fuzzy (ANFIS) using Vision in Detecting Color Objects. *International Journal of Recent Technology and Engineering*, 8(2S7), 224-227.
- Setyawan, T. A., Nugroho, A. S., Febyana, A., & Pramono, S. (2022). Multiple linear regression method used to control nutrient solution on hydroponic cultivation. *Journal of Engineering Science and Technology*, *17*(5), 3460-3474.
- Sheikh Ahmadi, S. H., Karami, M., Gholami, M., & Mirzaei, R. (2022). Improving MPPT performance in PV systems based on integrating the incremental conductance and particle swarm optimization methods. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, 1-13.
- Songara, J. C., & Patel, J. N. (2022). Calibration and comparison of various sensors for soil moisture measurement. *Measurement*, *197*, 111301.
- Sutyasadi, P., & Wicaksono, M. B. (2020). Joint control of a robotic arm using particle swarm optimization based H2/H∞ robust control on arduino. *Telecommunication Computing Electronics and Control*, *18*(2), 1021-1029.
- Suresh, P., Aswathy, R. H., Arumugam, S., Albraikan, A. A., Al-Wesabi, F. N., Hilal, A. M., & Alamgeer, M. (2022). IoT with Evolutionary Algorithm Based Deep Learning for Smart Irrigation System. *Computers, Materials & Continua*, Vol. 71, No. 1, pp. 1713-1728).
- Tedla, T. B., Bovas, J. J. L., Berhane, Y., Davydkin, M. N., & James, P. S. (2019). Automated granary monitoring and controlling system suitable for the sub-Saharan region. *International journal of scientific & technology research*, *8*(12), 1943-1951.
- Tupac-Yupanqui, M., Vidal-Silva, C., Pavesi-Farriol, L., Ortiz, A. S., Cardenas-Cobo, J., & Pereira, F. (2022). Exploiting Arduino features to develop programming competencies. *IEEE Access*, 10, 20602-20615.
- Ueyama, Y., Sago, T., Kurihara, T., & Harada, M. (2022). An Inexpensive Autonomous Mobile Robot for Undergraduate Education: Integration of Arduino and Hokuyo Laser Range Finders. *IEEE Access*, *10*, 79029-79040.
- Vidal, J., Bravo, D. A., & Rengifo, C. F. (2022). Teaching Physical Sciences Using Arduino Physics Lab at the Universidad Del Cauca. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, *17*(3), 223-229.

# Proposing a Training Model on Energy Management of Compressed Air Systems with Artificial Intelligence of Things

Noppadon MONOK<sup>a\*</sup>, Suppachai HOWIMANPORN<sup>a</sup> & Sasithorn CHOOKAEW<sup>a</sup>

<sup>a</sup> Department of Teacher Training in Mechanical Engineering, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Thailand \*s6502017910028@email.kmutnb.ac.th

Abstract: Technology 4.0 in the industry focuses on effective energy saving. There is an increase in the need for industrial worker training, especially in compressed air systems, air kept under more significant pressure than atmospheric pressure. It is an important medium for the transfer of energy in industrial processes. However, the new technology of energy management systems to enhance performance has yet to be widespread. Owing to the learning material for training is expensive. It is also not possible to simulate the operation of a comprehensive system. The employees lack an understanding of the big picture, causing problems in learning and lacking practical skills. Especially new employees may need more knowledge and experience and have operation errors or problem-solving skills. In this study, we proposed a training model consisting of a compressed air systems training kit based on the Artificial Intelligence of Things (AloT), and the energy-saving scenarios consist of 1) controlling the compressed air pressure fed to the air cylinder while being subjected to loads of different sizes. 2) controlling the compressed air pressure supplied to the air cylinder while operating the air cylinder without load. 3) controlling the compressed air supplied to the machine while it is stopped. In addition, we use an outcome-based approach to drive training activities. The results show that the training model can simulate the process of energy management systems of compressed air systems and show it on the dashboard. It can be implemented in employee or engineer practical training in the future.

Keywords: energy management systems, training kit, AloT, engineering education

# 1. Introduction

Nowadays, industrial manufacturing and companies increasingly recognize the financial and reputational benefits of meeting sustainable targets concerning waste and excessive energy consumption (Abela et al., 2020). Although compressed air in the industry is an important yet overlooked energy carrier. However, in the industrial sector, compressed air systems are one primary energy consumer, accounting for around 10% of the electricity consumed in the European Union and China, while the US, Malaysia, and South Africa account for 9% of total energy consumption (Bonfá et al., 2019; Eras et al., 2020) Therefore, reducing CAS energy consumption is crucial to meet the energy efficiency goals (Benedetti et al., 2019).

Compressed Air Systems (CAS) provide a significant improvement opportunity for energy saving. CAS refers to air kept under more significant pressure than atmospheric pressure. It is an important medium for the transfer of energy in industrial processes.

Many studies attempt to propose a methodology to develop the energy management of CAS. For example, Thabet et al. (2020) proposed real-time sensing and machine learning to increase CAS efficiency with algorithms that automate the detection of energy inefficiencies and make decisions regarding suitable troubleshooting procedures will be created. Sanders et al. (2020) proposed new intelligent techniques to save energy in compressed air systems with real-time ambient sensing via Artificial Intelligence and Knowledge Management to automatically improve efficiency in energy intensive manufacturing.

This paper uses Artificial Intelligence of Things (AloT) technology to develop the energy management of compressed air systems training model. AloT will reduce energy consumption in compressors by considering real-time circumstances and predicted needs. Sensor data will deliver real-time performance information, interpret the data, and then act automatically.

This training course focuses on new technology that will be applied to save energy in industrial manufacturing. Therefore, we proposed a training model consisting of a compressed air systems training kit based on the Artificial Intelligence of Things (AIoT) and the energy-saving scenarios to align the outcome-based approach.

# 2. Related work

#### 2.1 Training course of a compressed air system

For the industrial sector, CAS is necessary to offer versatility, reliability, and ease of use in various industrial applications. It is a cost-effective and efficient solution for powering tools, controlling processes, and facilitating numerous manufacturing and operational tasks. CAS is a network that generates energy from gas molecules being transported through and activating an engine or rotor. It uses natural air that is all around us that we inhale and exhale. The system uses compressed gas molecules to create sufficient pressure, which pushes through the rest of the circuit. Therefore, the concept of training is shown in Table 1.

Concepts	Leaning outcomes
1. Management of industrial	Understand the management compressed air systems
compressed air systems	
2. Programmable logic	Use PLC to control hardware.
controller (PLC)	
3. Ladder programming	Write Ladder logic programming with PLC.
4. Pneumatic cylinders	Write programming to control pneumatic cylinders.
5. Photoelectric sensor	Apply a photoelectric sensor to detect an object.
6. Vision inspection systems	Apply a vision inspection system to track an object.
7. Electro Pneumatic control	Write program control Electro Pneumatic control
8. Vacuum grippers control	Use Vacuum grippers to pick and plate the workpiece.
9. Artificial Intelligence	Apply Artificial Intelligence to save energy
10. Dashboard	Display information on the dashboard

Table 1. The training concepts of leaning outcomes

# 2.2 Artificial Intelligence of Things (AloT)

Artificial Intelligence of Things (AloT) refers to integrating artificial intelligence (Al) technologies with Internet of Things (IoT) devices and systems. AloT combines the power of Al algorithms and data analytics with IoT devices' connectivity and sensor capabilities to enable intelligent decision-making, automation, and advanced functionalities (Hu et al., 2021; Xian et al., 2023). Many studies attempt to present AloT applications in many fields. For example, El Himer et al. (2022) present applications of AloT technology to manage and save energy in renewable energy sources such as solar and wind. Shi (2021) presented AloT as a new technology for developing innovative, low-cost, reliable monitoring systems for smart home applications.

# 3. Training Model on Energy Management of Compressed Air Systems

#### 3.1 Proposed Training model

In this study, we designed a training model focusing on industrial employee practice training. A training course should have a hands-on activity that can operate related work. To understand the engineering problem of saving energy, we design the training model to perform the activities and appropriately prepare devices or materials based on a conceptual framework, as shown in Figure 1.



Figure 1. Training model framework

In addition, we design energy savings scenarios to consist of 1) controlling the compressed air pressure fed to the air cylinder while being subjected to loads of different sizes. 2) controlling the compressed air pressure supplied to the air cylinder while operating the air cylinder without load. 3) controlling the compressed air supplied to the machine while it is stopped, as shown in Figure 2.



Figure 2. Energy savings scenarios

#### 3.2 A compressed air systems training kit

To design the structure of the training kit, we identify the performance of training materials able to compare the scenarios in case of saving energy. Figure 3. shows a compressed air systems training kit consisting of a Programmable Logic Controller or PLC (Omron NX1P2 V1.50), a device that automatically controls a machine. A load cell sensor is an electromechanical sensor that measures force or weight. A distance sensor is a sensor that measures the distance by emitting a signal and measuring how long it takes to return to the

transducer. An air regulator is a pneumatic device that receives air at any pressure within its tolerance and then dispenses air at a pressure no more significant than its intended output. A DC-to-DC converter temporarily stores electrical energy to convert the direct current (DC) from one voltage level to another. A camera is an optical instrument that captures images. A cylinder is A device that converts energy from wind pressure into mechanical energy.



Figure 3. The component of a compressed air systems training kit

# 3.3 Neuro-fuzzy algorithm

Neuro-fuzzy is an artificial intelligence method effectively utilized in control engineering to demonstrate neuro-fuzzy usage in training for managing the energy of compressed air systems. Neuro-fuzzy algorithms, also known as fuzzy neural networks or adaptive neuro-fuzzy inference systems (ANFIS), combine elements of fuzzy logic and neural networks to create hybrid models that can handle complex and uncertain data. These algorithms aim to leverage the strengths of both fuzzy logic and neural networks to improve the accuracy and interpretability of the models (Tiruneh et al.,2020; De et al., 2020). In this study, PLC is necessary for building its control programs and is appropriately tied with AloT. In addition, the MATLAB program can run a Neuro-fuzzy algorithm, as shown in Figure 4.



Figure 4. (a) Example of Membership Functions in MATLAB (b) Example of PLC program

#### 3.4 Dashboard

A dashboard is a visual representation of essential information, data, metrics, and performance indicators displayed in a consolidated and easily understandable format. It provides a snapshot or summary of data to help users monitor, analyze, and make informed decisions based on the information presented. Node-RED is a development tool for connecting hardware devices, application programming interfaces, and online services (as shown in Figure 5). It has a browser-based editor to enable wiring together flows using the wide range of nodes in the palette. The industrial flows automation created in the Node-RED tool and a programmable logic controller (PLC) can be stored, imported, and exported for information-sharing as a dashboard in the gateway connected with the open platform communications unified architecture (Chookaew & Howimanporn, 2022).



Figure 5. The screen snapshot of Node-RED dashboard

# 4. Conclusion and Future work

In this study, we proposed a training model consisting of a compressed air systems training kit based on the Artificial Intelligence of Things (AloT), and the energy-saving scenarios consist of 1) controlling the compressed air pressure fed to the air cylinder while being subjected to loads of different sizes. 2) controlling the compressed air pressure supplied to the air cylinder while operating the air cylinder without load. 3) controlling the compressed air supplied to the machine while it is stopped. In addition, we use an outcome-based approach to drive training activities. The results show that the training model can simulate the process of energy management systems of compressed air systems and show it on the dashboard. It can be implemented in employee or engineer practical training. We plan to investigate the effectiveness of the proposed models by comparing pre- and post-test performance and evaluating attitudes toward training as part of future work.

# Acknowledgements

The part of the research was funded by SMC Pneumatic Automation Training Center Thailand, SMC (Thailand) Ltd., and we thank King Mongkut's University of Technology North Bangkok for their support.

#### References

- Abela, K., Refalo, P., & Francalanza, E. (2020). Design and implementation of an energy monitoring cyber physical system in pneumatic automation. *Procedia CIRP*, 88, 240-245.
- Benedetti, M., Bonfà, F., Bertini, I., Introna, V., Salvatori, S., Ubertini, S., & Paradiso, R. (2019). Maturitybased approach for the improvement of energy efficiency in industrial compressed air production and use systems. *Energy*, 186, 115879.
- Bonfá, F., Benedetti, M., Ubertini, S., Introna, V., & Santolamazza, A. (2019). New efficiency opportunities arising from intelligent real time control tools applications: the case of Compressed Air Systems' energy efficiency in production and use. *Energy Procedia*, 158, 4198-4203.
- Chookaew, S., & Howimanporn, S. (2022). Upskilling and reskilling for engineering workforce: implementing an automated manufacturing 4.0 technology training course. *Global Journal of Engineering Education*, 24(1), 34-39.
- De Campos Souza, P. V. (2020). Fuzzy neural networks and neuro-fuzzy networks: A review the main techniques and applications used in the literature. Applied soft computing, 92, 106275.
- Eras, J. J. C., Gutiérrez, A. S., Santos, V. S., & Ulloa, M. J. C. (2020). Energy management of compressed air systems. Assessing the production and use of compressed air in industry. *Energy*, 213, 118662.
- El Himer, S., Ouaissa, M., Ouaissa, M., & Boulouard, Z. (2022). Artificial Intelligence of Things (AloT) for Renewable Energies Systems. *In Artificial Intelligence of Things for Smart Green Energy Management* (pp. 1-13). Cham: Springer International Publishing.
- Hu, X., Li, Y., Jia, L., & Qiu, M. (2021). A novel two-stage unsupervised fault recognition framework combining feature extraction and fuzzy clustering for collaborative AloT. *IEEE Transactions on Industrial Informatics*, 18(2), 1291-1300.
- Sanders, D. A., Robinson, D. C., Hassan, M., Haddad, M., Gegov, A., & Ahmed, N. (2019). Making decisions about saving energy in compressed air systems using ambient intelligence and artificial intelligence. In Intelligent Systems and Applications: *Proceedings of the 2018 Intelligent Systems Conference (IntelliSys)* Volume 2 (pp. 1229-1236). Springer International Publishing.
- Shi, Q., Zhang, Z., Yang, Y., Shan, X., Salam, B., & Lee, C. (2021). Artificial intelligence of things (AloT) enabled floor monitoring system for smart home applications. *ACS nano*, 15(11), 18312-18326.
- Thabet, M., Sanders, D., Becerra, V., Tewkesbury, G., Haddad, M., & Barker, T. (2020, August). Intelligent energy management of compressed air systems. *Proceeding of 2020 IEEE 10th International Conference on Intelligent Systems (IS)* (pp. 153-158). IEEE.
- Tiruneh, G. G., Fayek, A. R., & Sumati, V. (2020). Neuro-fuzzy systems in construction engineering and management research. *Automation in construction*, 119, 103348.
- Xian, W., Yu, K., Han, F., Fang, L., He, D., & Han, Q. L. (2023). Advanced Manufacturing in Industry 5.0: A Survey of Key Enabling Technologies and Future Trends. *IEEE Transactions on Industrial Informatics*. 1 – 15.

# An Inquiry-based Learning Approach in Engineering Education regarding Simulation Practice on Automation Control.

#### Santi HUTAMARN<sup>a\*</sup>

<sup>a</sup> Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Thailand \* santi.h@fte.kmutnb.ac.th

**Abstract:** Automatic control is currently used widely, so it is in engineer curriculum of every major of institutions, which is instructed in most major curriculums. Each curriculum focuses on application of mathematics in high theory in order to describe the system behavior, which makes learners in automatic control subject. The learners face with the understanding of occurring after the input stimulation, which results the learner's lack of practice skills on instrument and control. Therefore, to promote learners to understand the system behavior, and gain practice skills, this research presents the instructional model focused on learner's vision on the system behavior by using computer program in order to simulate the abstract work on computer screen, and connect to concrete work from op-amp electrical circuit on PCB, and measure the system behavior from inquiry -based learning. Teachers are facilitators to monitor learners to connect the simulation with the testing of op-amp instrument. It is operated effectively by learners who are able to see the overall results from the experimental presentation through websites, which learners present through google sites. The researcher will take the instructional model to use with third-year learners.

Keywords: Inquiry-based Learning, Simulation, Operational Amplifier

#### 1. Introduction

The control system engineering (or automatic control) is an important work in the production process in order to produce work following the design effectively. It is an autonomous driving, which is driven with robot. Therefore, an engineering education needs to provide control system engineering in every major that learners have to learn about engineering. Learning in the past, teachers taught theory using equation in applied mathematics for describing phenomenon and behavior of (Pena et al., 2022) linking theory to practice. At present the computer program is used to reduce the gap between theory that learners need to have prior knowledge about mathematics before, it leads to real practice. Teachers have to use various active learning instructions in order to help learners who have poor basic mathematics to understand the system behavior. (Al-khazraji & Ismaeel, 2020) The software program sometimes used MATLAB as a device for learning. It helps learners to have more vision on system behavior, that learners are able to adapt parameter value in the same time as system response on computer screen. (Marin et al., 2020) When learners are able to connect theory to software computer, the connection to practice, it is developed for model in industrial work, (Hernandez et al., 2020; Kaluz, Klauco, & Cirka, 2019; Arrieta, Duque & Villa, 2019) which helps learners understand and connect theory to practice completely, in the same time the work of the model development of the system has the limitation that the learners need to have prior knowledge on electronic sensor, and electronic measurement, or microcontroller, which makes learners successfully-experimented. (Rojas-Moreno, 2019) According to the mention in this paper to present instructional model by using the computer program together with the electronics simulation in order to practice with theory in automatic control subject for third-year bachelor's degree student on the topic of transfer function of 1<sup>st</sup> order, and 2<sup>nd</sup> order system, and controller design using SISO function of MATLAB program in order to test with the system developed.

# 2. Related Works

# 2.1 Automatic Control Software Simulation

Taking software package to help in the simulation with the engineering education that it helps learners to connect theory of the complexed equation in applied mathematics with the computer program, which stimulates physical condition related to theory easily. (Kumar & Milanovic, 2022) Learners are able to see the phenomenon after the mathematics value adaptation on the computer screen immediately, and it also helps learners to understand in depth verv well. (Herrmann et al., 2022) As the simulation related with the control system engineering views on the behavior in the form of time response graph by comparing between electrical signal and electrics, which stimulate the system called input with the response system in the form of the electrical signal. If the systems are mechanical system, chemical system, or others that are converted to electrical system by sensor. Therefore, learners want to experiment on the behavior system, the learners have to learn about the sensor equipment, and electrical measurement device before. It is the difficulty of learners' learning who do not study in the electrical major before. So, taking the computer program to use in the simulation for experiment on the physics phenomenon for theory concepts to help to reduce the difficulty of learners' learning very well. (Fennell et al., 2019) The proper computer programs are various which are able to stimulate the work in the form of virtual laboratory in order to design the controller of the work system. Learners are able to view the response in the form of graph, and the illustrated photos perfectly, they can see the results each of point on the graph very perfectly in the form of real time, and it reduces the complexity, the danger which will occur with learners, and it also reduces the time length for developing the stimulation model. (Dosoftei & Cojocaru, 2020; Fan et al., 2020; Huilcapi et al., 2023; Vermunt, 2023) Therefore, the simulation is a proper option for learners who practice, and prepare the readiness before the experimentation with the real stimulation model. (Andreenkov & Shunaev, 2022)

# 2.2 Automatic Control Hardware Simulation

Taking the active device, especially the operational amplifier : Op-amp to develop circuits and take the software program to use in the work simulation. Besides, it helps learners to understand the phenomenon of work, and helps learners to create ideas of Circuits design. Learners are able to compute by using the formula, and measure the parameter value from the electric signal device, which shown in the program immediately to promote learners to have more confidence in the application of the parameter estimation in the Circuits. (Tsvetkova & Borodzhieva, 2021) Moreover, using the Op-amp instead of the second order system, which used in the control system engineering. It is the analog Circuits, leaners are able to connect the Circuits by themselves by using the Op-amp device. The advantages of this device are cheap, can be bought easily, and can be connected through protoboard Circuits. Besides, the Op-amp is able to developed to analog controller easily, and be satisfyingly stable. (Li et al., 2022) Using the Op-amp for the stimulation by developing the printed circuit board : PCB in order to reduce the mistake of connecting the Circuits, and the signal shape which has more noise than the stimulation from real system. (Lara et al., 2018)The experimentation of the measurement, and the parameter value testing from the signal shape easily, and relates to the theory. The learners are able to view the phenomenon, which occurs after controlling, which relates to theory and practice very perfectly. (Beloiu, 2015; Subtirelu, 2015)

# 2.3 Inquiry-based Learning

Inquiry-based learning is learning management, which focuses on child-center. It emphasizes learners who can search knowledge by themselves by using scientific method. Teachers change their roles to be facilitators and support resources for learning, which can be applied for learning management in engineering education concerning with engineering practice found that it can promote learners have positive attitude towards engineering learning, which uses mathematics equation and understanding in content. After that the design is taken to verify and observe behavior, then report the conclusion occurred from the design, which uses theory knowledge. Learners work in groups, brainstorm, and cooperate to find the answers from the designing step, which assigned by the teachers to try-out and test from these make cooperating skills to find the answers (Gero, Catz & Sabag, 2018), while they help learners have competency on cooperation work including report writing, and the interpreting scientific articles (Jarvisalo et. Al., 2020). Taking the inquiry-based learning to use for this learning, it is the talented issue for teachers in practicing subject (Brookes, 2018). Teachers help learners while problem occurring by learners classify problems and used for problem solving to prevent learners loose the concepts and feel discouraged.

#### 3. Learning approach for automatic control :

#### 3.1 Learning Approach

Learning approach in this research presents the model of the instruction by using the inquirybased learning together with the active device in the type of op-amp, which stimulated to be 1<sup>st</sup> order, and 2<sup>nd</sup> order system including the design of controlling Circuits with the op-amp device. The advantages of the learning of the inquiry-based learning. It helps learners to have good attitude towards learning. (Mueni et al., 2023) It is suitable for the instruction based on both theory and practice, which helps learners understand the science concept to promote the collaboration working, and team of learners. (Dobber et al., 2017) The steps are as follows: step 1 the teacher random the learner's group by each of group has not more than 3 members. Each week teacher defines the learning in each week. Besides, step 2 learners in each group operates the simulation with MATLAB Simscape program using the op-amp circuits, it is the main design by learners have to use the computing formula of the resistor, and the capacitor from theory which is suitable for the questions that the teacher defined in each group in order to connect to 1<sup>st</sup> order, and 2<sup>nd</sup> order system. Then measurement of the time constant value from the computer program is done. When the topic of the controlling design uses the order of SISO in the design of controller instead. When the learners understand the Circuits, they connect PCB Circuits by themselves, step 3 the learners measure the time constant from the PCB compared with the theory, it is the step 4 learners check the accuracy, and indicate the mistake in practice, and represent the computing formula, and the last step is learners present work, which is done in each week through webpage with google site program. The step shown in the figure 1. Beginning form step 1: the teacher defines the learning objectives, informs learning principles that learners must learn from the theory to practice, then use questions to encourage learners to find the answers, which are not covered the learners' knowledges, step 2: learners are divided into 3-5 people, after that the teacher presents theory principles, and have learners brainstorm by learners' prior knowledges to find the answers together with the teacher, step 3: expanding the theory concepts to simulate from MATLAB to practice with opamp on PCB, this step helps the teacher to guide new principles for occurred-problem solving. If the results of testing from the simulation, op-amp, and on PCB are not relevant to the theory, and step 4: learners must conclude the knowledge from the simulation, and practice with electronics gadgets, which how it is related to the theory. If problems are occurred during work, how to solve the problems, and the last step: the teacher must evaluate learner's knowledge, and skills. From the observation, how learners take concepts, and new skills to apply, including skills and team-work process. The role of learners for learning on inquiry-based learning, learners must pay attention to build concepts by themselves by before entering class, they should review the background knowledge to be ready for new lesson or prepare questions for asking in the classroom.



Figure 1. Learning Approach

# 3.2 Simulation Step

The MATLAB program in this research comprised 2 parts as follows:

#### 3.2.1 Simscape

Simscape is the program by using drag and drop device, and draw to connect together. The learners are able to use the electrical measurement device as the authentic device, which shown in figure 2. Learners use simscape to stimulate the work of op-amp Circuits before connecting the Circuits on the PCB sheet.



Figure 2. Simscape on MATLAB Program

# 3.2.2 SISOTOOL command

SISOTOOL command is the command in the text mode of the MATLAB program, which shown in figure 3. It is the type of graphic user interface : GUI, which users are able to fill the transfer function of the parameter value form the op-amp Circuits on simscape in order to design the

controller of parameter value of controller, which shown on computer screen. The users are able to take parameter value to design analog controlling Circuits by using the op-amp immediately.



Figure 3. GUI SISOTOOL on MATLAB Program

3.2.3 The op-amp Circuits is a Circuits of 1<sup>st</sup> order, and 2<sup>nd</sup> order system, which contained on PCB sheet, it has the point for line connecting, and the connecting point for measure electrical signal between input compared with output, which shown in figure 4. Moreover, it is able to take op-amp Circuits to develop to analog controller Circuits in order to be close loop control as well.



Figure 4. Op-amp Circuits on PCB

# 3.4 Assessment and Evaluation Step

The evaluation form comprises 2 parts as follows: skill evaluation form, which used for evaluating each learner's skills by this skill evaluation form will do after finishing the learning unit by using the rubrics score. Which is the criterion for the evaluation. After finishing of the skill evaluation, the learners' attitudes towards learning will be done using the criterion of Likert 5-level.

# 4. Conclusion & future work :

This research presents the model of instruction by using the inquiry-based learning model in an automatic control subject by group of learners help to find the information including with practice by using the MATLAB computer program in the simulation, and the controller together with practice the connecting on the PCB sheet. Learners measure the electrical signal on input compared with output in order to view the response of the system in the form of graph. Moreover, learners present the work via websites. After completing all topics, Learners do the skill test, and attitude evaluation towards learning. When the information is from the evaluation. The research is presented by using percentage, and standard deviation for further research. In the future these principles will be used in other related subjects with Automatic Control, which are more complexed by comparing with other models such as, project-based learning or flipped classroom.

#### Acknowledgements

This research was funded by Faculty of Technical Education, King Mongkut's University of Technology North Bangkok (KMUTNB) contract no. FTE-2566-02. We would like to thank all the people who concerned of this fund.

#### References

- Al-khazraji, A., & Ismaeel, S. (2020). Efficient practice on teaching modeling and simulation for engineers. *In 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE.
- Andreenkov, E., & Shunaev, S. (2022). Application of simulation modeling as a replacement for laboratory practice in engineering education. *In 2022 VI International Conference on Information Technologies in Engineering Education (Inforino)* (pp. 1-6). IEEE.
- Arrieta, M.P., Duque, J.E., & Villa, J.L. (2019). Implement a like-industrial process for control education. In 2019 IEEE 4<sup>th</sup> Columbian Conference on Automatic Control(CCAC) (pp.1-6). IEEE.
- Beloiu, R. (2015). Hands-on solutions for teaching basic introductory control systems course. *In 9th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*(pp. 89-93). IEEE.
- Brookes, W. (2018). Inquiry-based Learning of Database Concepts. In 17<sup>th</sup> International Conference on Information Technology Based Higher Education and Training (ITHET) (pp.144-149).
- Dobber, M., Zwart, R., Tanis, M., & Oers, B.V. (2017). Literature review : the role of teacher in inquirybased education. Educational Research Review 22(2017), 194-214.
- Dosoftei, C.C. & Cojocaru, A.E., (2020). Implementation of a virtual control lab to support teaching in engineering control. *In 2020 International Conference and Exposition on Electrical And Power Engineering (EPE)* (pp. 699-703). IEEE.
- Fan, H., Zhang, J., Li, Y., Feng, Q., Fang, K., Wen, H., Lin, L., Qi, X., Diao, X., Bonizzoni, E., Maloberti, F., Ghannam, R., & Heidari, H. (2020). Innovative engineering education in circuits & systems. *In 2020 IEEE International Symposium on Circuits and Systems (ISCAS)*(pp. 1-4). IEEE.
- Fennell, H.W., Lyon, J.A., Magana, A.J., Rebello, S., Rebello, C.M., & Peidrahita, Y.B. (2019). Designing hybrid physics labs : combining simulation and experiment for teaching computational thinking in first-year engineering. In 2019 IEEE Frontiers in Education Conference (FIE) (pp. 1-8). IEEE.
- Gero, A., Catz, B. & Sabag, N. (2018). Inquiry-Based Electronics Laboratory : Student's Attitudes. *In* 2018 28<sup>th</sup> European Association for Education in Electrical and Information Engineering (EAEEIE) Annual Conference. (pp. 1-9).
- Herrmann, S., Felder, D., Padligur, M., Brosch, S., Geiger, M., Stockmeier, F., Baitalow, K., Rall, D., Femmer, R., Roghmans, F., Hauser, M., Mehlis, J., Linkhorst, J. & Wessling, M. (2022). Handson kinetic measurements and simulation for chemical process engineering students. *Education for Chemical Engineers*, 41 (2022) 14-21.
- Huilcapi, V., Cajo, R., Ponguillo, V., & Gomez, G. (2023). Educational platform for teaching automation and control engineering : a pneumatic levitation system. *In 2023 IEEE World Engineering Education Conference (EDUNINE)* (pp. 1-6). IEEE.
- Jarvisalo, H., Korhonen, J., Naukkarinen, J. & Silventoinen, P. (2020). From Seminar Presentations to Inquiry-based Learning-Developing an Advanced Electronics Course. In An International Conference on Engineering, Technology and Education (TALE2020) (pp.626-629). IEEE.
- Kaluz, M., Klauco, M. & Cirka, L. (2019). Flexy<sup>2</sup>: A portable laboratory device for control engineering education. *IFAC PapersOnline* 52-9, 42-47.
- Kumar, S., & Milanovic, I. (2022). Simulation-supported engineering curriculum. In 2022 IEEE IFEES World Engineering Education Forum - Global Engineering Deans Council (WEEF-GEDC) (pp. 1-6). IEEE.

- Lara, V., Vargas, H., Castro, C., Chacon, J., & Torre, L. (2018). Design and development of flexible control laboratory plant for educational purposes. *In 2018 IEEE International Conference on Automation/XXIII Congress of the Chilean Association of Automatic Control (ICA-ACCA)*(pp. 1-6). IEEE.
- Li, J.H. (2022). Genetic algorithm PID control of second-order analog circuit system. *In 2022 International Conference on System Science and Engineering (ICSSE)* (pp. 022-025). IEEE.
- Marin, L., Vargas, H., Heradio, R., Torre, L., Diaz, J.M., & Dormido, S. (2020). Evidence-based control engineering education : Evaluating the LCSD simulation tool. *IEEE Access* 8 (2022). (pp. 170183-170194).
- Mueni, C.N., Rugano, P., Njoroge, J.M., & Muriithi, C.G. (2023). Inquiry-based learning and student's self-efficacy in chemistry among secondary school in Kenya. *Heliyon* 9(2023), e12672.
- Olivares Hernandez, A.M., Castellanos Tapia, E., Vega Martinez, A., Urzua Landa, F.A., Vasquez Lopez, V., & Villagomez Guerrero, L.E., (2020).Research-based learning to attract students to control engineering. *In 2020 IEEE Global Engineering Education Conference (CDUCON)* (pp.1451-1457). IEEE.
- Pena, D.M., Dominguez, M., Gomez-Estern, F., Reinoso, O., Torrres, F., & Dormido, S., (2022). Overview and future trends of control education. *IFAC PapersOnline 55-17*, 79-84.
- Rojas-Moreno, A., & Peru, L. (2019). Undergrad research in control engineer. *In 2019 IEEE World Conferece on Engineering Education(EDUNINE)*(pp.1-5). IEEE.
- Subtirelu, G.E. (2015). Simulation versus real world of operational amplifier circuits. *In 2015 9th International Symposium on Advanced Topics in Electrical Engineering (ATEE)* (pp. 828-833). IEEE.
- Tsvetkova, I., & Borodzhieva, A. (2021). Designing and simulation of bistable multivibrators with operational amplifiers in Multisim. *In 2021 29th National Conference with International Participation (TELECOM)* (pp. 145-148). IEEE.
- Vermunt, J.D. (2023). Understanding, measuring and improving simulation-based learning in higher education : student and teacher learning perspectives. *Learning and instrument* 86(2023), 101773.

# Promoting STEM Interest through Empathy and Creative Thinking in a STEM-based Community Service Program

Ma. Jenina N. NALIPAY<sup>a,b\*</sup>, Biyun HUANG<sup>b</sup>, Morris Siu-Yung JONG<sup>a,b</sup>, Ching Sing CHAI<sup>a</sup>, & Eric Tsun-Hin LUK<sup>b</sup>

<sup>a</sup>Department of Curriculum and Instruction, The Chinese University of Hong Kong, Hong Kong SAR, China <sup>b</sup>Centre for Learning Sciences and Technologies, The Chinese University of Hong Kong, Hong Kong SAR, China \*jnalipay@link.cuhk.edu.hk

**Abstract:** Students' interest in science, technology, engineering, and mathematics (STEM) tends to decline as they reach high school. Among the factors to which this decline is attributed are the perceptions about the lack of communality and creativity in STEM. To address this problem, we implemented an 8-week STEM-based community service program where high school students learned about STEM and community service. The students were asked to design possible solutions to problems faced by people from disadvantaged groups (e.g., visually impaired people). Using the post-test data from 275 high school students who participated in the program, we examined a model of empathy predicting STEM interest via creative thinking. Results of mediation analysis showed that empathy predicted STEM interest through creative thinking. The findings provided support for the important roles of empathy and creative thinking in the promotion of STEM interest. This information can be used as a basis for designing STEM-based community service programs for the enhancement of students' interest in learning STEM.

**Keywords:** STEM-based community service program, STEM interest, empathy, creative thinking

# 1. Introduction

Science, technology, engineering, and mathematics (STEM) education has been a globally important educational initiative (Jong et al., 2021; Lau & Jong, 2022; Nalipay et al., 2022); nevertheless, there is declining interest in STEM as students reach high school (Gok, 2022), especially among girls (Valenti et al., 2016). Among the reasons cited for this are the perceptions of lack of communality (e.g., Belanger et al., 2020) and creativity (Valenti et al., 2016) in STEM. STEM fields are often viewed as providing little opportunity for collaboration and helping others (Fuesting et al., 2021). Moreover, traditional STEM teaching tends to employ teacher-centered approaches that do not provide much opportunity for students to imagine, explore, and make mistakes, and hence, students tend to be less creative and engaged in their learning (Amran et al., 2021). To address this, we implemented a crossover program of STEM and community service education (Huang et al., 2022; Huang & Jong, 2020) that provided students with an opportunity to empathize and understand the problems of the people in the community, as well as practicing creative thinking in helping them find a solution to these problems. We proposed that empathy would be associated with creative thinking, and subsequently, interest in learning STEM among the high school students who participated in the STEM-based community service program.

# 1.1 STEM-based Community Service Program

We developed a STEM-based community service program based on the integrated STEMbased community service learning model (Huang et al., 2022). The integrated STEM-based community service learning model illustrates the key elements that should be emphasized to support students' development of creativity (Amabile & Pratt, 2016) and other twenty-first century skills in the design thinking process (Hasso Plattner Institute of Design, 2010). In the 8-week program, students were taught about STEM knowledge and community service, interacted with participants from disadvantaged groups, and designed possible solutions for the problems faced by them (Huang et al., 2022). We hypothesized that among the program participants, empathy would be related to increased creative thinking, and in turn, creative thinking would be associated with greater STEM interest.

# 1.2 Empathy and Creative Thinking

Empathy refers to one's efforts to comprehend other people's actions and motivations, physical and emotional needs, and their meaning-making. It allows designers to get to know and care about the product users and the context in which they live, and thus, create meaningful innovations (Hasso Plattner Institute of Design, 2010). There is evidence that STEM programs that facilitate empathy could promote STEM interest. Burns and Lesseig (2017) found that empathy-based lessons in STEM could increase students' sense of belongingness and interest in STEM careers. We proposed that empathy is linked to STEM interest via creative thinking. Creativity entails the development of new, novel, interesting, and effective ideas that have aesthetic sensibility (Smith & Henriksen, 2016). The National Academy of Engineering (2008) suggested that students must begin to associate the need for creativity and real world problem-solving skills with the opportunities in STEM fields. Previous studies have supported the link between empathy and creative thinking. For example, in an experiment by Anderson et al. (2023), participants who were instructed to be more empathic toward a subject demonstrated greater creativity.

# 1.3 Creative Thinking and STEM Interest

Students are usually highly interested in science and math in the primary grades, but their interest starts to decline as they approach high school. Although this is true for both genders, the decline is more apparent among girls (Valenti et al., 2016). To maintain students' interest in STEM, it has been suggested that STEM education should allow more room for creativity (Wieselmann et al., 2020). The relationship between creativity and STEM interest has been widely supported in the literature. Cooper and Heaverlo (2013) found creativity as a positive predictor of girls' STEM interests. In a study by Knezek et al. (2015), self-reported creative tendencies were found to predict STEM career aspirations in both males and females. These provide support for the likelihood that in the present study, students who demonstrate greater creative thinking would also be more interested in STEM.

# 1.4 The Present Study

In the present study, we investigated a model in which empathy is associated with creative thinking, and in turn, creative thinking is related to STEM interest among students who participated in a STEM-based community service program. Although past studies have mostly focused on the relationships of empathy and creativity to STEM career interest (e.g., Burns & Lesseig, 2017; Knezek et al., 2015), the present study looked into learning interest in STEM, which is more proximal to high school students and indicative of intrinsic motivation (Wigfield & Eccles, 2000). Specifically, we intended to address the following research questions:

- Does empathy predict STEM interest?
- Does creativity mediate the relationship between empathy and STEM interest?

We proposed empathy and creativity as contributing factors to STEM interest. In particular, we hypothesized that empathy would predict STEM interest, and this relationship

is mediated by creativity. The study could contribute valuable information for designing STEMbased community service programs that nurture students' interest in STEM.

# 2. Method

# 2.1 Participants and Procedures

The participants were 275 high school students from three schools in Hong Kong; 52% were boys and 48% were girls; and most were 12-13 years old (76%). They attended the STEM-based community service program described in Section 1.1. For the STEM knowledge aspect of the program, among the key topics covered were coding skills and computational thinking; Internet-of-Things concepts and applications; and smart home devices. For the community services; community housing residents and their needs; and presentation of design solutions and collection of feedback. The participants completed a survey measuring various twenty-first century skills and STEM learning outcomes, including the main variables of the study. The findings presented and discussed in this paper are based on the post-test data that were gathered from the program.

# 2.2 Measures

We used adapted scales to measure the main variables of the study. For empathy, we used a 5-item scale adapted from Vossen et al. (2015). For creative thinking, we used a 5-item scale adapted from Chai et al. (2020). For STEM interest, we used a 4-item adapted scale from (Paul et al., 2020). The three measures were rated from 1 = *strongly disagree* to 6 = *strongly agree*, and demonstrated adequate internal consistency, with Cronbach's  $\alpha$  = .903, .921, and .879, respectively. Schools 1 and 2 were coded 0 (*no*) or 1 (*yes*) (a response of 0 in both means they were from School 3). Gender was coded 1 (*boy*) or 2 (*girl*), while age was coded 1 (*12-13 years*), 2 (*14-15 years*), or 3 (*16-17 years*). We used the Chinese version of the questionnaires.

# 2.3 Data Analysis

To find out whether empathy predicts STEM interest via creative thinking, we performed regression-based mediation analysis using the PROCESS macro for SPSS (Hayes, 2018). Empathy was entered as the independent variable, STEM interest as the dependent variable, and creative thinking as the mediator. We also included schools, gender, and age as covariates. The sample was bootstrapped to 5000, with a 95% confidence interval.

# 3. Results

Table 1 shows the mean, standard deviation, and correlations of the study variables. The main variables empathy, creative thinking, and STEM interest were significantly and positively associated with each other. As for the covariates, those from School 1 tend to have higher levels of empathy, creative thinking, and STEM interest relative to those from other schools. Boys tend to have higher levels of creative thinking and STEM interest, whereas age did not correlate with the main variables of the study. Results of the mediation analysis showed direct ( $\beta = .261$ , p < .001) and indirect ( $\beta = .388$ , p < .001, 95% CI = .302, .476) relationships between empathy and STEM interest via creative thinking (see Figure 1; note: all coefficients reported are standardized). Empathy predicted STEM interest, and creativity mediated the relationship between empathy and STEM interest, hence confirming the hypotheses of the present study.

Tab	Table 1. Mean, Standard Devlation, and Conclations of the Study Vallables								
		Mean	SD	1	2	3	4	5	6
1.	Empathy	4.64	0.89						
2.	Creative thinking	4.36	0.99	.676***					
3.	STEM Interest	4.23	1.06	.672***	.790***				
4.	School 1			.125*	.142*	.174**			
5.	School 2			.110	.088	0.056	608***		
6.	Gender			.032	131*	126*	030	.062	
7.	Age			003	.080	.065	.123*	261***	075
NI-4									

Table 1. Mean, Standard Deviation, and Correlations of the Study Variables

Note: \**p* < .05, \*\**p* < .01, \*\*\**p* < .001

#### 4. Discussion

The study investigated whether empathy would be associated with creative thinking, and subsequently, STEM interest among high school students who participated in a STEM-based community service program. Results showed that empathy positively predicted STEM interest directly and indirectly via creative thinking. This indicates that students who try to better understand the people in the community and their needs and problems tend to be more interested in learning STEM, specifically, design principles of STEM products. Moreover, this relationship can be attributed partially to creative thinking, where they come up with new and diverse ideas and combine different subject knowledge to solve problems.



*Figure 1.* Results of mediation analysis (Note: p < .05, p < .01, p < .001).

The findings align with the idea that empathy is associated with an increased interest in learning STEM, as shown in past studies (e.g., Burns & Lesseig, 2017) and demonstrated that this could be due to creative thinking. Empathizing with the people for whom one designs a product for allows designers to gain a better understanding of their context, views, and needs, and hence, facilitate the creation of more meaningful innovations (Hasso Plattner Institute of Design, 2010). Trying to understand the needs of the people in the community could have facilitated the students to generate ideas to address their problems (Weng et al., 2022). Consequently, coming up with new and different ideas could have made the students more

interested to learn about STEM knowledge which empowered them to solve the problems of the people in the community. This is consistent with past studies linking creativity with greater interest in STEM (e.g., Knezek et al., 2015; Wieselmann et al., 2020). The underlying psychological mechanism is likely to be associated with the self-determination theory (Ryan & Deci, 2020). It is noteworthy that in the model, gender did not significantly predict the STEM interest of the students who participated in the STEM-based community service program ( $\beta$  = -.055, *p* = .129), providing support for the possible roles of empathy and creativity in addressing the gender gap in STEM (e.g., Wieselmann et al., 2020).

The study provided a better understanding of how the integration of STEM and community service learning promotes STEM interest among students through empathy and creative thinking. These findings can be utilized in designing STEM-based community service programs to improve students' interest in learning STEM. The study has some limitations. The use of cross-sectional data prevents cause-and-effect relationships from being inferred, and the use of self-report measures makes it prone to certain biases, e.g., social desirability. Despite its limitations, the study supported the importance of empathy and creativity in the promotion of students' interest in STEM learning.

#### Acknowledgement

The work was substantially supported by The Hong Kong Jockey Club Charities Trust (Project Title: Jockey Club Community Care and STEM in Action [Project S/N Ref: JC 2019/0112]).

#### References

- Amabile, T. M., & Pratt, M. G. (2016). The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior*, 36, 157– 183. https://doi.org/10.1016/j.riob.2016.10.001
- Amran, M. S., Abu Bakar, K., Surat, S., Mahmud, S. N. D., & Mohd Shafie, A. A. Bin. (2021). Assessing preschool teachers' challenges and needs for creativity in STEM education. *Asian Journal of University Education*, 17(3), 99. https://doi.org/10.24191/ajue.v17i3.14517
- Anderson, S., Cameron, C. D., & Beaty, R. E. (2023). Creative empathy. *Creativity Research Journal*, 1–23. https://doi.org/10.1080/10400419.2023.2229649
- Belanger, A. L., Joshi, M. P., Fuesting, M. A., Weisgram, E. S., Claypool, H. M., & Diekman, A. B. (2020). Putting belonging in context: Communal affordances signal belonging in STEM. *Personality and Social Psychology Bulletin*, 46(8), 1186–1204. https://doi.org/10.1177/0146167219897181
- Burns, H., & Lesseig, K. (2017). Infusing empathy into engineering design: Supporting underrepresented student interest and sense of belongingness. 2017 ASEE Annual Conference & Exposition Proceedings. https://doi.org/10.18260/1-2--28529
- Chai, C. S., Liang, J.-C., Tsai, C.-C., & Dong, Y. (2020). Surveying and modelling China high school students' experience of and preferences for twenty-first-century learning and their academic and knowledge creation efficacy. *Educational Studies*, *46*(6), 658–675. https://doi.org/10.1080/03055698.2019.1627662
- Cooper, R., & Heaverlo, C. (2013). Problem solving and creativity and design: What influence do they have on girls' interest in STEM subject areas? *American Journal of Engineering Education*, *4*(1), 27–38.
- Fuesting, M. A., Diekman, A. B., & Bautista, N. (2021). Integrating communal content into science lessons: An investigation of the beliefs and attitudes of preservice teachers. *School Science and Mathematics*, 121(3), 154–163. https://doi.org/10.1111/ssm.12457
- Gok, T. (2022). A comparison of students' attitudes towards STEM at middle school, high school, and vocational high school. *Journal of Science and Mathematics Education in Southeast Asia*, *45*, 184–206.
- Hasso Plattner Institute of Design. (2010). *An introduction to design thinking process guide*. https://web.stanford.edu/\$%5Csim\$mshanks/%0AMichaelShanks/files/509554.pdf
- Hayes, A. F. (2018). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach (2nd Edition). Guilford Press.
- Huang, B., & Jong, M. S.-Y. (2020). Exploring the integration of social care education with STEM: A

social-scientific maker curriculum. 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), 991–994. https://doi.org/10.1109/TALE48869.2020.9368475

- Huang, B., Jong, M. S.-Y., King, R. B., Chai, C.-S., & Jiang, M. Y.-C. (2022). Promoting secondary students' twenty-first century skills and STEM career interests through a crossover program of STEM and community service education. *Frontiers in Psychology*, 13. https://doi.org/10.3389/fpsyg.2022.903252
- Jong, M. S., Song, Y., Soloway, E., & Norris, C. (2021). Editorial note: Teacher professional development in STEM education. *Educational Technology & Society*, *24*(4), 81–85.
- Knezek, G., Christensen, R., Tyler-Wood, T., & Gibson, D. (2015). Gender differences in conceptualizations of STEM career interest: Complementary perspectives from data mining, multivariate data analysis and multidimensional scaling. *Journal of STEM Education*, 16(4), 13– 19.
- Lau, W. W. F., & Jong, M. S. (2022). Typology of teachers' stages of concern for STEM education. *Research in Science & Technological Education*, 1–19. https://doi.org/10.1080/02635143.2022.2064447
- Nalipay, M. J. N., Jong, M. S.-Y., Chiu, T. K. F., & Chai, C. (2022). Teachers' technological pedagogical content knowledge for integrative science, technology, engineering, and mathematics education. 2022 International Symposium on Educational Technology (ISET), 37–41. https://doi.org/10.1109/ISET55194.2022.00016
- National Academy of Engineering. (2008). *Changing the conversation" Messaging for improving public understanding of engineering*. The National Academies Press.
- Paul, K. M., Maltese, A. V., & Svetina Valdivia, D. (2020). Development and validation of the role identity surveys in engineering (RIS-E) and STEM (RIS-STEM) for elementary students. *International Journal of STEM Education*, 7(1), 45. https://doi.org/10.1186/s40594-020-00243-2
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860
- Smith, S., & Henriksen, D. (2016). Fail again, fail better: Embracing failure as a paradigm for creative learning in the arts. *Art Education*, *69*(2), 6–11. https://doi.org/10.1080/00043125.2016.1141644
- Valenti, S. S., Masnick, A. M., Cox, B. D., & Osman, C. J. (2016). Adolescents' and emerging adults' implicit attitudes about STEM careers: "Science is not creative." *Science Education International*, 27(1), 40–58.
- Vossen, H. G. M., Piotrowski, J. T., & Valkenburg, P. M. (2015). Development of the Adolescent Measure of Empathy and Sympathy (AMES). *Personality and Individual Differences*, 74, 66–71. https://doi.org/10.1016/j.paid.2014.09.040
- Weng, X., Chiu, T. K. F., & Jong, M. S. Y. (2022). Applying relatedness to explain learning outcomes of STEM maker activities. *Frontiers in Psychology*, *12*, 800569.
- Wieselmann, J. R., Roehrig, G. H., & Kim, J. N. (2020). Who succeeds in STEM? Elementary girls' attitudes and beliefs about self and STEM. *School Science and Mathematics*, *120*(5), 297–308. https://doi.org/10.1111/ssm.12407
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

# A Preliminary Study on Knowledge Reconstruction Activity for Fostering Cognitive Presence in Online Discussion

Lintang Matahari HASANI<sup>a\*</sup>, Kasiyah JUNUS<sup>b</sup>, Lia SADITA<sup>c</sup>, Tsukasa HIRASHIMA<sup>a</sup>, & Yusuke HAYASHI<sup>a</sup>

<sup>a</sup>Graduate School of Advanced Science and Engineering, Hiroshima University, Japan <sup>b</sup>Faculty of Computer Science, Universitas Indonesia, Indonesia <sup>c</sup>National Research and Innovation Agency, Indonesia \*lintang.matahari97@gmail.com

Abstract: Asynchronous online discussion activity has become a popular form of online collaborative learning (OCL) and has been widely adopted in higher education. In the discussion, learners are expected to effectively co-construct knowledge together through engaging in problem solving discourse. Such an ability to construct and confirm understanding while working together to solve a problem is defined as cognitive presence, which plays a central role in the success of OCL. However, despite various facilitation strategies suggested in prior studies, it is challenging to ensure that learners show sufficient cognitive presence. Providing a preparatory activity that increases engagement with learning materials before participating in the discussion is another strategy that could foster cognitive presence. This study aims to investigate the effectiveness of providing two kinds of knowledge reconstruction activities, namely kitbuild concept mapping (KBCM) and summary writing, as a discussion preparatory activity. The results show that using KBCM for reconstructing basic knowledge related to discussion topic has a potential to be better for fostering cognitive presence compared to summary writing. The KBCM activity provides opportunities to engage more with basic knowledge related to the discussion topic, to confirm understanding, and to familiarize themselves with common terms or keywords used in the discussion through reconstructing a teacher-created map. Some directions for future studies are discussed.

**Keywords:** Mathematics education, concept mapping, online discussion, community of inquiry

# 1. Introduction

The emergence of various online learning technologies has paved the way for the popularity of online collaborative learning (OCL) in higher education. As an alternative to the traditional classroom, collaborative learning has been extensively discussed and promoted in several studies (Laal et al., 2012). In OCL, learners construct knowledge through online collaborative work in which they explore, invent, and apply ways to solve problems through active discourse, usually through an online asynchronous discussion forum (Bates, 2022; Harasim, 2017).

Conducting effective OCL requires an effective framework that guides the design of the activity. Effective OCL is characterized by learners' active participation in OCL processes, such as generating and organizing ideas to achieve consensus about what the problems are, how to solve them, and the solutions (Harasim, 2017). Learners' active participation is important in OCL due to its influence on learning performance (Ng et al., 2022). As a result, learners' inactivity may hinder the effectiveness of OCL (Hew et al., 2010; Mazzolini & Maddison, 2003). Unfortunately, learners' inactivity in online discussion forums has been widely reported in many studies (Fung, 2004).

To avoid or minimize inactivity and realize a successful OCL, a framework that addresses the issue of facilitation or provides facilitation strategies could be implemented. One of the OCL frameworks that addresses this issue and has been widely adopted in online learning is the Community of Inquiry (CoI) framework (Valverde-Berrocoso et al., 2020). This framework defines three elements of an asynchronous text-based OCL that make up learners' educational experience: Social presence, cognitive presence, and teaching presence (Garrison et al., 2001). These presences are the abilities that are exhibited by the participants during the collaborative activity. Social presence is related to social interaction that creates a sense of community among group members. Teaching presence is related to the design and facilitation of the activity. On the other hand, cognitive presence is related to the ability to engage and progress through inquiry steps to solve problems, i.e., to explore, integrate, challenge, and defend ideas.

A successful Col-based discussion is characterized by the active participation of learners in sustained knowledge sharing and building activities, i.e., the existence of a series of agreements or disagreements, a high amount of information sharing related to problems, etc. This is characterized by the existence of social and teaching presences as well as a high level of cognitive presence, because serial monologues with low participation in sharing and challenging ideas are not enough for deeper learning (Garrison & Cleveland-Innes, 2005). However, realizing a successful Col-based discussion is challenging, especially fostering a high level of cognitive presence. There is a possibility that learners are not actively participating in the discourse or producing a low-quality discussion. Simply creating discussion tasks is not enough to encourage idea integration and the discovery of novel insights (Moore & Miller, 2022).

Providing a training as a preparatory activity, which could make the learners aware of what kind of activities they could do to facilitate the discussion and understand the importance of having a sense of belonging as a group before the inquiry process begins, could foster teaching presence and social presence (Junus et al., 2017). Nevertheless, in some cases, providing training alone is insufficient for fostering high cognitive presence. Such training is predicated upon the assumption that participants are expected to provide effective direct instructions to sustain the inquiry process. A recent study showed that the effectiveness of some forms of facilitation to improve cognitive presence is still inconclusive (Moore & Miller, 2022).

Moreover, fostering cognitive presence is more challenging because providing training alone as a preparatory activity is insufficient. In addition, the forum itself, as a mere learning environment, does not automatically foster a high level of cognitive presence (Moore et al., 2019; Sadaf & Olesova, 2017). Based on this issue, there is a need to foster cognitive presence for successful Col-based discussion by improving the preparatory activity.

A prior study suggested that providing individual preparatory activities could improve discussion quality, which is characterized by a higher ability to express and elaborate ideas (Sadita et al., 2020). One possible alternative for improving the preparatory activity is providing certain activities or technologies into the process to make learners more engaged with learning contents (i.e., basic knowledge required for understanding the problem given in the discussion (Saadatmand et al., 2017). An activity in which learners organize information related to the basic knowledge of the discussion topic could be offered as a preparatory activity. Such activity is intended to familiarize the learners with concepts and terms used in the discussion and give them a "bigger picture" of what they are going to discuss through reviewing related learning materials.

Summary writing is one of the strategies to increase engagement with the learning contents related to discussion topic from knowledge generation through reviewing and expressing concepts using learners' own word (Richards et al., 2023). The other alternative is concept mapping, which could increase engagement with relevant materials by constructing an external representation of related knowledge. Several studies suggested that a well-designed external representation, such as concept maps, is useful and could be used for improving the accuracy of problem solving (Hu et al., 2021; Kremer, 1988; Zhang, 1997). A concept map could visually represent important concepts for problem solving that could be more easily retrieved and shared to aid the inquiry process, thus may increase the cognitive
presence. However, skimming the concept map alone is insufficient. Learners need to actively organize information by creating a concept map. In the context of collaborative learning, recreating concept map form given components based on teacher-created map is more effective in increasing problem-related conversations in discussion (Pinandito et al., 2021). Moreover, using a teacher-created map is an effective approach to concept mapping (Willerman & Mac Harg, 1991).

A type of concept mapping known as kit-build concept mapping (KBCM) enables learners to reconstruct concept maps from components based on a teacher-created map, receive immediate feedback, and reflect their state of understanding (Yamasaki et al., 2010). Therefore, it is one of the promising alternatives to be included in preparatory activities for increasing cognitive presence. Moreover, learners could also be provided access to the goal map during the discussion for reference that scaffolds their discussion process, i.e., learners may share the map when encountering difficulties during the discussion because the map is worth sharing.

KBCM has been used for supporting discussion in past studies, e.g., Pinandito et al. (2021). However, whether KBCM as a preparatory activity before the discussion, as well as the provision of a teacher-created goal map as a reference during the inquiry process, is more effective compared to other alternatives needs to be investigated. To address this gap, this study aims at comparing the use of KBCM and summary writing in a Col-based discussion for fostering cognitive presence. The following research question is addressed: How different is the cognitive presence shown in the discussion by learners who did the KBCM activity compared to learners who did summary writing?

## 2. Related Works

## 2.1 Strategies for Fostering Cognitive Presence

One of the challenges for realizing Col-based discussion is preventing inactivity and increasing the quality of discussion; hence, fostering cognitive presence is necessary. Prior studies have proposed several approaches for addressing this issue, namely: (a) facilitation by teacher; (b) facilitation by peers; and (c) designing the activity by adding certain extra activities or certain technologies, e.g., the use of certain tools for aiding the discussion, etc. (Moore & Miller, 2022). Providing facilitation by peers and instructors during the discussion are possible but insufficient.

Despite being the most popular form of facilitation, the effectiveness of instructor facilitation is still inconclusive according to Cho and Tobias (2016). Furthermore, instructor facilitation during the discussion requires the instructor to gain the current state of the learners' understanding and constantly monitor their progress during the discussion, which is very impractical when a large number of learners are involved. In the case of peer facilitation, it is difficult to prepare all participants before discussion to take on the role of facilitator, and there is no correlation between it and some processes that are categorized as cognitive presence, according to a study by Chen et al. (2019). Learners' unpreparedness to be facilitators may also hinder the effectiveness of the facilitation.

Another alternative is providing extra activities or technologies into the discussion activity to make learners more engaged with the contents (Saadatmand et al., 2017). This comprises of activities or the use of technologies before and during the discussion. This study focuses on investigating the effectiveness of a preparatory activity using summary writing and KBCM for fostering cognitive presence.

## 2.2 Kit-Build Concept Map

KBCM is an approach to concept mapping in which learners try to reconstruct a teachercreated map (goal map) using given components without seeing the teacher-created map (Yamasaki et al., 2010). In the process, KBCM enables automatic diagnosis of understanding which provides the learners with feedback on whether they correctly or incorrectly link the given components (Hirashima et al., 2011). Moreover, after the reconstruction activity, the goal map could be provided to be accessed throughout the discussion, which may be shared and referred to when learners need to confirm their understanding of relevant concepts during the discussion. Prior studies have shown evidence that KBCM could improve learning achievements, for example, learners who used KBCM showed better knowledge retention and higher-order thinking in problem solving (Alkhateeb et al., 2015; Nurmaya et al., 2023). KBCM is proposed as a preparatory activity for improving cognitive presence based on the assumption that KBCM could increase awareness of common terms and concepts related to the discussion topic. As a result, it is expected that the learners could productively discuss the given problem, thus achieving a high level of cognitive presence.

## 3. Methods

#### 3.1 Learning Context

Two linear algebra classes (Class A & Class B) consisting of computer science undergraduate students at an Indonesian public university were involved in this study. Both classes consisted of different numbers of students ( $n_A = 70$ ;  $n_B = 20$ ) because the students could freely register for classes based on their convenience as long as the classes were available at the beginning of the academic semester. Some students did not participate fully in the learning activity. Therefore, only a part of the students who did all the necessary activities were included in the analysis ( $n_A = 48$ ;  $n_B = 13$ ). Class A was designated as the experimental (KBCM) group, while Class B was designated as the control (summary writing) group. The second and the third authors worked together in designing the course, developing the test items, and teaching the classes. They are both domain experts in linear algebra.

Both classes received identical learning materials and teaching styles. The learning activities of both classes included lecture sessions and asynchronous online discussion activities offered in the Moodle forum (https://moodle.org/). The learning materials used in the activity related to this study were about vectors and general vector space. The learning goal of this topic was the ability to demonstrate whether a mathematical object could be viewed as a vector through sustained discussion via the forum, in which collaborative knowledge construction is expected to happen. The Col framework was implemented to design the discussion activity to achieve this goal. To conduct the discussion activity, students in the KBCM group were divided into 10 discussion groups, while the students in the summary writing group were divided into three discussion groups. Each group consisted of six to seven participants. Both classes had similar level of basic knowledge regarding the discussion topic at the beginning of the activity.

#### 3.2 Procedures

The experiment conducted in this study took place in a real class environment. The activities undertaken and the different treatments received by the control group and the experimental group in this study are illustrated in Figure 1.



Figure 1. Activities Undertaken in This Study

At the beginning of this experiment, both the KBCM and summary writing groups participated in classroom lecture sessions that introduced them to the concepts of vectors. After attending the class, both groups participated in a session in which a pretest was administered. The pretest served as a baseline for: (a) measuring the basic knowledge about vectors (Test 1), which is expected to be mastered by the students before the discussion; and (b) measuring the procedural knowledge about mathematical proof related to vector space (Test 2), which is expected to be acquired after the discussion (see Subsection 3.3.1).

After completing the pretest, the students in the KBCM group were instructed to individually reconstruct a concept map using given components, while the students in the summary writing group were instructed to write a structured summary about the learning materials. Both groups received identical learning materials. The concept map contains keywords derived from the learning materials, while the summary template provides subtopics as hints for making the summary. Both the concept map and the summary template cover the same learning topics and scope. Figure 2 illustrates the comparison of KBCM and summary writing activities.



Figure 2. Comparison of KBCM and Summary Writing Activity

Afterwards, both KBCM and summary writing groups participated in discussion orientation provided via Moodle forum for practicing Col-based discussion and doing ice-breaking activities as preparation before the main discussion. The discussion orientation is a cognitive apprenticeship training provided to the students to familiarize and train them with Col processes as suggested in prior studies (Junus, 2017; Junus et al., 2019). This is intended to foster their social and teaching presence.

The first posttest was administered after the discussion orientation, in which only the items from Test 1 were used. The main discussion started as soon as the students completed the first posttest. In the main discussion, the teachers started the discussion thread with a triggering question that needed to be collaboratively solved by the discussion groups. The triggering questions given to each discussion group in the main discussion are different, and their cases are not identical with the cases given in the discussion orientation. For example, this triggering question was posed in the discussion orientation: "Suppose you are given the following entity:  $a = (1, 5, 12, \frac{x}{27})$  and  $k, x \in \mathbb{R}$ . What is a?" In contrast, the following is used in the main discussion: "You are given the set C[0, 2] (the set of all continuous functions on the interval [0, 2]). How can you think of a function, such as  $y = 3x + \cos x$ , as a vector? Relate it to what you have learned in the concept map activity."

Moreover, the students from the KBCM group were provided with the teacher map and were explicitly asked to refer to the concept map. In contrast, the students from the summary writing group were also instructed the same but an instructor-created summary was provided instead. From this point, the experimental group could access the teacher map to aid their discussion process. Both the teacher map and teacher-generated summary have identical contents related to the discussion. At the end of the experiment, both groups were instructed to complete the second posttest consisting of Test 2 items.

### 3.3 Data Analysis

#### 3.3.1 Learning Performance: Pretest & Posttest Questions

Two tests were administered to measure students' learning achievement. Test 1 consists of 10 multiple-choice question (MCQ) items that measure students' understanding of basic knowledge about vectors (i.e., how to represent a vector in  $R^2$ , basic arithmetic involving 2 vectors, etc.). The test items were self-developed, and their content was validated by the teachers of the linear algebra classes (the second & the third authors). On the other hand, Test 2 comprises 10 MCQ items measuring students' understanding of procedural knowledge, that is, their ability to analyze the logically valid and correct mathematical proofs showing a mathematical object as a vector. Tests of significant differences between the pretest and posttest scores for Test 1 were used to investigate whether the KBCM group improved their understanding of basic knowledge before conducting the discussion compared to the summary writing group. On the other hand, Test 2 results were used to investigate whether there are differences between the groups in terms of the ability to analyze a mathematical proof at the end of all activities.

#### 3.3.2 Discussion Quality Instrument: Cognitive Presence Coding Scheme

The coding scheme was developed to identify messages that showed cognitive presence based on Garrison et al. (2001) and Junus (2023). In this preliminary study, the coding process was conducted using MAXQDA by a coder who has experience in the Col framework. The final codes were analyzed by obtaining the frequency of cognitive presence messages for each student, and the average was calculated for the KBCM and summary groups. A significance test was conducted on the data to answer the research question.

#### 4. Results and Discussion

#### 4.1 Pretest and Posttest Scores

The pretest and posttest scores for both Test 1 and Test 2 were analyzed. The descriptive statistics for the pretest and posttest scores are shown in Table 1.

Group	Tests	Data	Min	Max	Median	Mean	S.D.
KBCM	Test 1	Pretest	50.00	100.00	90.00	90.42	11.66
( <i>n</i> = 48)		Posttest	40.00	100.00	100.00	96.67	10.59
		Learning Gain	-0.56	1.00	0.00	0.42	0.52
	Test 2	Pretest	0.00	100.00	80.00	66.25	35.05
		Posttest	0.00	100.00	100.00	79.58	36.26
		Learning Gain	-1.00	1.00	0.00	0.25	0.63
Summary	Test 1	Pretest	70.00	100.00	80.00	83.08	11.09
Writing		Posttest	80.00	100.00	100.00	93.85	7.68
( <i>n</i> = 13)		Learning Gain	-0.20	1.00	0.67	0.61	0.45
	Test 2	Pretest	0.00	100.00	50.00	50.77	34.99
		Posttest	0.00	100.00	80.00	60.77	40.92
		Learning Gain	-1.00	1.00	0.25	0.35	0.65

 Table 1. Descriptive Statistics of Pretest and Posttest Scores

Non-parametric tests (Mann-Whitney & Wilcoxon tests) were conducted on the pretest and posttest scores for both Test 1 and Test 2 to compare the learning achievement of KBCM and summary writing groups due to the existence of a non-normal distribution in some data. To minimize the probability of false positives in multiple comparisons, all *p* values from Test 1 and Test 2 data analyses were adjusted using Holm's method. An overview of the results of the non-parametric tests is shown in Figure 3.

As shown in Figure 3, both groups showed the same level of understanding regarding basic and procedural knowledge prior to and after the discussion. Moreover, there is also no significant difference in the learning gain of both groups for Test 1 and Test 2. There is an indication that the large difference in sample sizes between both groups affected this result. This is a limitation of this study due to the experimentation conducted in real-class settings, which made controlling the number of students who enrolled in a class to ensure a similar sample size rather difficult. Thus, whether the treatment contributed to these results is inconclusive. However, despite having the same level of basic and procedural knowledge, the KBCM group showed a different discussion output in terms of cognitive presence.



Figure 3. Pretest and Posttest Scores Comparison

# 4.2 Cognitive Presence Frequency

At the end of the activity, in total, 256 messages were retrieved (237 messages in the experimental group's discussion threads; 19 messages in the control group's discussion threads). In addition, the coding process resulted in 243 messages categorized as posts containing cognitive presence. After the discussion transcripts of all groups were coded, the frequency of messages that contain cognitive presence was calculated for each participant. The means of the frequency from both groups were compared using Mann-Whitney test. The descriptive statistics of the result are presented in Table 2.

Group	Number of CP Messages	Min*	Max*	Median*	Mean*	S.D.*
KBCM ( <i>n</i> = 48)	225	1.00	18.00	4.00	4.69	3.55
Summary Writing $(n = 13)$	18	0.00	3.00	1.00	1.38	0.77

Table 2. Descriptive Statistics of Cognitive Presence Messages

*Note.* \* Statistics presented are based on the number of messages per participants

Both groups showed an identical ratio of cognitive presence messages compared to noncognitive presence messages (95% of messages containing cognitive presence & 5% of messages excluding cognitive presence). This indicates that both groups were focused on discussing the problem. However, further analysis regarding the frequency of cognitive presence messages frequency per participant revealed that the intensity of the discussion differed. There is a significant difference between both groups' frequency of cognitive presence messages per participant, as shown in Figure 4.

The KBCM group significantly had a higher frequency compared to the summary writing group (p < 0.0001). This result showed that, despite having an identical ratio of cognitive presence messages, the KBCM group showed significantly more interaction related to understanding and solving the given problem. Participants in the KBCM group contributed more to the discussion about the problem than the summary writing group. This indicates that they have a better ability to discuss the problem in more detail than the summary writing group. This result supports a prior study that showed KBCM as an effective tool to keep learners focused on discussing the learning contents (Pinandito et al., 2021).



Figure 4. Comparison of the Cognitive Presence Messages per Participant

The characteristic of the instructor-generated concept map as an effective external representation of basic knowledge for the discussion might have contributed to this finding. The discussion required the students to connect various concepts (i.e., the definition of vectors, their axioms, & how to prove certain axioms). By referring to the instructor-generated map during the discussion, the learners could easily resolve misconceptions about core concepts related to the problems they were discussing. In contrast, by referring to the text representation of the instructor-provided summary, the students who did not use KBCM might have more difficulty in finding the connections between important concepts. Prior studies suggested that visual representation of knowledge has the potential to be better than textual representation (Barideaux et al., 2013; Kremer, 1988).

In addition, the characteristics of the KBCM as an activity for reconstructing a teachercreated concept map might also have contributed to this finding. The concept map, as a teacher-provided external representation of knowledge that needs to be reconstructed by learners from given components, provides an opportunity to make the students aware of the correct terms used in discussion (i.e., what a vector is, how it is represented, etc.) better than just seeing an external representation. They could be aware of the common terms used to describe something related to the topics. Even though confusion might occur during the discussion, they could use the same terms to refer to an object when engaging in the discussion. Moreover, having experience in reconstructing the map just before the discussion might reinforce memorization and awareness of what concepts are relevant to the discussion.

A prior study had suggested the effectiveness of KBCM in improving knowledge retention through increasing access to long-term memory (Alkhateeb et al., 2015). This might have resulted in better retention of basic knowledge during the discussion period. However, this assumption needs further investigation as the delay test on the basic knowledge (Test 1) was not yet included after the discussion. As a result of being enabled to be aware of common terms, learners in the KBCM group could resolve differences, reach consensus easily, focus the discussion on how to solve the problem, and explore many possible solutions, instead of being passive due to a lack of clarity about what and how to discuss. In short, by using the same terms and achieving intellectual convergence on basic knowledge before the discussion, reaching a common understanding about what to do to solve the given problem might become easier. This result suggests that KBCM has a potential enable learners to have a more productive problem-solving discussion. In this case, KBCM could provide learners with the opportunity to develop ability to inquire more when conducting collaborative problem solving, which is one of the necessary characteristics that the learners should acquire in higher-education.

## 5. Conclusion

This study compared the use of KBCM and summary writing activities as discussion preparation aimed at increasing learners' cognitive presence and learning achievements (basic & procedural knowledge related to the discussion topic). The results showed that both

experimental and control groups had an identical level of understanding about basic knowledge after the treatment as well as an identical level of procedural knowledge after the discussion. However, those who did KBCM activity showed a significantly higher rate of cognitive presence messages in the discussion due to increased awareness of common terms represented in the concept map. To conclude, KBCM presents a potential for increasing learners' cognitive presence in a Col-based discussion.

However, the result could not be generalized due to several factors that might affect the outcomes. There are some limitations that resulted from conducting the experiment in realclass environments. Firstly, there is a difference in the number of students in both classes due to the university's policy regarding enrollment. Secondly, the experiment is limited to the subject of linear algebra. Thus, conducting the experiment in different educational contexts by involving a near equal number of participants in both the experimental and control groups is needed. Moreover, future studies could also investigate the effect of the proposed treatments on particular practical inquiry processes of cognitive presence (i.e., exploration, integration, etc.).

#### Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP23H01003.

#### References

- Alkhateeb, M., Hayashi, Y., Rajab, T., & Hirashima, T. (2015). Comparison between kit-build and scratch-build concept mapping methods in supporting EFL reading comprehension. *The Journal* of *Information and Systems in Education*, 14(1), 13–27. https://doi.org/https://doi.org/10.12937/ejsise.14.13
- Barideaux, K., Maass, J., & Pavlik, Phil. (2013). A comparison of concept maps and text summaries: The benefits for learning and transfer. *54th Annual Meeting of the Psychonomic Society*. https://www.researchgate.net/publication/258883084\_A\_Comparison\_of\_Concept\_Maps\_and\_T ext\_Summaries\_The\_Benefits\_for\_Learning\_and\_Transfer
- Bates, A. W. (2022). Teaching in a digital age: Guidelines for designing teaching and learning (3rd ed.). Tony Bates Associates Ltd.
- Chen, Y., Lei, J., & Cheng, J. (2019). What if online students take on the responsibility: Students' cognitive presence and peer facilitation techniques. *Online Learning Journal*, *23*(1), 37–61. https://doi.org/10.24059/olj.v23i1.1348
- Cho, M.-H., & Tobias, S. (2016). Should instructors require discussion in online courses? Effects of online discussion on community of inquiry, learner time, satisfaction, and achievement. *The International Review of Research in Open and Distributed Learning*, 17(2). https://doi.org/10.19173/irrodl.v17i2.2342
- Fung, Y. Y. H. (2004). Collaborative online learning: Interaction patterns and limiting factors. Open Learning: The Journal of Open, Distance and e-Learning, 19(2), 135–149. https://doi.org/10.1080/0268051042000224743
- Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education*, *15*(1), 7–23. https://doi.org/10.1080/08923640109527071
- Garrison, D. R., & Cleveland-Innes, M. (2005). Facilitating cognitive presence in online learning: Interaction is not enough. *American Journal of Distance Education*, 19(3), 133–148. https://doi.org/10.1207/s15389286ajde1903 2
- Harasim, L. (2017). *Learning theory and online technologies* (2nd ed.). Routledge.
- Hew, K. F., Cheung, W. S., & Ng, C. S. L. (2010). Student contribution in asynchronous online discussion: a review of the research and empirical exploration. *Instructional Science*, 38(6), 571– 606. https://doi.org/10.1007/s11251-008-9087-0
- Hirashima, T., Yamasaki, K., Fukuda, H., & Funaoi, H. (2011). *Kit-build concept map for automatic diagnosis* (pp. 466–468). https://doi.org/10.1007/978-3-642-21869-9\_71
- Hu, L., Chen, G., Li, P., & Huang, J. (2021). Multimedia effect in problem solving: A meta-analysis. *Educational Psychology Review*, 33(4), 1717–1747. https://doi.org/10.1007/s10648-021-09610-z
- Junus, K. (2017). *Metode pembekalan model community of inquiry dengan cognitive apprenticeship pada forum diskusi online asinkronus* [Doctoral dissertation]. Universitas Indonesia.
- Junus, K. (2023). Pembelajaran kolaboratif daring asinkron. UI Publishing.

- Junus, K., Santoso, H. B., Sadita, L., R-Suradijono, S. H., & Suhartanto, H. (2017). The community of inquiry model training for beginners: Patterns of interaction and student learning strategies. 2017 7th World Engineering Education Forum (WEEF), 343–348. https://doi.org/10.1109/WEEF.2017.8467110
- Junus, K., Suhartanto, H., Suradijono, S. H. R., Santoso, H. B., & Sadita, L. (2019). The community of inquiry model training using the cognitive apprenticeship approach to improve students' learning strategy in the asynchronous discussion forum. *Journal of Educators Online*, *16*(1). https://doi.org/10.9743/jeo.2019.16.1.7
- Kremer, R. (1988). Visual languages for knowledge representation. *KAW'98: Eleventh Workshop on Knowledge Acquisition, Modeling and Management.* http://kremer.cpsc.ucalgary.ca/papers/KAW98/visual/kremer-visual.html
- Laal, M., Laal, M., & Kermanshahi, Z. K. (2012). 21st century learning; Learning in collaboration. *Procedia - Social and Behavioral Sciences*, 47, 1696–1701. https://doi.org/10.1016/j.sbspro.2012.06.885
- Mazzolini, M., & Maddison, S. (2003). Sage, guide or ghost? The effect of instructor intervention on student participation in online discussion forums. *Computers & Education*, 40(3), 237–253. https://doi.org/10.1016/S0360-1315(02)00129-X
- Moore, R. L., & Miller, C. N. (2022). Fostering cognitive presence in online courses: A systematic review (2008-2020). In *Online Learning Journal* (Vol. 26, Issue 1, pp. 130–149). The Online Learning Consortium. https://doi.org/10.24059/olj.v26i1.3071
- Moore, R. L., Oliver, K. M., & Wang, C. (2019). Setting the pace: Examining cognitive processing in MOOC discussion forums with automatic text analysis. *Interactive Learning Environments*, 27(5– 6), 655–669. https://doi.org/10.1080/10494820.2019.1610453
- Ng, P. M. L., Chan, J. K. Y., & Lit, K. K. (2022). Student learning performance in online collaborative learning. *Education and Information Technologies*, 27(6), 8129–8145. https://doi.org/10.1007/s10639-022-10923-x
- Nurmaya, Pinandito, A., Hayashi, Y., & Hirashima, T. (2023). Investigation of the Relationship Between Map Quality and Higher-Order Thinking in Kit-Build Concept Map. In H., Mori & Y. Asahi (Eds.), *Human Interface and the Management of Information. HCII 2023. Lecture Notes in Computer Science* (pp. 363–377). https://doi.org/10.1007/978-3-031-35129-7 27
- Pinandito, A., Prasetya, D., Khudhur, N., Hayashi, Y., & Hirashima, T. (2021). The effectiveness of collaborative concept map recomposition and discussion with kit-build concept map in online learning. *Proceedings of the 29th International Conference on Computers in Education*, 169–174. https://icce2021.apsce.net/wp-content/uploads/2021/12/ICCE2021-Vol.I-PP.-169-174.pdf
- Richards, A. H., Ercan-Dursun, J., Suh, J. K., Hand, B., & Fulmer, G. (2023). Anchoring phenomena and summary writing working together to improve student learning. *Science Activities*, *60*(2), 51–57. https://doi.org/10.1080/00368121.2023.2168242
- Saadatmand, M., Uhlin, L., Hedberg, M., Åbjörnsson, L., & Kvarnström, M. (2017). Examining learners' interaction in an open online course through the community of inquiry framework. *European Journal of Open, Distance and E-Learning*, 20(1), 61–79. https://doi.org/10.1515/eurodl-2017-0004
- Sadaf, A., & Olesova, L. (2017). Enhancing Cognitive Presence in Online Case Discussions With Questions Based on the Practical Inquiry Model. *American Journal of Distance Education*, 31(1), 56–69. https://doi.org/10.1080/08923647.2017.1267525
- Sadita, L., Hirashima, T., Hayashi, Y., Furtado, P. G. F., Junus, K., & Santoso, H. B. (2020). The effect of differences in group composition on knowledge transfer, group achievement, and learners' affective responses during reciprocal concept mapping with the Kit-Build Approach. *Research and Practice in Technology Enhanced Learning*, *15*(1). https://doi.org/10.1186/s41039-020-00133-9
- Valverde-Berrocoso, J., Garrido-Arroyo, M. del C., Burgos-Videla, C., & Morales-Cevallos, M. B. (2020). Trends in educational research about e-Learning: A systematic literature review (2009–2018). *Sustainability*, *12*(12), 5153. https://doi.org/10.3390/su12125153
- Willerman, M., & Mac Harg, R. A. (1991). The concept map as an advance organizer. *Journal of Research in Science Teaching*, *28*(8), 705–711. https://doi.org/10.1002/tea.3660280807
- Yamasaki, K., Fukuda, H., Hirashima, T., & Funaoi, H. (2010). *Kit-build concept map and its preliminary evaluation*. Asia-Pacific Society for Computers in Education.
- Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, *21*(2), 179–217. https://doi.org/10.1207/s15516709cog2102\_3

# Technological Tools for the Teaching and Learning of Statistics

#### Mark Anthony C. TOLENTINO<sup>a,\*</sup>, Juan Carlo F. MALLARI<sup>a</sup>, Maria Alva Q. ABERIN<sup>a</sup>, Ma. Louise Antonette N. DE LAS PEÑAS<sup>a</sup>, Agnes D. GARCIANO<sup>a</sup>, Jumela F. SARMIENTO<sup>a</sup>, Debbie Marie B. VERZOSA<sup>b</sup>

<sup>a</sup>Department of Mathematics, Ateneo de Manila University, Philippines <sup>b</sup>Department of Mathematics and Statistics, University of Southern Mindanao, Philippines \*mtolentino@ateneo.edu

**Abstract:** Statistics and its applications form an integral part of STEM education. In the literature, it is shown that technology is valuable in supporting the teaching and learning of statistics. This paper discusses some technological tools that have been developed to support statistics education in the grade school, junior high school, and senior high school levels. It describes the design and pedagogical basis of these tools, and how these may be integrated in the classroom.

Keywords: Mathematical apps, statistics, STEM, technological tools in statistics

## 1. Introduction

Statistics continues to play an important role in STEM education as it can be applied to all STEM fields (Adams, 2017). In fact, Ben-Zvi and Garfield (2008) have stated how statistics is becoming a necessary area of study that teaches students how to react intelligently to quantitative information. Mastery in statistics allows students to apply different statistics concepts and techniques across STEM projects or within their disciplines when analyzing data generated from investigations (Watson et al., 2020). Thus, statistics is usually featured in the mathematics curricula of different countries and is often included in multiple grade levels (https://timssandpirls.bc.edu/timss2015/encyclopedia/). In the Philippines, for example, the K to 12 Basic Education Curriculum (Department of Education, 2016) includes *Probability and Statistics* as one of the five main content strands.

Given the importance of statistics education, efforts and resources to enhance or to support teaching and learning statistics remain valuable. The use of technology is usually emphasized in this regard; for instance, the Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II) highlights "the recognition that modern statistical practice is intertwined with technology, and the importance of incorporating technology as feasible" (Bargagliotti et al., 2020, p.2). There are several methods on how technology is used in teaching and learning statistics, and how these impact student learning (Chance et al., 2007). Studying statistics in an interactive environment using technology tools supports the understanding of statistical concepts (Prodromou, 2015).

While there are already various technological tools available for statistics education, most existing software (e.g., spreadsheet software, programming language) can be too sophisticated or have irrelevant functionalities for beginner learners. In the context of developing countries such as the Philippines, many of these software may also not be affordable or may be more demanding in terms of device requirements. Thus, we have developed a number of technological tools to support statistics education. In this paper, we discuss pedagogical bases and various features of these tools and how these tools can be concretely integrated into statistics classes.

# 2. Pedagogical Basis

Statistics is often seen to be a branch of mathematics, but it is a distinct field of study requiring both mathematical and non-mathematical skills (Wild et al., 2018). It covers five interrelated themes—variation, expectation, distribution, randomness and informal inference (Watson et al., 2020). Moore (1998, p. 1257) argued that statistical thinking "is a general, fundamental, and independent mode of reasoning about data, variation, and chance." To promote statistical thinking and develop the notion of statistical inference, it is important to provide simulation and randomization-based activities in the classroom.

Rossman and Chance (2014) compiled a list of textbooks that include randomization in the early parts of statistical courses, as well as technological tools that allow students to perform simulations. They also described their own introductory statistics course where students initially perform simulations of random processes using coin flipping, compare their simulations with those of their classmates, and later perform the same simulations using technology. In their curriculum, simulations not only illustrate statistical inference but become the principal mechanism through which they learn to perform statistical inference themselves.

Another important pedagogical consideration, as emphasized in GAISE II (Bargagliotti et al., 2020), is the role of context in statistics. Aside from dealing with the variability of data sets, students should also recognize the context(s) of these data sets so that they can craft interpretations and conclusions appropriately. Hence, to develop such competencies, it is vital that students have access to authentic data sets that are contextually clear and relatable to them. Using such data sets, students can practice procedural skills in statistical analysis and experience contextual learning (Samo et al., 2017).

In designing and developing technological tools, as far as possible, we were guided by Rossman and Chances's (2014) recommended list of features—ease of use, mimicking of byhand simulations, and consistency of user interface. While the first and third features are doable, we had to do away with the second feature because we wanted the tools to run even on modest devices. To manage this limitation, the tools are accompanied by teachers' guides and student worksheets to enable teachers to use the tools more readily. These guides and worksheets are aligned with the most essential learning competencies (MELCs) specified in the official Philippine mathematics curriculum (Department of Education, 2020).

# 3. Technological Tools for Statistics

The technological tools we have developed include apps, applets, a web-based platform, and videos. The apps have been developed using Unity Editor Version 2021.3.25f1 and are available both for Windows and Android devices. The Windows versions can run on 32- or 64-bit systems and have modest requirements: Windows 7 or later, 1 GHz processor or higher, 4 GB RAM, and approximately 300 MB available storage. On the other hand, the Android versions can run on devices with Android 5.1 until 13, 2 GB RAM, and approximately 120 MB available storage. The applets and web-based platform do not require installations and can run on most internet browsers, with only the platform requiring internet connectivity.

# 3.1 Tools for Simulation and Data Generation

GAISE II emphasizes the importance of the context and variability of data sets that are used in statistics education (Bargagliotti et al., 2020). One aspect of this involves giving students opportunities or experiences of dealing with uncertainty and randomness. In the classroom, it is possible to achieve this by having activities in which students perform repeated trials of actual physical probabilistic experiments (e.g., rolling dice, drawing cards). However, in some cases (e.g., remote learning, due to time or resource constraints), these physical experiments may not be feasible. Thus, we have developed the app *Probability Simulator* that allows students to simulate different probabilistic experiments using only their electronic devices (e.g., computer, tablet, smartphone). In *Probability Simulator*, the following probabilistic experiments are available for simulation (Figure 1(a)): rolling a fair die; rolling unfair dice; spinning a number wheel; and drawing 1 to 5 cards from a standard deck. In addition, multi-stage probabilistic experiments can also be simulated by setting up a sequence of different tasks. The student can then specify the number of times that the task(s) must be repeated.



*Figure 1.* Screenshots from *Probability Simulator*. (a) task selection, (b) animated depictions of tasks simulation, (c) outcomes window.

When the student taps/clicks the Simulate Tasks button, a short animation (Figure 1(b)) is played to depict the simulation of the selected tasks. After the animation, the outcomes of the tasks are displayed in a scrollable window (Figure 1(c)). If the number of trials is too many, the scrollable window may become inconvenient for viewing all the outcomes; thus, there is also an option to save the results in a spreadsheet file (csv) that can then be processed or analyzed further using other software. In the following section, we will illustrate how *Probability Simulator* can be used for learning activities across different grade levels.

Another tool that we have developed is the web-based platform *Senso Eskwela Pilipinas* (*SEP*) (De Las Peñas et al., 2020; Tolentino et al., 2022). In *SEP*, students themselves become part of the data collection process by answering the *SEP* survey and having their anonymous responses become part of the *SEP* database. In turn, the database becomes a source of readily available, authentic, and relatable data sets. Based on initial integration into two Grade 11 classes, *SEP* has shown positive benefits in students' learning (Tolentino et al., 2022).

## 3.2 Tools for Data Analysis

In real-world contexts, analyzing different data sets, especially large ones, is usually done using different statistics software. Thus, it is also important to give students opportunities to perform different data analysis activities using technology.

Focusing on competencies for Grades 7, 8, and 10, we have developed a data analysis tool subject to two important considerations. First, we want the tool to focus only on selected functionalities that are relevant to the aforementioned grade levels. This is to make the tool relatively simple and easy to use so that students and teachers are not overwhelmed with many functionalities and complicated commands. Secondly, we want to maximize the accessibility of the tool; that is, it must be freely available, uses relatively low storage, and can run offline in a wide range of low-cost electronic devices.

Given the above considerations, we have developed the app *Mathplus Statistics*, a tool for computing descriptive statistics and generating data visualizations, which can be accessed via Windows or Android. Its functionalities include the generation of histograms (Figure 2(a)), frequency distribution tables, and box plots (Figure 2(b)), as well as the calculation of measures of central tendency and variability and measures of position.

Students may upload their own data (in csv format) or use the app's built-in datasets. The app is also compatible with SEP in the sense that data sets downloaded from the *SEP* platform are already properly formatted and may be uploaded directly to the app. Data filtering is also available and can be applied to data to keep (or exclude) observations based on various variable-dependent criteria.



Figure 2. Screenshots from Mathplus Statistics: (a) histogram, (b) box plot.

In addition to *Mathplus Statistics*, we have also developed two Grade 11 tools or applets using the free software *GeoGebra* (https://www.geogebra.org/): *Mathplus Tool for Normal Distribution* and *Mathplus Tool for t Distribution*. The applet *Mathplus Tool for Normal Distribution* allows students to compute probabilities involving the normal distribution for which the mean and variance are user-inputted (Figure 3(a)). A graph of the normal probability distribution function and the corresponding area under it are also displayed. Lastly, there is also an inverse function (Figure 3(b)) in which a student can input a desired probability then obtain the corresponding value at which the normal cumulative distribution function gives this probability. *Mathplus Tool for t Distribution* has similar functions but for the *t* distribution.



Figure 3. Screenshots from the Mathplus Tool for Normal Distribution applet.

Aside from being an alternative to printed out distribution tables, the applets can also be used for strengthening students' understanding of these probability distributions. Particularly, the visualization of the graphs and probabilities-as-areas can help students master different properties of computing probabilities and of the probability distributions.

#### 3.3 Instructional Resources

In addition to the data generation and analysis tools discussed above, we have also created resources that can contribute to statistics instruction. These have been made according to the list of most essential learning competencies of the Department of Education (DepEd, 2020).

One such resource is the app Just Keep Solving 2 (JKS 2), a game that helps students develop skills in carrying out computations under chosen topics in probability and statistics. As the name implies, JKS 2 is an expansion of an earlier app Just Keep Solving! (Garciano et al., 2023). The Statistics topic in JKS 2 has three levels covering measures of central tendency, measures of variability, and the identification of correct statistical measures in data analysis and interpretation. In the game, questions (Figure 4(a)) are shown as bombs that will explode upon reaching the bottom of the screen (Figure 4(b)). Players tap on these bombs to pause their movement and see the questions they are challenged to answer. Players must answer a set number of questions before losing all lives or answer enough questions to survive within a specific time limit. Garciano et al. (2023) provides additional details of JKS 2.

Some videos have also been produced to complement some of the apps discussed above. The videos have been developed to include a demonstration of an app and a concise discussion of the relevant math lesson(s). The project's mascot *Banoy Bilang* plays the role of a teacher's assistant in most of the videos. A particular example is the *Experimental Probability with Banoy Bilang* video that is intended for Grade 8 students. In the video, the *Probability Simulator* app is used to review the computation of experimental probabilities (Figure 4(c)).



*Figure 4.* Screenshots from *Just Keep Solving 2*: (a) sample question, (b) main level screen; (c) screenshot from the video *Experimental Probability with Banoy Bilang.* 

## 4. Integration of the Technological Tools

A possible obstacle in using any technological tool for statistics education is teachers' "lack of awareness and comfort with new technologies" (Chance et al., 2007, pp. 18-19). Additionally, teachers are often busy enough with their usual teaching responsibilities and other ancillary activities (Batilantes, 2021) to have sufficient time for learning or designing new activities that use technology. At the same time, students also need direction or guidance when it comes to learning activities that they can do independently.

Given the above considerations, we have also developed teachers' guides and student worksheets detailing learning activities that make use of the discussed technological tools. The teachers' guides contain complete information (e.g., competencies covered, materials needed, instructions) related to implementing concrete learning activities. On the other hand, the student worksheets contain instructions for doing the activities; the worksheets are readyto-use and can also serve as students' answer sheets, if applicable.

In Grades 1-3, for example, one set of learning competencies under the *Probability and Statistics* strand (Department of Education, 2020) is focused on introducing and developing students' notion of uncertainty and likelihood. In the teacher's guide *Introduction to Uncertainty* and its corresponding student worksheet, we present a learning activity that makes use of the *Probability Simulator* app. This allows the students to perform probabilistic experiments more practically, compared to using physical objects (e.g., dice, cards) that may be too time-consuming. Other examples of teacher's guides and student worksheets can be accessed for free from mathplusresources.wordpress.com.

## 5. Conclusion and Future Direction

In this paper, we recognize the significance of statistics in students' STEM education and how the teaching and learning of statistics can be improved by using technology. We have developed a number of technological tools that can support different aspects of statistics education. The tools can be used for simulation and data generation (*Probability Simulator* app and *Senso Eskwela Pilipinas*) or for data analysis (*Mathplus Statistics* and *GeoGebra* applets for normal and *t* distributions). There are also resources (the game *Just Keep Solving 2* app and instructional videos) that can support instruction. These resources can easily be downloaded and used on low-cost mobile devices, making them accessible and suitable for the use of most students in the Philippines. We have also developed teachers' guides and student worksheets that present concrete learning activities that use the different technological tools or resources.

Among the activities for future work is Investigating the effectiveness of the tools presented in this paper in terms of increasing student engagement, interest in learning and strengthening statistical knowledge and skills. Another future activity is gathering data on usage and acceptance of the technological tools on the part of students and teachers. Moreover, other ways of integrating the use of these tools in the classroom can be further explored in collaboration with Filipino high school teachers. A greater knowledge of the

requirements for statistics education for both students and teachers will result from such collaboration.

#### Acknowledgements

This paper is one of the outputs of the Ateneo Mathplus Resources Laboratory housed at the Department of Mathematics, School of Science and Engineering, Ateneo de Manila University (ADMU). The authors thank the Department of Science and Technology-Philippine Council for Industry, Energy, and Emerging Technology Research and Development and the University Research Council, ADMU for supporting the development of the resources for Grades 1 to 11 Mathematics. We thank our app developers Logik, Inc, Mr. Victor Antonio M. Ortega, Mr. Nigel Benedict Cargo, Mr. Jose Teodoro Lacson, and our artist Mr. Amiel Damian F. Justiniani.

#### References

- Adams, C. (2017). *The 7 most important STEM skills we should be teaching our kids*. WeAreTeachers. https://www.weareteachers.com/important-stem-skills-teaching-kids/
- Bargagliotti, A., Franklin, C., Arnold, P., & Gould, R. (2020). *Pre-K-12 guidelines for assessment and instruction in statistics education II.* American Statistical Association.
- Batilantes, S. P. (2021). Project VLOGI (Video Lectures on Giving Instructions): Effects on learners' performance in Probability and Statistics. *International Journal of Educational Studies in Mathematics*, *8*(4), 299-315.
- Ben-Zvi, D., & Garfield, J. (2008). Introducing the emerging discipline of statistics education. *School Science and Mathematics*, *108*(8), 355–361.
- Chance, B., Ben-Zvi, D., & Garfield, J., & Medina, E. (2007). The role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education*, 1(1).
- De Las Peñas, M. L. A. N, Verzosa, D. M. B., Sarmiento, J. F., Tolentino, M. A. C., & Loyola, M. L. (2020). Designing mobile apps to promote numeracy and statistical reasoning. In W.-C. Yang & D. Meade (Eds.), *Electronic proceedings of the 25th Asian technology conference in mathematics* (pp. 167-182). Mathematics and Technology, LLC.
- Department of Education (2016). *K to 12 curriculum guide: mathematics.* www.deped.gov.ph/wp-content/uploads/2019/01/Math-CG\_with-tagged-math-equipment.pdf
- Department of Education. (2020). *K to 12 most essential learning competencies with corresponding CG codes.*
- Garciano, A. D., Verzosa, D. M. B., De Las Peñas, M. L. A. N., Aberin, M. A. Q., Mallari, J. C. F., Sarmiento, J. F., & Tolentino, M. A. C. (2023). Practice through play using mobile technology. In P. Kommers, I. A. Sánchez, & P. Isaías (Eds.), *Proceedings of the international conferences on esociety 2023 and mobile learning 2023*. International Association for Development of the Information Society.
- Moore, D. S. (1998). Statistics among the liberal arts. *Journal of the American Statistical Association*, 93(444), 1253–1259.
- Prodromou, T. (2015). Teaching statistics with technology. *Australian Mathematics Teacher*, 71(3), 32-40.
- Rossman, A. J., & Chance, B. L. (2014). Using simulation-based inference for learning introductory statistics. *Wiley Interdisciplinary Reviews. Computational Statistics, 6*(4), 211-221.
- Samo, D. D., Darhim, Kartasasmita, B. (2017). Developing contextual mathematical thinking learning model to enhance higher-order thinking ability for middle school students. *International Education Studies*, *10*(12), 17-29.
- Tolentino, M.A.C., De Las Peñas, M.L.A.N., Aberin, M.A.Q., Garciano, A.D., Loyola, M.L., Sarmiento, J.F., Mallari, J.C.F., & Verzosa, D.M.B (2022). Engaging learners through data: Senso Eskwela Pilipinas. In W.-C. Yang, M. Majewski, D. Meade, & W.K. Ho (Eds.), *Proceedings of the 27th Asian technology conference in mathematics* (pp. 351-362). Mathematics and Technology, LLC.
- Watson, J., Fitzallen, N., & Chick, H. (2020). What is the role of statistics in integrating STEM education? In J. Anderson & Y. Li (Eds.), *Integrated Approaches to STEM Education*. Springer, Cham.
- Wild, C.J., Utts, J. M., & Horton, N. J. (2018). What is statistics? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 5-36). Cham, Switzerland, Springer.

# A Visualization App on Proving Geometric Concepts

#### Ma. Louise Antonette N. DE LAS PEÑAS<sup>a\*</sup>, Debbie Marie B. VERZOSA<sup>b</sup>, Maria Alva Q. ABERIN<sup>a</sup>, Agnes D. GARCIANO<sup>a</sup>, Jumela F. SARMIENTO<sup>a</sup>, Mark Anthony C. TOLENTINO<sup>a</sup> & Juan Carlo F. MALLARI<sup>a</sup>

<sup>a</sup>Department of Mathematics, Ateneo de Manila University, Philippines <sup>b</sup>Department of Mathematics and Statistics, University of Southern Mindanao, Philippines \*mdelaspenas@ateneo.edu

**Abstract:** This paper discusses the description, design and pedagogical basis of a mathematical app called *Two Column Proof*, which provides students a framework for writing proofs of geometric statements. The app focuses on proving concepts on: properties of parallelograms, conditions that determine when a quadrilateral is a parallelogram, properties on trapezoids and kites for high school mathematics. It employs visual representations for students to understand the statement or reason of each given line in the proof, strengthening their logical and mathematical knowledge needed in the proof construction.

Keywords: Mathematical app, two column proof, STEM, mathematical proofs

#### 1. Introduction

The abilities to carry out analytical and critical thinking are necessary for students to succeed in STEM education. In the STEM fields, analytical and critical thinking skills are used to find patterns in data, determine the meaning of these patterns, and then use the data to support a claim (Adams, 2017). Mathematics educators have been recommending increasing the role of proofs in the mathematics education of students at all levels (Stylianou et al, 2009). Moreover, it has been suggested that the role of proofs be opportunities for students to observe patterns, make conjectures and provide justification, not just merely exercises in logical thinking (De Villers, 2003). Geometry is one of the subject areas where mathematical argumentation and proving is taught, developing students' deductive reasoning and critical thinking (Kunimune, Fujita, & Jones, 2010). The various geometric facts and theorems may not have much use to students in all aspects of STEM education, but the critical thinking capability developed through the rigor of mathematical proofs would certainly make them more successful problem solvers. Moreover the concept of proof is used frequently in STEM education, where students need to write reasons and justifications based on evidence in a logical manner and organized manner.

This paper discusses a mathematical application (app) called *Two Column Proof* (TCP), designed and created to support the proving of geometry concepts. Zbiek (2003) emphasizes that the design and appropriate use of technology is essential in developing mathematical intuition, understanding math concepts and proving assumptions. TCP is one of the mathematical resources developed under government-funded projects (De Las Peñas et al., 2022, 2023) with one of the primary objectives to strengthen mathematical learning competencies in students. In engagement in schools under these projects, one of the greatest challenges of teachers is the teaching of geometric proofs. Students have difficulties completing proofs. There is not much student engagement with the topic, and the students have difficulty learning as well as enjoying the learning process.

This paper presents the design and pedagogical basis for the TCP app, where the visualization of geometric representations is integrated into the app with the purpose of helping students understand the proof construction process; allowing them to have a better

understanding of the arguments to use to support statements and claims. Employing visualization technological tools stimulates interest in learning and developing cognitive interest (Andreev et al., 2021). Visualization representations and the role of visualization play an important role in geometry learning (Žakelj & Klančar, 2022). The use of visualization in proofs was found to be effective in discovery learning models (Kristiyajati & Wijaya, 2019). Visualization is a powerful tool to explore mathematical problems, give meaning to mathematical concepts and find relationships between them (Rösken & Rolka, 2006).

The TCP app was designed to run primarily on mobile devices (Android) and desktops (Windows). The Android version can run on devices with Android 5.1 until 13, 2 GB RAM, and approximately 120 MB available storage. The Windows version can run on 32- or 64-bit systems with Windows 7 or later, 1 GHz processor or higher, 4 GB RAM, and approximately 300 MB available storage. In the past two years of the pandemic, the app has been useful in support of the blended learning modality in schools called for by the Philippine Department of Education (DepEd). The app can be used with minimal supervision, and is apt for remote learning. An interested reader is invited to visit the Mathplus website at https://mathplusresources.wordpress.com/ for access to the app and other mathematical resources developed by the authors.

## 2. Scientific Framework and Pedagogical Basis

The design of the *Two Column Proof* app was grounded on the use of visualization and imagery in mathematics, which has drawn the attention of experts in the field of psychology of mathematics education (Presmeg, 2006). Drawing on the extensive body of work on mathematical visualization, Rivera (2011) presented a strong rationale for a visually-oriented mathematics curriculum. An important aspect of this curriculum is the use of diagrams. Pantzaria et al. (2009) found that diagrams were not universally helpful. In fact, the presence of diagrams constrained the range of strategies employed by students when solving problems. However, presenting a problem together with a diagram significantly enabled students with initial difficulties to solve problems successfully. Rivera argued that the main issue is not whether to present diagrams or not, but more importantly, how diagrams are used by the students. Drawing on the idea of diagrams as effective visual scaffolds provided that students use them meaningfully, the app presents a diagram integrated into a feedback mechanism representing a target theorem. More specific details on how the app facilitates effective use of diagrams towards the proof of a theorem are discussed in the next section.

## 3. The Two Column Proof (TCP) App

#### 3.1 App Description

A *two column proof* is a structured way of coming up with a proof of a mathematical result, where in one column, a statement is written, and in the second column a corresponding justification and reason is given for the statement. The two-column proof remains the most common proof format in many secondary school textbooks. The first encounter of students in proving geometry concepts is usually through the two-column proof.

The *Two Column Proof* (TCP) app was designed primarily to help students in writing a geometric proof. The application is divided into three main topics: parallelograms, trapezoids and kites (Figure 1(a)). The app addresses the following learning competencies in the Philippine high school curriculum: i) "determine the properties of a parallelogram"; ii) "determine the conditions that make a quadrilateral a parallelogram"; and "prove theorems on trapezoids and kites" (Department of Education [DepEd], 2020). The first part of the parallelogram sub-menu gives statements pertaining to properties of parallelograms; the second part pertains to the converse of these statements, where a property of a quadrilateral is given and one has to prove the quadrilateral is a parallelogram. For the trapezoid and the kite sub-menus, statements on properties of trapezoids and kites are given respectively. While

writing a proof in the app, the student clicks the "Next" button to proceed to the next Statement/Reason pair. A line of the proof appears, and the student types the answer in the blank provided. To visualize a statement/phrase, the right arrow button is clicked (Figure 1(b)).



Figure 1. (a) Main menu; (b) "Statement" and Reason" columns with the Next/Back buttons

## 3.2 Design

The key element considered in the design of the app is *visualization*. Žakelj & Klančar (2022) defined visualization as the creation, application and reflection of diverse representations. In the process of visualization, the student creates, identifies or shapes visual representations and applies them to solve problems. In the design of TCP, geometric concepts can be visualized with constructed representations that are dynamic in nature. In particular TCP provides visual feedback for a particular line in the proof. Each line in the proof is detailed, in order to depict visually the necessary connections between each one. The visualization element is intended to guide the student in writing the mathematical statements, providing the

Task	Visualization	Example from the app
To write mathematical statements in the "Statement" column where the justification provided in the "Reason" column is the definition of a geometric figure, such as a parallelogram, isosceles trapezoid or kite.	Relevant parts of the geometric figure are highlighted such as pairs of sides or angles. Color flashes are provided as visual hints.	In proving "The diagonal of a kite is an angle bisector of a pair of opposite angles," in the "Reason" column the definition of kite is given. To visualize the properties that distinguish a kite from other quadrilaterals, the animation involves the first pair of sides AB and AD (repeated flashes in blue) (Figure 2(a)); followed by the second pair of sides, BC and DC (repeated flashes in red) (Figure 2(b)). Congruence of sides is hinted at by highlighting the congruence symbols.
To write mathematical statements in the "Statement" column where the justification provided in the "Reason" column is a geometric result a student has been previously taught.	An animation that will illustrate the theorem is shown.	In proving "Opposite sides of a parallelogram are congruent," given in the "Reason" column is "If two parallel sides of a parallelogram are cut by a transversal then alternate interior angles are congruent". The animation first shows one pair of parallel sides AD and BC (Figure 2(c)), followed by a transversal BD (Figure 2(d)) then the pair of alternate interior angles ADB, CBD (Figure 2(e)). The process is repeated for the other pair of parallel sides AB, DC, the transversal BD, then the pair of alternate interior angles ABD, CDB (Figure 2(f)).
To provide the correct triangle congruence theorem or postulate in the "Reason" column that justifies the given statement in the "Statement" column.	Animation is provided on the corresponding parts of both triangles; then the two triangles are highlighted.	For the Angle-Side-Angle (ASA) congruence used in proving a property of a parallelogram "Opposite sides are congruent," an animation is provided in the following sequence: a pair of angles ADB then CBD from two different triangles (Figure 2(g)), an included side (Figure 2(h)), a second pair of angles ABD then CBD of the same two triangles (Figure 2(i)). Then the two triangles are highlighted, one after the other, first triangle ABD followed by triangle CDB (Figure 2(j)).

Table 1. Visualization Representations in the Two Column Proof App

correct reasons to the given statements or understanding the logical order of the proof. We outline in Table 1 examples of visualization or visual representations used in the app.



*Figure 2.* Animation for (a)-(b) definition of a kite; (c)-(f) "If two parallel sides of a parallelogram are cut by a transversal then alternate interior angles are congruent"; and (g)-(j) Angle-Side-Angle (ASA) congruence.

One of the features of TCP is that a student may save the answers to the lines of the proof. An answer key is provided by the app once he/she has finished drawing up a complete proof and saved the work; The student can rework the proof even after viewing the answer key and save a revised version of the proof.

#### 4. Integration and Use of Two Column Proof (TCP) App

TCP can be used as a tool in teaching based on the RAT (Replacement, Amplification, Transformation) framework (Hughes, Thomas, & Scharber, 2006) where the app may replace the conventional method of teaching proofs without affecting the content or the learning objectives associated with proofs or the act of proving. It is common for teachers to write two columns on the board in the classroom to start the process of proving geometric statements. This practice is now replaced by the interface in the app. The task of completing the proof of a theorem remains the same but the static method of presenting them is now changed to the colorful and dynamic animation of geometric definitions and properties. Such features of the app can be used by the teacher as a scaffold to help students understand relationships between parts of the geometric figure.

Additionally, the app's integration can go beyond simple replacement. The app can amplify learning. The app offers a visualization and animation of angles and sides of geometric shapes that can communicate more effectively and clearly the connections between the angles and sides. The correct properties and definitions to employ as justifications in the proof are subsequently identified by the students with the help of these visual clues.

Last but not least, technology may transform teaching strategies and the learning process. The app may encourage teachers to shift from teacher-centric or teacher-controlled instruction to student-centered instruction. As student engagement is enhanced by the feedback and save-answer mechanisms, the learning process may change. To increase their exposure and practice, students can simply repeat and rework the proof. By doing this, they will be better able to comprehend the logical sequence of a proof. They can also compare kinds of proofs and proof strategies such as that of a theorem and its converse by repeatedly going through the lines of the proofs and their corresponding visual images. Also, the app can be utilized as a tool for group work in which students express their ideas and experiences while using it. This kind of activity also encourages collaboration among students and fosters a supportive learning atmosphere. Further, the app can be utilized in a face-to-face or distance learning set-up.

In terms of the advantage and disadvantage of using the Android and desktop versions of TCP: the desktop version has the advantage of letting the students type the proofs into the app with more precision and at times, more speed as the desktop device uses a mouse and a keyboard. The animations would also be noticeably clearer as compared to portable devices with smaller screens. On the other hand, the primary advantage of using TCP in a mobile device is its portability.

In the implementation of the government funded projects (De Las Peñas et al., 2022, 2023) we observed the use of TCP in a Philippine high school in school years 2021-2022 (offsite because of the pandemic) and 2022-2023 (onsite) for selected Grade 9 classes. One of the ways the teachers use the app in class is when introducing the lesson in proving a geometric theorem or when discussing an assigned proving homework. The students found the animations from the app useful during synchronous classes via google classroom or in the classroom when the two column proof was taught for the first time. The app was also helpful for asynchronous classes that followed. The teachers were assured of a way they could check on how the students practiced proving, and they could assign several problems where the students could verify the proofs. A common feedback from the students is that the animations, when carried out repeatedly, were able to guide them in writing down sentences of the proof. Moreover, it was a big help to rely on the feedback of the app at the time when they were not with their teachers.

A study was conducted to assess the effectiveness of the app in the learning of Grade 9 students in a public high school in Quezon City, Philippines during the school year 2022-2023. Two pairs of control-experimental classes were given a pretest and a posttest on the competencies: (i) proves theorems on the different kinds of parallelograms (rectangle, rhombus, square), (ii) proves theorems on trapezoids and kites. The first pair's control class (38 students), where the app was not introduced and only traditional instruction was applied, recorded a mean gain score of 30.94%. On the other hand, the experimental class (40 students), where the app was used in synchronous lessons and in homework, got a mean gain score of 44.90%. The second pair of control and experimental classes recorded a wider gap between mean gain scores: 25.73% for the control (25 students) and 79.29% for the experimental (27 students). These results suggest the potential of the app to help students improve in the mentioned competencies.

## 5. Conclusion and Future Direction

The *Two Column Proof* app was designed as a technological tool to aid students in learning the process of proving geometric statements. Utilizing visualization features, the app shows definitions or geometric statements in a step-by-step animated format, actively involving students in an entire proving task. While the app focuses on specific learning competencies on parallelograms, kites and trapezoids it can also benefit students in other grades who require guidance in constructing geometric proofs. Studies have shown that using visualization tools helps pupils better understand geometric ideas and properties. In the Two-Column Proof app, students are assisted in developing mathematical arguments and constructing proofs in a clear and coherent manner. This support is provided through the app's visual components, which highlight the relationships between geometric objects such as lines, segments, angles, triangles, and quadrilaterals.

As a next step, a more detailed analysis is being carried out on the effectiveness of the app on student learning and competencies with geometry proofs. There are also plans to extend the app to include proofs of other geometric theorems. Adding more animations and interactive elements can help improve further the learning experience of students.

## Acknowledgements

This paper is one of the outputs of the Ateneo Mathplus Resources Laboratory housed at the Department of Mathematics, School of Science and Engineering, Ateneo de Manila University. The authors thank the Department of Science and Technology-Philippine Council for Industry, Energy, and Emerging Technology Research and Development (DOST-PCIEERD) and the University Research Council (URC), Ateneo de Manila University for the support of the

development of the mathematical resources for Grades 1 to 11 Mathematics. Acknowledgement also goes to Dr. Ma. Mercedes T. Rodrigo and the Ateneo Laboratory for the Learning Sciences (ALLS) for providing technical assistance. We thank Mr. Nigel Benedict Cargo and Mr. Amiel Damian F. Justiniani for the development of the app and art assets, respectively.

### References

- Adams, C. (2017). *The 7 most important STEM skills we should be teaching our kids*. WeAreTeachers. https://www.weareteachers.com/important-stem-skills-teaching-kids/
- Andreev, A. N., Emtseva, O. V., Babkina, E. S., Dolzhenkov, V. N., & Kudryashova, N. V. (2021). The effectiveness of using visualization tools and forms in distance learning. *Revista Tempos e Espaços Em Educação*, 14(33). https://doi.org/10.20952/revtee.v14i33.16053
- De Las Peñas, M. L. A. N., Aberin, M. A. Q., Garciano, A. D., Sarmiento, J. F., Tolentino, M. A. C., Verzosa, D. M. B., Mallari, J. C. (2023). *Terminal report (Department of Science and Technology-Philippine Council for Industry, Energy and Emerging Technology Research and Development): Mathematical resources for distance learning utilizing community LTE networks and television frequencies.*
- De Las Peñas, M. L. A. N., Verzosa, D.M.B., Aberin, M.A.Q., Tolentino, M.A.C., Sarmiento, J. F., Loyola, M. & Garces, L. P. D. M. (2022). Terminal report (Department of Science and Technology-Philippine Council for Industry, Energy and Emerging Technology Research and Development): Technology interventions for mathematical reasoning, statistical thinking and assessment.

De Villiers M. D. (2003). Rethinking proof with Geometers Sketchpad. Key Curriculum Press.

- Department of Education. (2020). *K to 12 most essential learning competencies with corresponding CG codes.*
- Hughes, J. E., Thomas, R., & Scharber, C. (2006). Assessing technology integration: The RAT Replacement, Amplification, and Transformation – Framework. In C. M. Crawford, R. Carlsen, K. McFerrin, J. Price, R. Weber, & D. A. Willis (Eds.), *Proceedings of SITE 2006--Society for Information Technology & Teacher Education international conference* (pp. 1616-1620). Orlando, Florida, USA: Association for the Advancement of Computing in Education (AACE).
- Kristiyajati, A., & Wijaya, A. (2019). The effectiveness of visualization of proofs in learning mathematics by using discovery learning viewed from conceptual understanding. *Southeast Asian Mathematics Education Journal*, 9(1), 37-44.
- Kunimune, S., Fujita, T., & Jones, K. (2010). Strengthening students' understanding of 'proof' in geometry in lower secondary school. In V. Durand-Guerrier, S. Soury-Lavergne, & F. Arzarello (Eds.), *Proceedings of the sixth congress of the European Society for Research in Mathematics Education* (pp. 756–765). Lyon: ERME.
- Pantziara, M., Gagatsis, A., & Elia, I. (2009). Using diagrams as tools for the solution of non-routine mathematical problems. *Educational Studies in Mathematics*, *72*, 39-60.
- Presmeg, N. (2006). Research on visualization in learning and teaching mathematics. In A. Gutierrez & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education: Past, present and future* (pp. 205-235). Rotterdam, the Netherlands: Sense Publishers.
- Rivera, F. (2011). Towards a visually-oriented school mathematics curriculum: Research, theory, practice, and issues. Dordrecht, the Netherlands: Springer.
- Rösken, B., & Rolka, K. (2006). A picture is worth a 1000 words the role of visualization in mathematics learning. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proceedings 30th conference of the International Group for the Psychology of Mathematics Education Vol. 4* (pp. 457-464) Prague: PME.
- Stylianou, D. A., Blanton, M. L., & Knuth, E. J. (2011). *Teaching and learning proof across the grades: a K-16 perspective*. Routledge.
- Žakelj, A., & Klančar, A. (2022). The role of visual representations in geometry learning. *European Journal of Educational Research*, *11*(3), 1393–1411.
- Zbiek, R. M. (2003). Using technology to foster mathematical meaning through problem solving. In H.
   L. Schoen, & R. I. Charles (Eds.), *Teaching mathematics through problem solving* (pp. 93–104).
   The National Council of Teachers of Mathematics, Inc.

# AquaFarm Ace: A Game on the Intensive Aquaculture Process of Oreochromis niloticus

## Gerick Jeremiah Niño GO<sup>a\*</sup> & Angela Nicole YAP<sup>a</sup>

<sup>a</sup>Ateneo Laboratory for the Learning Sciences, Ateneo de Manila University, Philippines \*gjgo@ateneo.edu

Abstract: As Aquaculture accounts for 52.4% of the fishing industry in the Philippines, it boasts potential in providing a stable source of food and livelihood. AquaFarm Ace is a 2D Time-Management simulation game on the intensive aquaculture process. The player takes the role of an aquaculturist as they go through the various stages of aquaculture. The game showcases hatchery, transportation, grow-out, and brood stock stages in the aquaculture stages of Oreochromis noliticus. This study aims to contribute to the scholarly community of aquaculture and serious game design. The game was developed using the following tools: Aseprite, beepbox.co and Unity 2020.3.181f and is made available through Unity Play. The researchers selected volunteer participants to play test the game and answer pretest, post-test, and game experiences questionnaires to gather data in relation to the effectiveness of the game. The results show that AquaFarm Ace contributes to raising awareness of aquaculture and can play a key role in highlighting its increasing demand in the aquatic scene in the Philippines to alleviate the socioeconomic problems such as food and job security. To note, play testers retained information about the aquaculture process of O. niloticus after playing the game. As shown on the Game Experience survey, the testers see the game as an enjoyable experience and a potential tool to teach aquaculture.

Keywords: Serious game design, aquaculture, time management games

## 1. Introduction

The Philippines, being an archipelagic state, has an abundance of water resources that spans about 2,000,000km<sup>2</sup> of its territorial area while its land area only covers about 301,000km<sup>2</sup> and a coastline of 17,460km<sup>2</sup>. With this, the Philippines is considered one of the leading producers of fishery products, ranking 7th in the world. Much of its production is consumed locally (Rivera et al., 2002; Romana-Eguia et al., 2022).

As of 2010, the country has been experiencing a decline in fish production due to the exploitation of its water resources by capture fisheries and other fishing grounds (Romana-Eguia et al., 2020). Concurrently, the Philippine aquaculture scene has significantly increased during the last decade and is expected to address the shortfall in fish production as well as issues pertaining to poverty and food insecurity. Due to its dynamic production system, aquaculture has steadily contributed to the fishing industry (Lieberman, 2001; SEAFDEC , 1993).

With the richness of the Philippines' marine resources, aquaculture is the raising of aquatic flora and fauna in areas where they can be harvested (SEAFDEC, 1993). Ayson claims that aquaculture is the fastest-growing food production sector as of 2015. Ayson further expounds that the impacts of aquaculture provide a vital role in providing food sources and income. Aquaculture provides 3.5% of the national GDP of the Philippines. This industry provides about 28.7% of the GDP production to the agriculture GDP (Ayson, 2015; Lieberman, 2001). In 1947, the Bureau of Fisheries and Aquatic Resources (BFAR) was established to facilitate the growth of the aquaculture industry of the Philippines. As of 1997, BFAR ranks the

Philippines as the 3rd biggest producer of seaweed that contributes to 9.3% of the world's seaweed production (Romana-Eguia et al., 2020).

As discussed by Teves and Ragaza, aquaculture contributes to the fish supply that alleviates the problems of overfishing, habitat degradation, and diminishing fish stocks (Teves & Ragaza, 2014). As fish is part of the daily diet of poorer countries in Asia, fish is one of the more widely used sources of animal protein, the demand for aquaculture products provides opportunities for profit (Brummett, 2003).

The research covers Philippine aquaculture with a particular focus on the production process of cultivating aquatic plants and animals as a medium to spread awareness, expand understanding, and encourage positive actions toward Philippine aquaculture. The game encompasses the following: (1) how these processes occur, (2) corresponding challenges to these processes, and (3) the overall production of aquatic resources through aquaculture.

The research focuses on tilapia species, particularly the Nile tilapia *Oreochromis niloticus*. Additionally, tilapia is readily available and relatively easier to market due to its inexpensiveness, international demand, and flexibility to be sold as a product (Teves & Ragaza, 2014; SEAFDEC, 1993). This showcases aquaculture practices and processes of the Philippines in hopes to spotlight the difficulties, challenges, and potential of aquaculture which, in turn, provide a deeper understanding, appreciation, and awareness to aquaculture. This is done by creating a game that presents the aquaculture process and highlights the challenges of fulfilling the demand for fish resources.

This paper will answer the following questions: (1) What aspects of aquaculture in the Philippines should be translated into the game? What game mechanics can be of use? How will these aspects be depicted in the game? (2) How can the game highlight the demand for aquatic resources and raise awareness of the aquaculture processes in the Philippines? (3) How can the game be engaging and fun while providing insights into the state of aquaculture? (4) How effective is the game in teaching players about aquaculture in the Philippines?

## 2. The Aquaculture Process

To supplement the world food production process, animal husbandry and capture fisheries thrive as the main sources of animal protein. However, as the world population increases, these methods become increasingly inadequate for the actual food demand. With this, world aquaculture production rapidly develops to gain recognition as the main source of seafood production. By Food and Agriculture Organization (FAO, n.d.-a) definition, aquaculture is the farming of aquatic organisms such as fish, mollusks, crustaceans, and aquatic plants (Lucas & Southgate, 2012). Unlike capture fisheries, where the seafood is caught in the wild, aquaculture intervenes with the rearing process; dealing with the practice of artificial breeding, raising, and stocking of aquatic animals to enhance seafood production.

In this paper, the researchers identified the general domestication process for intensive finfish aquaculture, specifically the aquaculture of *O. niloticus* tilapia. Intensive farming refers to the type of culture system where it maximizes the stocking rate by fully engineering the growth process of the fish by feeding them with optimal feeds and controlling water nutrients and organic matter within small tanks (Oddsson, 2020). General aquaculture with market objectives follows specific phases such as hatcheries, transporting, grow-out systems, and brood stock management (Funge-Smith & Phillips, 2001).

#### 2.1 Oreochromis niloticus Aquaculture Process

Tilapia is one of the most important fish produced in the Philippines. It is a major component of the Philippines' subsistence fisheries and aquaculture due to its prolific nature. Compared to other culture fishes, domesticating tilapia is a simple and efficient job as it is naturally fast-growing and adaptable, requiring low investment for fish farmers (Gabbadon, 2008).

#### 2.1.1 Hatchery

Hatchery systems are farm laboratories where eggs are kept in a conditioned environment similar to their natural breeding process. As the basic unit of fish husbandry, the fish egg quality is a critical step for successful fish production. Upon arrival from the broodstock unit, a quality control check is performed to ensure the absence of contaminated stocks. The eggs are classified into stages using egg graders, and healthy eggs are weighed and disinfected before incubation (Primavera, 1995).

Incubation tanks are set up with a gentle flow, as Tilapia are mouthbrooders, which means that natural incubation occurs orally. (Morretti et al., 1999). In the rearing tanks, fry undergo sex reversal through administration of male hormonal steroid to prevent uncontrolled reproduction. Proper grading is done to discard abnormal or female fry, and healthy male fish are fed larger powdered feed for 3 weeks before transport to the nursery and grow-out systems (Ali et al., 2020; SEAFDEC, 1993)

#### 2.1.2 Transportation

When fish are being transported from one environment to another, it is important to imitate the previous water conditions the fish were exposed to. This is because fish are susceptible to becoming hyperactive, increasing their respiration rate and metabolic excretion when being transferred (Primavera, 1995). Additionally, aside from physical injuries and shock from handling and packaging, transportation of live fish may also result in progressive changes in the quality of water and fluctuating concentration of chemicals. Any means of transportation can be used when transporting live fish, but special container vehicles are often used by established companies. In most cases, polyethylene (PE) bags, which are filled with water and pure oxygen, placed in insulated polystyrene cardboard boxes act as the specialized containers in special vehicles (FAO, n.d.-b; Morretti et al., 1999). No matter how short or long-distance travels are, it is always required to provide special care and attention to the fish.

#### 2.1.3 Grow-out Systems

High-quality seed stock necessitates rigorous quality control throughout the grow-out process, involving constant monitoring of juveniles in concrete tanks after being conditioned for optimal growth. Fingerlings are regularly fed and sorted by size, with aquaculturists choosing total or partial harvesting based on market demand. Harvesting involves using seines to capitalize on fish migratory instincts while minimizing stress through cooler conditions. The harvested fish are transported to processing facilities or holding cages before reaching marketplaces. After each harvest, tank cleaning and draining are vital to prevent contamination and cannibalism in subsequent spawns (Gabbadon, 2008; Romana-Eguia et al., 2020).

#### 2.1.4 Broodstock Management

A properly managed broodstock unit is necessary to ensure quality seed stock for any aquaculture system. Broodstock refers to the matured fish used for breeding to enhance fish egg quality and quantity. It is important for spawning units to be kept separated from the hatcheries to avoid disturbance and contamination. After carefully selecting the best fish in a group, it is important to check the sex of each fish since there is a prescribed stocking ratio in spawning tanks. For tilapia, the best stocking ratio is 2-3 females for every male. To identify the sex of the fish, gently hold the fish and flip them to their stomach to examine their genitals near the anus. Males have pointed genitals, while females have rounded ones (Eguia, 2007; Romana-Eguia et al., 2020).

Males are often placed inside the tank to aid them to establish their territory prior to defending and attracting the females. Subsequently, female fish lay eggs inside the pot before leaving them to the male fish to be fertilized. After this, the female will pick up the fertilized eggs in her mouth. After spawning, eggs are collected from the female's mouth carefully by hand, opening up the female's mouth and gently dipping her mouth into transport containers,

releasing the eggs. These eggs are then transported to the hatchery units for artificial breeding. After the eggs are harvested, the female broodfish is transferred to a different tank for rest and recuperation. This broodfish can be used in future spawning sessions after 2-3 months and must be cared for regularly (FAO, n.d.-a).

## 2.2 Games on Aquaculture

Among games that feature aquaculture, two games stand out. Both games show promise on the resource management of aquaculture, but lack in showing each step of the process of aquaculture. These games are more focused on the management of the farm than showing the realism of the process of the aquaculture.

First, AquaKultor, dubbed as the first-ever published aquaculture simulation game, has inspirations of resource management games such as Theme Park, Theme Hospital, and Sid Meier's Civilization. The gameplay features fast-paced farming once the fish eggs are placed. The game shows realistic elements of managing the environmental impact as an effect of the process. Additionally, there are minigames inside the main game itself, such as cleaning the nets under a time limit (Evans, 2018).

Aquaculture Land immerses players in a pond management simulation where they begin with a plot of land, a worker, and some initial funds. The game involves regulating various pond components to optimize fish growth and quality. Players can also introduce machines to aid in fish care, and the gameplay includes intricate tasks like employee management, breeding, harvesting, and selling fish, complete with a quest system for selling or fulfilling customer requests. However, it's worth noting that the game is still in early access and lacks certain features, primarily in post-processing, offering extensive and semiintensive fish farming for specific species (Lopez, 2020).

## 3. Design and Development

#### 3.1 Game Mechanics

The game mechanics were designed to provide players with a simplified and engaging experience that mimics real-life aquaculture practices. The developers opted for a time management simulation game style that would allow players to complete tasks quickly and efficiently. It was developed using Unity game engine version 202.3.181f and is made available through Unity Play.

General aspects of aquaculture such as waiting for the fish to grow, regularly feeding the fish, monitoring the fish, and moving the fish around the laboratory are present in all stages of the game. The developers chose to incorporate these to commonly known game mechanics to ensure that players could quickly understand and engage with the game. This approach allowed the developers to focus on integrating educational elements seamlessly into the gameplay, without the added burden of complicated controls or mechanics hindering the player's ability to learn.

To simulate the waiting time involved in real-life aquaculture, the average waiting time in the game has been narrowed down to five to fifteen seconds, allowing players to engage in other tasks. Additionally, repetitive actions such as harvesting, feeding, and monitoring fish are simplified by merely interacting with the game object

Actions are translated into a "hold button" game mechanic to signify the energy expenditure required. Furthermore, complex aspects of aquaculture are simplified using popup windows with a drag-and-drop feature to represent the aquaculturist's movement. Overall, these design choices ensure that players can easily understand and engage with the game, while still incorporating the necessary educational elements.

#### 3.2 Translation of Game Mechanics

The game aims to teach players that aquaculture is a meticulous and time-consuming process, but one that is ultimately efficient and sustainable. Through engaging gameplay mechanics, players will learn about the different stages of aquaculture, the importance of attention to detail, and gain an appreciation for this industry. The incorporation of different aspects of aquaculture to the game are as follows:

At the start of the game, the player must input their name and choose between a male or a female character. This element advocates for gender inclusivity in the world of aquaculture. Upon receiving the message from a quality control officer, the player sees a map of the aquaculture farm and proceeds to start the actual gameplay. There are 5 stages in the game which correspond to the 4 phases of aquaculture, mainly the hatchery, transportation, growout systems, and broodstock selection. Different phases have different tasks and stations around the map. The player is provided with a starter item and must carry out a series of tasks on it through aquaculture processing techniques to create a new item. The goal is to complete as many items to move onto the next phase of aquaculture. At the end of the stage, the final score received will award the player points. Figure 1 shows a visualization of the workflow of the game mechanics for each stage.



Figure 1. Game Mechanics Flowchart

In the Hatchery stage, the player receives a collection of fish eggs from the receiving station. The player must place the eggs in the quality control station where he will be grading the fish eggs for any possible infested batch. A mini-game screen will pop up at this point where the player must quickly identify the contaminated samples to be discarded. The player then disinfects the seedstock. The player then transfers the seedstock to the incubation period for hatching. At this point, the player must wait until the eggs hatch. While managing this, more seedstock arrives for grading and collecting newly-hatched fry. Players then move the fry to the sex reversal station, where they must feed them when a "feed" notification appears. By this period of the stage, the player must be managing multiple seedstock at once. After the sex reversal station, the player transfers the collected fry in the final grading station ready for transport.

In the Transportation phase, the map and stage are divided into two. During the first map, the player must collect the fry in the receiving station and transfer it to the filling station where a meter appears to indicate the water levels. Next, the player must go to the oxygen station and hold the interact button, making sure that the oxygen fills up to only the PE bag. After filling multiple bags, they move to the vehicle. While the player keeps in check with the driver, the player must also monitor several bags at once. A notification pops up indicating problems have risen. When a notification pops up, the player must attend to the bag by pressing the interact button. The player constantly monitors both the driver and the bags until the stage ends.

The Grow-out Systems and Market Selection stage are primarily like the Hatchery stage with some modifications. The player initially takes the PE bags from the receiving station. The player places the fish bags in tanks as it is important to only let the fish out once after conditioning them. Upon conditioning, the player places the fish in their respective grow-out tanks. Different tanks represent different methods of harvesting. In total harvest, the fish will be collected in one single order. In partial harvest, the fish will be harvested one at a time until all have been gathered. There are different containers depending on the tanks that the fish will be stored at, and it is the player's responsibility to match the correct container to its respective tank. After each order, the harvest collection data is logged in the computer system.

In the brood stock unit, the player receives healthy quarantined fish from the receiving station. The player then separates female fish to male fish by playing a mini-game that lets him identify the fish through their genitals. Pointed "V" openings for male, and "U" openings for females. The player places fish in separate spawning tanks according to the sex ratio and waits until the female fish holds her eggs in her mouth. Once it is done, a notification will pop out from the tank indicating that the eggs are ready for collection. These processes are done while the player feeds, like the previous stages, the potential/resting brooders in the other tanks.

## 3.3 Game Aesthetics

The visual graphics of the game were created with the intention of engaging the players and helping them immerse themselves into the aquaculture experience. The design aspect of the game was based on Gameboy Advance RPG games such as *Pokémon, Harvest Moon*, and *The Legend of Zelda*. The game allows the player to control a 2D character from a non-isometric third-person, overhead perspective. The design specification follows an 8x8 pixel tiling scheme with a limited color palette for a clean aesthetic as shown on Figure 2. The idea for the game design was to generate the emotional response of comfort, ease, and nostalgia, and to simplify the complexity of the aquaculture process for better interpretation and retention.



Figure 2. Hatchery Scene.

The game is set in the Philippines and is centered around the Philippine aquaculture of the Tilapia fish. Despite working inside the laboratory, the artistic choice of including the Filipino *salakót* is due to the goal of representing the Filipino culture, especially those who work in the agricultural sector as seen on Figure 2. With the simplification of the actual processes in the game mechanics, the aquaculture facilities, stations, and items were also reduced to follow the 8x8 tiling scheme. All design assets follow the similar color pattern wherein the top portion of each station or item have lighter shades while the bottom portion have darker ones. Light edges give definition to the items to imitate the real-life light shadowing.



Figure 2. Level Selection Scene.

An AquaFarm Ace computer system concept was implemented to simplify some of the processes in aquaculture and to minimize the addition of guide characters. As shown on Figure 3, the developers opted to let a computer system handle most of the communication

and dynamics of the game. The character customization, introduction, tutorials, and level summary all utilize the same computer system theme to give the players information.

maseta martine RECORDS/ ×	□/AQUA ACE FARM >	2 2
REEDROS REEDROS CONCUSION: REEDROS READY FOR TRANSPORT READY FOR TRANSPORT READY FOR TRANSPORT	<text><text><text><text><text><text></text></text></text></text></text></text>	RETRY Next >

Figure 3. Computer System for feedback and tutorial.

# 4. Results and Discussion

### 4.1 Field Test

The research underwent two testing cycles with voluntary testers. The researchers took account the comments and suggestions of testers and have improved results between test cycles. Both testing cycle run on a 30-minute testing cycle. The testers were given a 5-minute pretest, were asked to play the game for 15 minutes and answer posttest and the Game Experience Questionnaire for the rest of the time.

The game was hosted on Unity Play through a WebGL adapter. The playtesting session was done at an schedule agreed upon by the researchers and the tester. The study made use of Google Forms for questionnaires. Each tester was tasked to do the following: answer a pre-test questionnaire, play the game, answer a post-test questionnaire, and answer the game experience survey taken from the Game Experience Questionnaire by IJsselsteijn et al. (2013) after playing the game. The pre-test and post-test questions ask the tester to self-evaluate their knowledge on aquaculture. Furthermore, the posttest asks the following open-ended questions to ask the testers what they recall from the game: (1) What aquaculture concepts have you learned after playing the game?; (2) What about the aquaculture process have you learned after playing the game?; (3) Finish the sentence: "Aquaculture is important because..."

## 4.2 Game Testing Results

The game was evaluated through three questionnaires given to the testers. These questionnaires are a game experience questionnaire, and the pretest and post-test on knowledge. There were a total of 7 testers per testing cycle. Mean was used as the measure of central tendency for the data set of the Game Experience Questionnaire, as per recommendation (IJsselsteijn et al., 2023). For the pretest and post-test questionnaires, median was used as the measure of central tendency due to the ordinate nature of the questions and to avoid outliers that may affect the central tendency.

#### 4.2.1 Aquaculture Knowledge

Tester knowledge self-evaluation of knowledge on aquaculture and its process and importance, particularly aquaculture, and the demand for aquatic resources. Table 1 shows a comparison of the results on tester knowledge of the two testing cycles. The table shows that there is an increase in the self-evaluated knowledge of the testers in both cycles indicated by the higher scores in the post-test results.

knowledge item	cycle 1		cycle 2	
	pretest	post-test	pretest	post-test
aquaculture	4	5	4	5
demand for aquatic resources	4	4	3	5
aquaculture process	3	5	1	5
tilapia culture	2	5	1	5
importance of aquaculture	4	4	3	5

#### Table 1. Comparison of tester self-evaluation on knowledge

The testers were then asked to answer questions on what aquaculture and its process they learned after playing the game. Each tester had different answers which they had recalled from playing the game. These concepts were knowledge of the different stages of the aquaculture process, the segregation of male and female *O. niloticus*, and on maintaining the quality of fish. The open-ended question "Aquaculture is important because..." the testers provided more detailed and more specific feedback on why Aquaculture is important in the posttest. These answers include the effectiveness of fish harvest, the preservation of natural resources, and the importance of aquaculture in food security.

### 4.2.2 Game Experience

The Game Experience Questionnaire encompasses seven distinct components, each designed to gauge different aspects of player experiences. Illustrated in Table 2, the mean score of individual items within these components serves as a measurement of the players' experiences. These components, along with their corresponding items, are as follows: (1) Positive Effect - assesses player enjoyment, (2) Competence - measures the player's eagerness to excel, (3) Sensory and Imaginative Immersion - focuses on the player's sense of wonder within the game, (4) Flow - examines the player's state of mind while actively engaged in the game, (5) Tension or Annoyance - delves into emotions such as anger and frustration directed at the game, (6) Negative Effect - This component evaluates feelings of boredom or disinterest during gameplay, (7) Challenge - explores the player's perception of being tested in terms of their skills. These components are rated on a 6-item Likert scale, with a score of 6 representing the highest level of each aspect (IJsselsteijn et al., 2013). This comprehensive questionnaire offers a multifaceted approach to understanding the diverse dimensions of the player experience.

The testers mentioned that the gameplay, and aesthetics such as the design choice and background music of the game is enjoyable despite the amount of information the testers need to know at the beginning of the game. The visual cues on the icons help them navigate through the game. Additionally, the instructions manual provide aid to the players in providing information that they can refer to on the steps of the process.

Component	mean	mean		
	cycle 1	cycle 2		
Positive Effect	5.17	5.00		
Competence	4.17	4.29		
Sensory and Imaginative Immersion	5.21	4.86		
Flow	4.46	4.40		
Tension and Annoyance	2.14	2.57		
Negative Effects	1.90	2.43		
Challenge	4.14	4.50		

Table 2. Game Experience Questionnaire results

## 5. Conclusions and Recommendations

#### 5.1 Effectivity on Learning

As shown on Table 1, there is an increase in the self-evaluation of testers on their knowledge of aquaculture, the aquaculture process, and tilapia culture for both testing cycles. This shows that there is an improvement in tester knowledge upon playing the game.

The open-ended questions show recall of the topics and concepts introduced in the game. It was observed that the testers of Cycle 2 retained more specific information from the game as compared to Cycle 1 testers.

## 5.2 Game Experience

Results from the Game Experience Questionnaire indicate an enjoyable game experience. Sensory and imaginative immersion, positive effects, flow were the highest categories among the players' experience in cycle 1. The testers were not affected by the negative experiences of the game and were not annoyed while playing the game indicated by below the average scores in the questionnaire.

With both results of negative effect and tension or annoyance scoring below average on the game experience survey and the open-ended comments of the testers, it is concluded that the game provides an engaging and fun experience.

### 5.3 Recommendations

To enhance the player experience, the tutorial system should be made more intuitive and interactive for better information retention. One potential solution to address this issue involves implementing a tutorial run of a level before the player embarks on the first official level. This tutorial serves a dual purpose: it allows players to familiarize themselves with the game's mechanics and, simultaneously, gain an understanding of the aquaculture processes involved. This approach empowers players by providing them with the knowledge necessary to perform various tasks effectively. For instance, consider the hatchery stage. Prior to the timer commencing, a brief guided walkthrough of the game is presented. During this phase, players must follow the UI indicators, which guide them on where to go and what actions to take next. This introductory experience equips players with valuable insights into recurring UI cues that will continue to be relevant throughout the game.

Another recommendation is the addition of a well-defined storyline element. This will better immerse the player into the world of aquaculture as a story can highlight its demand and importance to societal issues, such as food or job security. As an example, the development team has already initiated work on a preliminary storyline. In this narrative, following the main gameplay, players are presented with various neighboring islands grappling with aquatic resource management challenges. The player's role then shifts towards aiding these islands and imparting essential aquaculture knowledge to help them overcome their difficulties. This approach not only enriches the player's understanding of why aquaculture matters but also fosters a more immersive and profound gaming experience. By intertwining the game's objectives with real-world issues, players can appreciate the broader significance of aquaculture in addressing societal challenges, although a thorough research on the stories of the people working in aquaculture is suggested to add more depth and relevance to the game.

Given the game's early stage, it holds potential. Researchers recommend improving the game's communication with the player: including tutorials, storytelling, and refining rewards and consequences to be more impactful. Awarding bonus points for consecutive perfect fish processing, deductions for missed station cooldowns should be considered. Moreover, adding animations and background sound effects can enhance immersion.

### Acknowledgements

The authors extend their heartfelt gratitude to Dr. Maria Mercedes T. Rodrigo and Mr. Walfrido David A. Diy, for their invaluable guidance and unwavering support throughout the authors' research and development journey. Additionally, the authors also thank Dr. Janice A. Ragaza for being the Subject-Matter Expert and igniting the passion for aquaculture in the Philippines, as well as the anonymous testers for their insightful feedback, which broadened the authors' perspectives on AquaFarm Ace. Finally, the authors thank their professors, mentors, family, and friends for their continuous support, encouragement, and influence in making this project a success!

## References

- Ali, M. S., Griffiths, D., & Turner, W. A. (2020). Practical training manual: Tilapia breeding and all-male fry production.
- Ayson, F. (2015). Aquaculture contributes to food security and poverty alleviation. Retrieved from https://www.seafdec.org.ph/2015/aquaculture-contributes-food-security-poverty-alleviation/
- Brummett, R. (2003). Aquaculture and society in the new millennium.
- Eguia, R. V., & Eguia, M. R. R. (2007). Tilapia broodstock and hatchery management. Aquaculture Department, Southeast Asian Fisheries Development Center.
- Evans, O. (2018). AquaKultor review: world's first aquaculture video game is pretty spot on. SalmonBusiness News. Retrieved from https://salmonbusiness.com/aquakultor-review-worldsfirst-aquaculture-video-game-is-pretty-spot-on/
- FAO (Food and Agriculture Organization). (n.d.-a). Fishculture manual. Retrieved from http://www.fao.org/3/ae381e/ae381e01.pdf
- FAO (Food and Agriculture Organization). (n.d.-b). Transporting live fish. Retrieved from http://www.fao.org/fishery/docs/CDrom/FAO\_Training/FAO\_Training/General/x6709e/x6709e14. htm
- Funge-Smith, S., & Phillips, M. J. (2001). Aquaculture systems and species.
- Gabbadon, P. (2008). Tilapia farming: A manual for commercial tilapia production in Trinidad and Tobago. Institute of Marine Affairs. Retrieved from https://www.academia.edu/7105829/TILAPIA\_FARMING\_A\_Manual\_for\_Commercial\_Tilapia\_Pr oduction in Trinidad and Tobago
- IJsselsteijn, W. A., de Kort, Y. A. W., & Poels, K. (2013). The Game Experience Questionnaire. TechnischeUniversiteit Eindhoven.
- Lieberman, D. A. (2001). Management of chronic pediatric diseases with interactive health games: Theory and research findings. The Journal of ambulatory care management, 24(1), 26-38.
- Lopez, C. (2020). Aquaculture Land is a good catch but has some fishues too. Virtual SEA Games from Southeast Asia. Retrieved from https://virtualseasia.com/2020/02/17/aquaculture-land-review/.
- Lucas, J. S., & Southgate, P. C. (2012). Aquaculture: Farming Aquatic Animals and Plants (2nd ed., Vol. 2). Wiley-Blackwell.
- Morretti, A., Fernandez-Criado, M. P., Cittolin, G., & Guidastri, R. (1999). Manual on hatchery production of seabass and gilthead seabream (Vol. 1). Food & Agriculture Organization.
- Oddsson, G. V. (2020). A definition of aquaculture intensity based on production functions—The aquaculture production intensity scale (APIS). Water, 12(3), 765.
- Primavera, J. H. (1995). Mangroves and brackishwater pond culture in the Philippines. Hydrobiologia, 295, 303-309.
- Rivera, R., Turcotte, D., Boyd-Hagart, A., Pangilinan, J., & Santos, R. (2002). Aquatic resources in the Philippines and the extent of poverty in the sector.
- Romana-Eguia, M. R. R., Eguia, R. V., & Pakingking Jr, R. V. (2020). Tilapia culture: The basics. Aquaculture Department, Southeast Asian Fisheries Development Center.
- SEAFDEC (Southeast Asian Fisheries Development Center, Aquaculture Department). (1993). Producing 100% male tilapia. Aquaculture Asia, 11(3), 7. Retrieved from https://repository.seafdec.org .ph/bitstream/handle/10862/2560/AFNv11n03\_p07.pdf?sequence= 1.
- Teves, J. F. C., & Ragaza, J. A. (2014). A Fish Farmer's Role in Sustainable Aquaculture: An Overview of Philippine Aquaculture. World Aquaculture, 11.

# Designing a Training Tool for an Industrial Robot Operating with a Programmable Logic Controller

Porramut Watanakul<sup>a,b\*</sup>, Suppachai HOWIMANPORN<sup>a</sup> & Sasithorn CHOOKAEW<sup>a</sup>

<sup>a</sup>Department of Teacher Training in Mechanical Engineering, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Thailand <sup>b</sup>Taksin Rayong College of Technology, Thailand \*porramut.watanakul@gmail.com

**Abstract:** The industry sector has a complex manufacturing process and technology, mainly utilizing robots and automation, which are necessary for factory operation. Many vocational colleges try to use methods to promote the students' experience of learning automation technology to control robots with a programmable logic controller (PLC). Preparing students who will use technology in Industry 4.0 with robot arm operation is essential. However, the learning tools used to practice based are high-cost, and the technology is constantly changing and evolving. Hence, this research aims to design a training tool for an industrial robot operating using a programmable logic controller (IRO-PLC). The proposal shows that the design training model can simulate steps of the learning activity to control the robot with PLC and then display it on the dashboard with an IoT device. It can increase the student's performance in practical vocational education training in future studies.

Keyword: industrial robot, PLC, learning factory

## 1. Introduction

Nowadays, manufacturing in Industry 4.0 must refer to the digital database connection and take advantage. It is a significant challenge for the industry. Consequently, training students and professionals must address these new demands (Fuertes et al., 2021). It should be absorbed and applied in education, primarily Industry 4.0 learning and 21st-century teaching focusing on sustainable development goals for innovative knowledge and skills (Nurtanto et al., 2020).

Using robots for education has a rapidly increasing rate. Due to the increasing interest in STEM (Science, Technology, Engineering, and Mathematics) education (Mamatnabiyev et al., 2023). Many studies presented that robot learning activities, such as gamified educational robots, improve student achievement and learning motivation (Yang et al., 2023). Furthermore, using students' guidance with web-based personalized learning improves vocational students' performance in industrial robotics (Wanichsan et al., 2021). The use of educational robots has increased rapidly over the past few years due to the growing interest in STEM education. Primarily, the educational robot for practicing the Internet of Things (IoT) using hands-on experiences increased (Mamatnabiyev et al., 2023).

Programmable Logic controllers (PLC) currently form a crucial technological basis for the automation of industrial processes. Hence, industry 4.0 assumed that these controllers would continue to be important and required to a considerable extent for the manufacturing process. (Langmann, et al., 2019). Some study designed of electro-pneumatic controlled color selector robot arm was performed with PLC (ŞAHİN et al., 2020).

The learning material or training tool was developed to solve the significant problems of the students' difficulty in understanding the basic concepts of the lesson and the lack of learning tools in the classroom. It was developed to solve the significant problems of the students' difficulty in understanding the basic concepts of the lesson and the lack of learning tools in the classroom (Abdullah et al., 2021). However, the existing training tool does not meet the course content requirements, and students need to understand due to the limitations of the training tool that unconnected actual situations and learning experiences.

This study focuses on designing a training tool to simulate the technology of controlling industrial robots' operation with PLC in manufacturing process. In the design process, we use a training tool consisting of a Mitsubishi FX5U PLC and a Dobot Magician 4-axis robot, and a belt conveyor.

## 2. Related work

### 2.1 Training of Industry 4.0 technology

Learning factories are understood to be environments that provide practical experiences to these professionals, preparing them in the best way possible for the requirements of Industry 4.0, such as information and communication technology, algorithms, and automation. Industry 4.0 integrates various sectors' digital technologies, automation, data exchange, and advanced manufacturing techniques. It represents a significant transformation in how products are designed, manufactured, and delivered, which are some key aspects of Industry 4.0. (Kipper et al., 2021; Li, 2022; Pozzi et al., 2023). The practical integration of the manufacturing processes and cutting-edge technologies of intelligent manufacturing is important in engineering education and related files (Wang et al., 2021). Therefore, learning industrial robots operating using PLC requires knowledge and practical skills efficiently learning outcomes, as shown in Table 1.

Concepts	Learning outcomes
Programmable logic control	Use programmable logic control hardware.
Ladder programming	Write a PLC ladder program.
Industrial Robot	Use industrial robots in many tasks.
Connect the PLC to the robot	Control the industrial robot to complete the mission with PLC.
IoT dashboard	Display information on the dashboard

Table 1. The learning outcomes of training concepts.

In addition, a dashboard displays essential information such as data, figures, and performance indicators in an easy-to-understand format. It provides generalizations to help users review, analyze, and decide based on the information presented. Node-RED is a development tool for connecting hardware devices. The application interface and the online service have a browser-based editor to enable co-wiring using a variety of nodes in the palette. The industrial flows automation created in the Node-RED tool and a programmable logic controller (PLC) can be stored, imported, and exported for information sharing as a dashboard in the gateway connected with the open platform communications unified architecture (Chookaew & Howimanporn, 2022).

# 3. Industrial Robot Operating System using Programmable Logic Controller

### 3.1 Proposed Training model

In this study, we designed a training tool that emphasizes the practice of students or employees based on Industrial 4.0. The training course contains practical activities that can perform the relevant tasks. To understand how to control robots with programmable logic controllers. We have training activities to carry out and prepare appropriate equipment or materials based on the concept. After that, the data shows on the IoT dashboard, as shown in Figure 1.



Figure 1. Training model framework.

## 3.2 Structure of Training kit

Structural design of the training kit, we have identified the performance of the training equipment that can be compared to the situation of robot control. Figure 2 shows the structure of training kits consisting of PLC (Mitsubishi FX 5 U) as a device that automatically controls machinery and Dobot Magician 4-axis robot arms pick and place the box in the pallet of missions.



Figure 2. The structure of training kits.

## 3.3 Training Activity

We focus on supporting the vocational education students' understanding and practicing with robot control activities using PLC programs. They must work in groups to achieve the tasks that control the robot based on conditions and mission. They can set up the PLC station's structure and connect the conveyor belt and robot arm to pick and place, as shown in Figure 3.



*Figure 3.* Example of the student participating in training activities.

# 4. Conclusion and Future Work

This study proposes a training model consisting of a training tool for a robot control system using a programmable logic controller. Hence, this research aims to design a training tool for an industrial robot operating using a programmable logic controller (IRO-PLC). The proposal shows that the design training model can simulate steps of the learning activity to control the robot with PLC and then display it on the dashboard with an IoT device. It can increase the students' learning achievement who participate in training activities. We plan to develop the training tool and investigate the effectiveness of the proposed. In addition, we will compare the pre-and posttest students' performance. Furthermore, it evaluates attitudes and involvement toward training as part of future work.

# Acknowledgements

The authors would like to acknowledge Taksin Rayong College of Technology (TRCT) and the Division of Mechatronics and Robotics Engineering, Department of Teacher Training in Mechanical Engineering, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok for supporting.

# References

Abdullah, N. S., Rahman, K. A. A., Sumarwati, S., Amiruddin, M. H., Ismail, M., Aziz, A., & Hassan, N. E.
 S. (2021, September). Programmable Logic Controller (PLC) Automation Control Learning Kit
 Programmed: Using The CX-Programmer Software. In 2021 4th International Symposium on
 Agents, Multi-Agent Systems and Robotics (ISAMSR) (pp. 119-123). IEEE.

- Chookaew, S., & Howimanporn, S. (2022). Upskilling and reskilling for engineering workforce: implementing an automated manufacturing 4.0 technology training course. Global Journal of Engineering Education, 24(1), 34-39.
- Fuertes, J. J., Prada, M. Á., Rodríguez-Ossorio, J. R., González-Herbón, R., Pérez, D., & Domínguez, M. (2021). Environment for education on industry 4.0. IEEE Access, 9, 144395-144405.
- Kipper, L. M., Iepsen, S., Dal Forno, A. J., Frozza, R., Furstenau, L., Agnes, J., & Cossul, D. (2021). Scientific mapping to identify competencies required by industry 4.0. Technology in Society, 64, 101454.
- Mamatnabiyev, Z. (2023, June). Design and Implementation of an Open-Source Educational Robot for Hands-On Learning Experiences in IoT. In 2023 17th International Conference on Electronics Computer and Computation (ICECCO) (pp. 1-4). IEEE.
- Nurtanto, M., Fawaid, M., & Sofyan, H. (2020, July). Problem based learning (PBL) in Industry 4.0: Improving learning quality through character-based literacy learning and life career skill (LL-LCS). In Journal of Physics: Conference Series (Vol. 1573, No. 1, p. 012006). IOP Publishing.
- Langmann, R., & Stiller, M. (2019). The PLC as a smart service in industry 4.0 production systems. Applied Sciences, 9(18), 3815.
- Li, L. (2022). Reskilling and upskilling the future-ready workforce for industry 4.0 and beyond. Information Systems Frontiers, 1-16.
- Pozzi, R., Rossi, T., & Secchi, R. (2023). Industry 4.0 technologies: critical success factors for implementation and improvements in manufacturing companies. Production Planning & Control, 34(2), 139-158.
- ŞAHİN, H., GÜNTÜRKÜN, R., & Osman, H. I. Z. (2020). Design and Application of PLC Controlled Robotic Arm Choosing Objects According to Their Color. Electronic Letters on Science and Engineering, 16(2), 52-62.
- Wanichsan, D., Pitanon, K., & Chookaew, S. (2021, November). Design Web-based Personalized Environment for Industrial Robots Learning. In 2021 29th International Conference on Computers in Education (ICCE). (pp. 643-648).
- Wang, S., Jiang, L., Meng, J., Xie, Y., & Ding, H. (2021). Training for smart manufacturing using a mobile robot-based production line. Frontiers of Mechanical Engineering, 16, 249-270.
- Yang, Q. F., Lian, L. W., & Zhao, J. H. (2023). Developing a gamified artificial intelligence educational robot to promote learning effectiveness and behavior in laboratory safety courses for undergraduate students. International Journal of Educational Technology in Higher Education, 20(1), 18.

# Designing a Sorting System using Machine Vision Training Kit for Mechatronics and Robotics Engineering Students

Pakorn MUANGSUK <sup>a,b\*</sup>, Suppachai HOWIMANPORN <sup>a</sup> & Sasithorn CHOOKAEW <sup>a</sup>

<sup>a</sup>Department of Teacher Training in Mechanical Engineering, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Thailand <sup>b</sup>Ayutthaya Technical College, Thailand \* pakorn.mk@gmail.com

**Abstract:** Machine vision is even more critical for sorting workpieces in quality control of automated production systems for Industry 4.0. Use the image processing system principle to inspect the workpiece with a high-speed vision sensor camera. It results in high accuracy of sorting performance and can replace labor in automated production processes. The learning process of vision sensors used for inspecting workpieces with industrial cameras for mechatronics and robotics engineering students is concerned. Most barriers need more teaching materials due to the relatively high cost of the training kit. That is causing students' lack of knowledge and practical skills during the professional training experience in the workplace. This study proposed designing a sorting system using a machine vision (SSMV) training kit for mechatronics and robotics engineering students. It can be a learning tool to promote the students' learning in related fields, automated manufacturing systems, and engineering education.

Keywords: Machine Vision, Industry 4.0, engineering education

# 1. Introduction

Current technology machine vision will have many of these tasks in production by quality control to inspect the workpiece. Sorting workpieces are a topic in the production process with automatic control systems to sustainably increase the country's competitiveness and elevate the country's industry into Industry 4.0 (Eduardo et al., 2020). An automated control system for sorting people with machine vision uses a programmable logic controller to centralize automation (Li, 2020; Li et al., 2023). In addition, the sorting system with machine vision data results are displayed in real-time on the dashboard. At present, many members are often referred to as open (OPC-UA) on the Node-Red dashboard (Chookaew & Howimanporn, 2022). Many studies have found that learning the factory concept appropriate for engineering education, especially teaching, and learning about vision sensors, can develop the students' competency by learning from the training kit (Louw & Droomer, 2019). Penumuru et al. (2020) propose an automated material identification method with machine vision and machine learning algorithms to enhance the cognitive abilities of machine tools and material handling devices in Industry 4.0.

This study proposed designing a sorting system using a machine vision (SSMV) training kit for mechatronics and robotics engineering students. It can be a learning tool to promote the students' learning in related fields, automated manufacturing systems, and engineering education.
## 2. Related Works

#### 2.1 Machine vision

Machine vision or computer vision is a field of artificial intelligence (AI) and computer science that focuses on enabling machines to interpret and understand the visual world. It involves the development of algorithms and technologies that allow computers to process and analyze visual information, such as images and videos (Javaid et al.,2022). Machine vision is significant in sorting workpieces for the quality control of automated production systems for industry 4.0 technology (Li et al., 2023; Chookaew et al., 2022). Machine vision application in the industrial sector is talent learning and training in colleges and universities for educator design learning process (Chaoqun et al., 2023). Benbarrad et al. (2021) presented using machine vision for predicting the most suitable parameters of production processes to obtain a defect-free item by identifying the defective products and continuously improving manufacturing processes based on quality management requirements in Industry 4.0.

#### 2.2 Learning outcome

We analyzed the curriculum's learning objectives for mechatronics and robotics engineering students and requirements from the industrial sector about applying industrial technology and devices, especially vision sensors, for inspecting workpieces with industrial cameras. Hence, learning outcomes are essential to measure student achievements after completing the learning activity. In this study, the learning outcomes aligned with the concept of a sorting system using a machine vision training kit for students are shown in Table 1.

Table 1. The learning buccon	
Concepts	Learning outcomes
1. Machine Vision system	Able to use a vision sensor for inspection to sort materials.
2. Programmable Logic	Able to use PLC to control the operation of hardware devices.
Controller (PLC)	
3. Ladder programming	Able to write ladder logic programming for operating with PLC.
4. Pneumatic cylinders	Able to write programming to control pneumatic cylinders for
	sorting materials.
5. Dashboard	Able to display the data or information on the dashboard with
	Node-RED

#### Table 1. The learning outcomes

## 3. Designing the SSMV training kit

#### 3.1 Elements of Training

This study focuses on designing the training processes to foster students' understanding of the principle of machine vision and the operation of sorting object systems. Figure 1 show the process diagram of the SSMV consisting of three elements as follows:

- *Input:* This phase is first in the production line that uses a vision sensor to inspect types of material with a machine vision system. The student can understand and use devices to input image data.
- *Process:* This phase is process control. We employ the programmable Logic controller (PLC) model Omron NX1P2 to process the sorting system machine vision between input and output. The student can program to control the sorting system.
- *Output:* This phase shows the data output after the complete process. A Node-RED dashboard receives data from PLC (Omron NX1P2) to display the visual inspection data on a monitor or a web browser. The student can show all data on a mobile device.



Figure 1. The process diagram of the sorting system using machine vision.

## 3.2 Design of training kit

Figure 2 shows the structure of the machine vision training kit, we have designed the machine vision process based on factory learning concept. This training kit consist of:

- A structure body is an aluminum profile, conveyor, storage materials, storage OK materials, and storage OK Materials.
- Switch control is four switches (Emergency switch, stop switch, start switch, and reset switch) for control machine vision training Kit.
- The sorting unit consists of Cylinder 1 for sorting the OK part and Cylinder 2 for sorting the NG part.
- The workpiece detection unit uses a vision sensor (Keyence IV-500CA).
- Programmable Logic Controller (PLC) (Omron NX1P2) is an element that automatically controls the sorting system using machine vision.



Figure 2. The structure of the machine vision training kit

#### 3.3 Dashboard

A dashboard is a user interface that provides a consolidated visual representation of various data and metrics. It is designed to give users a quick and comprehensive overview of the state of an industrial process that involves data-driven insights. In this study, we use a Node-Red tool then create a dashboard that displays the number of inspections of 3 types of workpieces from the vision sensor. Each class will separate the characteristics of good (OK), defect (NG), and the total number of workpieces in each sorting. However, it allows students to study and analyze the data to determine the efficiency of the SSMV (see Figure 3).



Figure 3. The screen snapshot of the Node-RED dashboard.

## 4. Conclusion and Future work

This paper presents designing a sorting system using a machine vision (SSMV) training Kit for mechatronics and robotics engineering students. This proposed application of machine vision to visual inspection material for automatic sorting is implemented for mechatronics and robotics engineering students based on factory learning. In addition, to encourage students to master using vision sensors, PLC programming skills, and building dashboards on a Node-Red tool. To ensure that this learning kit can improve the students' learning achievement and motivation, we plan to implement it with the students in future study.

## Acknowledgements

The authors thank the Ayutthaya Technical College and the Division of Mechatronics and Robotics Engineering, Department of Teacher Training in Mechanical Engineering, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok for support.

## References

- Benbarrad, T., Salhaoui, M., Kenitar, S. B., & Arioua, M. (2021). Intelligent machine vision model for defective product inspection based on machine learning. *Journal of Sensor and Actuator Networks*, 10(1), 7.
- Chaoqun, L., Zhenhua, C., & Jie, L. (2022, November). Exploration of Machine Vision Curriculum Construction Facing on Industry Demand and Talent Training Demand of Higher Education. *Proceedings of the 37th Youth Academic Annual Conference of Chinese Association of Automation (YAC)* (pp. 920-924). IEEE

- Chen, S. (2021). Research and Practice of Industrial Sorting Technology Based on Machine Vision. *Proceedings of* IEEE 6th *International Conference on Intelligent Computing and Signal Processing (ICSP* 2021), (pp.1075-1078).
- Chookaew, S., & Howimanporn, S. (2022). Upskilling and reskilling for engineering workforce: implementing an automated manufacturing 4.0 technology training course. *Global Journal of Engineering Education*, 24(1), 34-39.
- Chookaew, S., Raijaidee, P., Khanthinthara, W., Howimanporn, S., & Sootkaneung, W. (2022, November). Learning Factory: A Proposed Framework for Engineering Learning Ecology by Automated Manufacturing System Kits. In 2022 30th *International Conference on Computers in Education (ICCE)*, (pp. 352–357).
- Eduardo, Z., Henrique O. Martins., Fernando P. Lopes., Fernando A.T.V. da Silva Neto., (2020). Machine Vision applications in a Learning Factory. Procedia Manufacturing, 45, 516-521.
- Howimanporn, S., Chookaew, S., & Silawatchananai, C. (2022). Implementation of Real Time Data Collection Process Automation Control Using IIoT Applications. *Journal of Advances in Information Technology*, 13(2), 117-124.
- Jiang, C., Wan, Y., Zhu, Y., & Wang, R. (2022). Machine Vision Algorithm Training Course Construction with PBL. International Journal of Information and Education Technology,12(10), 1050-1055.
- Javaid, M., Haleem, A., Singh, R. P., Rab, S., & Suman, R. (2022). Exploring impact and features of machine vision for progressive industry 4.0 culture. *Sensors International*, 3, 100132.
- Konstantinidis, F.K., Mouroutsos, S.G., & Gasteratos, A. The Role of Machine Vision in Industry 4.0: An automotive manufacturing perspective (2021). ST 2021 – IEEE International Conference on Imaging Systems and Techniques, Proceedings of Conference on Mechatronics and Automation, (ICMA), (pp.1075-1079).
- Kumar, R., Patil, O., Nath S, K., Sangwan, K.S., & Kumar, R. (2021). A Machine Vision-based. Cyber-Physical Production System for Energy Efficiency and Enhanced Teaching-Learning Using a Learning Factory. Procedia CIRP, 98, 424–429.
- Laszlo, R., Holonec, R., Copîndean, R., & Dragan, F. (2019). Sorting System for e-Waste Recycling using Contour Vision Sensors. *Proceedings of 2019 8th International Conference on Modern Power Systems, MPS 2019*, 8759739.
- Li, J. (2020). Application Research of Vision Sensor in Material Sorting Automation Control System. IOP Conference Series: *Materials Science and Engineering*, 782(2), 022074.
- Li, J., He, M., Su, J., Wang, B., & Li, Z. (2023, May). Design and Implementation of Machine Vision Experiment Platform for Virtual Production Line. In 2023 9th International Conference on Virtual Reality (ICVR) (pp. 452-460). IEEE.
- Louw, L., & Droomer, M. (2019). Development of a low cost machine vision based quality control system for a learning factory. *Procedia Manufacturing*, 31, 264-269.
- Penumuru, D. P., Muthuswamy, S., & Karumbu, P. (2020). Identification and classification of materials using machine vision and machine learning in the context of industry 4.0. *Journal of Intelligent Manufacturing*, 31(5), 1229-1241.
- Wang, X., Chen, J., Li, C., & Qi, Z. Song, L. (2022). Design of Intelligent Logistics Delivery Experiment Teaching System based on Machine Vision. *Proceedings of the International Conference on Mechatronics and Automation (ICMA)* (pp.1075-1079).
- Yang, H., Hu, G., & Lu, L. (2022). Research on an Automatic Sorting System Based on Machine Vision. *Proceedings of International Seminar on Computer Science and Engineering Technology,* (pp. 27-31).

## Computer-Supported Collaborative Work in Academics During the COVID-19 Pandemic

#### Bjorn Svetlana NG<sup>a</sup>, John Michael CALVARA<sup>b</sup> & Judith AZCARRAGA<sup>c\*</sup> <sup>a,b,c</sup>De La Salle University, Philippines \*judith.azcarraga@dlsu.edu.ph

**Abstract:** Computer-Supported Collaborative Work (CSCW) has been present as early as in the 1980's and is still being used even during the COVID-19 pandemic. This is mainly because it allows people to effectively use different collaborative platforms to work together and become more productive in their work. This study explores how CSCW affected the academic performances and group work of students during the COVID-19 pandemic. A total of forty (40) senior high school students under STEM in a private school participated in this study. Results have shown that collaborative tools enabled remote learning, allowing students to participate in group projects, discussions, and access resources regardless of their location. Students developed important collaboration skills and gained experience in communication and problemsolving. These tools also provided support and guidance from professors to students transitioning to remote learning. Moreover, CSCW tools have helped instructors assess students' understanding, offer additional resources, and provide feedback on assignments.

Keywords: collaborative work, online learning, COVID-19

## 1. Introduction

Computer-Supported Collaborative Work (CSCW) provides a means for students and even professionals to work collaboratively across different regions and countries. CSCW allows users to share the workload and enhance teamwork, thereby increasing productivity. It provides an ideal work environment for remote workers due to its convenience and flexibility. The demand for CSCW surged during the COVID-19 pandemic, benefiting the education sector by enabling teachers to conduct classes remotely (Talan, 2021.) CSCW's flexibility allows users to contribute collaboratively at their preferred time, pace, and location. The pandemic has prompted a massive increase in the use of collaboration tools in various work settings (Bullinger-Hoffmann et al., 2021.)

CSCW utilizes collaborative technologies like Zoom, Google Docs, and Canva to support individuals, groups, and organizations in working together on projects and research (Goff, 2022.) It is utilized in various fields such as medicine, accounting, business, multimedia, and education.

This study reviews different Computer-Supported Collaborative Work platforms for students, examines their utilization during the COVID-19 pandemic, and evaluates their effects on academic performance and the learning experience. Forty (40) Senior High School students under STEM Strand in a private school participated in the study. The study has evaluated the effectiveness of the identified tools by gathering the respondents' opinions on changes in their academic performance and collaborative work.

Cloud computing tools have emerged as valuable resources during the COVID-19 pandemic, facilitating remote work and enabling the storage and accessibility of various forms of digital data. Studies have explored the relationship between technology and work performance, emphasizing the importance of designing information and communication technologies (ICTs) based on task-technology fit principles to enhance productivity and

alleviate negative psychological outcomes among individuals working from home. Collaborative software tools have played a significant role in supporting remote work, fostering productivity, and maintaining a sense of belonging and work culture (Bullinger-Hoffmann et al., 2021.) Several organizations and institutions have shifted to online meetings, video conferences, online collaboration and online teaching where they have learned to adapt to new work and learning environments. Furthermore, studies have explored the impact of CSCW in education, with a focus on the opportunities and challenges associated with cloud computing tools, as well as the effectiveness of collaborative platforms in facilitating online collaborative learning and interaction among students and teachers (Faja, 2013.) Collectively, these findings explore the potential of cloud computing tools and online collaborative environments in strengthening information sharing, accessibility, and student engagement in various educational and professional settings, particularly during times of crisis such as the COVID-19 pandemic (Al-Samarraie et al., 2018.)

## 2. Related Works

## 2.1 CSCW During the COVID-19 Pandemic

During the COVID-19 pandemic, several studies explored the role of CSCW in academics and remote work. The study in Abelsen et al. (2021) focused on the relationship between technology and work performance. The paper suggests how designing information and communication technologies (ICTs) based on task-technology fit principles improved performance and reduced feelings of loneliness among individuals working from home. Vasilijevic et al. (2021) highlighted the significant impact of collaborative software tools in facilitating remote work, promoting productivity, and maintaining a sense of belonging and work culture. CSCW concepts and theories were reviewed by Bullinger-Hoffmann et al. (2021) emphasizing coordination support as a mechanism for efficient and reliable cooperation. Organizations have employed digital transformation during the pandemic, utilizing video conferences, digital collaboration, and online teaching and learning, which enhanced collaboration and fostered a unified campus experience. Additionally, Alharbi et al. (2021) explored the challenges faced by older adults due to pandemic-driven technology, with a focus on the usability and functionality of collaborative platforms. These studies collectively underscore the importance of task-technology fit, collaborative software tools. coordination support, and addressing the specific needs of diverse user populations in facilitating effective CSCW during the pandemic.

## 2.2 Online Collaborative Learning Environments

Cloud computing tools are valuable for storing and accessing digital data, ranging from simple text files to videos, serving as online libraries for both public and private use. Al-Samarraie et al. (2018) conducted a comprehensive review of several sources, categorizing cloud computing tools into Social Networking, Synchronized, and Learning Management Systems. The study outlined the opportunities and challenges associated with these tools, enabling a better understanding of their usage. Biasutti (2017) investigated the effectiveness of CSCW tools using Moodle, an online course management system. Engaging 87 students in asynchronous activities over five weeks, the study explored interdisciplinary teaching strategies and curriculum planning. Feedback collected through questionnaires highlighted the impact of collaborative tools on student interaction and learning outcomes. This convergence of research demonstrates the potential of cloud computing tools and online collaborative learning environments in facilitating enhanced information sharing, accessibility, and student engagement.

## 3. Methodology

In order to investigate the effectiveness of CSCW among students during the pandemic, a survey was administered to forty (40) senior high school (grades 11-12) STEM students, 60% of which are 18 years old and above. Google Form was utilized in collecting both quantitative and qualitative data on the various CSCW platforms used during the COVID-19 pandemic and its impact on their academic performance. Following the ethics guidelines of the school in the conduct of an experiment, consent forms were obtained from participants, including parental consent for minors. Additionally, ten participants were interviewed to gather further insights.

Quantitative research through a survey focuses on determining the specific collaborative platforms students used during the COVID-19 pandemic and their effects on academic and group work performances. Questions include the various collaborative platforms used, their effects on academic performance, and the factors that made them enjoyable and beneficial form the survey. The questions are based on the Unified Theory of Acceptance and Use of Technology (UTAUT) model and its variant UTAUT2 which are considered as technology acceptance models where factors such as performance and effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit are assessed (Venkatesh, et al., 2003).

#### 4. Results and Discussions

The survey instrument is divided into five sections: Basic Information of a respondent, Collaborative Platforms used, Utilization, Features, and Feedback. Table 1 presents the usage frequency of the different collaborative platforms. Based on the results, the top 3 platforms that are most frequently used are Zoom, Google Docs and Google Drive. Zoom is found to be the most popular platform for the following reasons: 1) free of use; 2) allows the collaborators to meet virtually, and 3) supports all the necessary features in a virtual meeting. Google Docs allows the collaborators to share and edit the same document at the same time where the revisions are reflected in real-time. Google Drive has become popular as well since collaborators can share the same drive and easily upload and revise the uploaded files.

Collaborative Platform	Never	Least Frequent	Less Frequent	Moderate	Frequent	Most Frequent
Google Docs		0%		1 (2.5%)	4 (10%)	35 (87.5%)
Microsoft 365	9 (22.5%)	6 (15%) 5 (12.5%)		9 (22.5%)	8 (20%)	3 (7.5%)
Canva		0%	2 (5%)	5 (12.5%)	13 (32.5%)	20 (50%)
Zoom		0%	1 (2.5%)	0%	2 (5%)	37 (92.5%)
Google Meet	2 (5%)	10 (25%)	9 (22.5%)	10 (25%)	6 (15%)	3 (7.5%)
Google Drive		0%	1 (2.5%)	4 (10%)	6 (15%)	29 (72.5%)
Discord	2 (5%)	3 (7.5%)	3 (7.5%)	10 (25%)	9 (22.5%)	13 (32.5%)
Notion	12 (30%)	5 (12.5%)	5 (12.5%)	6 (15%)	4 (10%)	8 (20%)
Adobe Shared Projects	24 (60%)	6 (15%)	6 (15%)	3 (7.5%)	1 (2.5%)	0%
Jamboard	2 (5%)	9 (22.5%)	12 (30%)	11 (27.5%)	4 (10%)	2 (5%)

Table 1. Usage Frequency of Each Collaborative Platform

UTAUT and UTAUT2 models are employed to measure the performance and effort expectancy, social influence, facilitation conditions, hedonic motivation, price value, and habit. *Performance expectancy* is the degree to which an individual believes that using a computerized system will assist him or her in increasing his or her job performance. Out of the 40 respondents, 39 agreed that these tools had been helpful on their education. With regard to speed in task completion when these tools are used, 85% agreed that the tools allowed them to complete their tasks faster. 75% believed that the tools made them more productive, with 22.5% remaining neutral and 2.5% to be disagreeing. Similarly, 77.5% believed that collaborative tools improved their academic performance, while 20% were neutral and the remaining have disagreed. Among the respondents, 87.5% believe that there was a change in their utilization of collaborative tools in their academic tasks, 10% were neutral, and the remaining participant did not perceive any change. Overall, results have shown that 75% and above of the participants believed that CSCW platforms were found to be helpful and effective in their education, accomplishment of tasks, productivity and academic performance.

For the *effort expectancy* factor, the results indicate that using collaborative tools and platforms requires low effort. A significant majority of the survey participants, 75%, found it easy to acquire skills in using these platforms, while 25% of the participants remained neutral. Regarding group work and collaborative activities, 85% agreed that collaborative platforms facilitated easier group cooperation. Additionally, 82.5% did not encounter difficulties in navigating through collaborative platforms, with 17.5% remaining neutral. When it comes to comfort level, 90% of respondents expressed being more comfortable using these tools, while 7.5% felt neutral and 2.5% disagreed. Finally, 87.5% agreed that collaborative tools made their learning experience during the pandemic easier, while 12.5% remained neutral. Results have shown that with CSCW platforms, at least 75% of the participants believe that with such technology, have helped them to become skillful, collaborative and navigate the such platforms.

With regard to the *social influence* on the usage of CSCW, majority of the participants, 85%, perceive that their peers have an influence on their utilization of collaborative tools and platforms. All participants unanimously agreed that their university has provided support for the use of these tools in their education. This suggests that the university services play a big role in the utilization of these platforms and in assisting the students adapt to the online learning environment during the pandemic.

For the *facilitation conditions* factor, results have shown that majority of participants, (97.5%), claimed to have the necessary resources for using collaborative tools. This is a positive finding, indicating that most participants had access to the technology and resources required. Again, support of the university on these services may have impacted the high score for this factor. Access to these resources is crucial for active participation in collaborative work and online classes. Only one participant remained neutral, possibly indicating they had access to the resources but faced difficulties in acquiring them. Additionally, 92.5% of the participants considered the use of collaborative platforms relevant and suitable for other aspects of their work, while three participants neither agreed nor disagreed. This brings to light the broader applications of these tools beyond education. A significant majority, 95% of participants, expressed their willingness to recommend these tools to others, while 5% remained neutral in their opinion.

For the *hedonic motivation* factor which is a type of motivation that aims to maximize pleasure and minimize pain, 26 participants or 65% feel more motivated to work when using collaborative tools, while the rest remained neutral in their response. However, 96% of the participants, indicated a preference for working with others through collaborative platforms compared to other modes of group work. This finding highlights that more than half of the participants recognized the value and benefits of utilizing collaborative tools when engaging in collaborative work with their peers.

85% of the participants believe that gaining access to collaborative tools is essential to their education. Results also show that all survey participants had access to collaborative tools during the COVID-19 pandemic. This may imply that students were able to adapt to the changes and challenges posed by the pandemic by having the essential tools for online learning.

For the *habi*t of using CSCW platforms, majority of the respondents i.e. 78% reported using collaborative tools not only for educational purposes but also in various fields of work. This finding suggests that collaborative tools have become integral to daily work life across

different professions. The COVID-19 pandemic likely accelerated this trend as remote work and reliance on technology for collaboration and communication increased. However, it is noteworthy that a small number of respondents had a neutral or disagreeing stance, indicating the need for ongoing evaluation and improvement of these platforms to ensure effectiveness and accessibility for all users. Additionally, 90% of the participants agreed that collaborative tools were helpful in assisting them with their work during the pandemic.

Table 2 presents the effectiveness and useful features of CSCW platforms. The most useful features of collaborative tools were identified as file sharing (100%) and screen sharing (82.5%) which facilitated document sharing and real-time collaboration. Feedback from the survey participants indicated that collaborative tools had a positive impact on their education, improving their productivity and the quality of their work. They valued the convenience, accessibility, and social interaction provided by these tools.

Features	Agree
File Sharing	40 (100%)
Screen Sharing	33 (82.5%)
Internal Communication Input/ In-platform Chat	31 (77.5%)
Annotation	27 (67.5%)
Project Management	27 (67.5%)
Cloud-based File Storage	27 (67.5%)
Video Sharing	23 (57.5%)

Table 2. Effective and Useful Features

With the qualitative data collection thru interviews, results have shown that CSCW further supported the positive experiences with collaborative tools, emphasizing their convenience and impact on performance. Technical difficulties and lack of guidance were mentioned as challenges during the pandemic. However, all interviewees recognized the essential role of collaborative tools in achieving quality education during that time.

#### 5. Conclusion and Future Works

The COVID-19 pandemic forced educational institutions to adopt alternative modes of learning. Collaborative tools and platforms played a crucial role in mitigating the effects of this disruption on students' academic performance. This study explores how CSCW platforms have affected the academic performances and collaboration of students with fellow students during the COVID-19 pandemic.

Forty (40) STEM high school students participated in this study. Results have shown that collaborative tools enabled remote learning, allowing students to participate in group projects, discussions, and access resources regardless of their location. Students have developed important collaboration skills and gained experience in communication and problem-solving. Feedback from the survey participants indicated that collaborative tools had a positive impact on their education, improving their productivity and the quality of their work. They valued the convenience, accessibility, and social interaction provided by these tools. Moreover, these tools have provided some support and guidance from professors to students transitioning to remote learning. This had even helped the instructors assess students' understanding, offer additional resources, and provide feedback on assignments.

Furthermore, students have utilized these tools in order to stay connected with their peers during the pandemic.

However, there are limitations and areas for improvement in collaborative platforms. Accessibility and cost remain as challenges since not all students have equal access to the necessary tools and platforms. This worsens pre-existing inequalities and requires educational institutions to ensure equal access for all students. Another concern is the potential for distractions and reduced concentration during remote learning. Students need guidance and support to navigate this new learning environment effectively. Our findings highlight the importance of acknowledging and addressing the limitations of collaborative platforms in order to enhance their utilization among students in educational settings. This study serves as a stepping stone for future advancements in the field, providing guidance to platform creators on areas that require focus and improvement. Despite the lockdown in the country caused by the pandemic, students have found some means to collaborate and learn with a group through these platforms. By highlighting the necessity and benefits of CSCW in education, the study may encourage educational institutions to support and invest in these collaborative tools.

Future work may involve more students to share their experiences and thoughts on the impact of such platforms on one's academic performance and productivity.

#### References

- Abelsen, S. N., Vante, S.H., Mikalef, P. & Choudrie, J. (2021). Digital working during the COVID-19 pandemic: how task-technology fit improves work performance and lessens feelings of loneliness. *Information Technology & People*. 36(15)
- Alharbi, R., Altayyari, F., Alamri, F. & Alharthi, S. A. (2021). Pandemic-Driven Technology During COVID-19: Experiences of Older Adults. CSCW '21: Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing (pp.5-9).
- Bullinger-Hoffmann, A., Koch, M., Moslein, K., & Richter, A. (2021). Computer-Supported Cooperative Work – Revisited. *i-com*, 20(3), 215-228
- Al-Samarraie, H., & Saeed, N. (2018). A systematic review of cloud computing tools for collaborative learning: Opportunities and challenges to the blended-learning environment. *Computers & Education*, 124, 77–91.
- Bullinger-Hoffmann, A., Koch, M., Moslein, K., & Richter, A. (2021). Computer-Supported Cooperative Work Revisited. i-com, 20(3), 215-228.
- Faja, S. (2013). Collaborative learning in online courses: Exploring students perceptions. *Information Systems Education Journal*, 11(3), 42–51.
- Goff, A. (2022) 23 Collaboration Tools For Teams To Work Together Online. https://hive.com/blog/collaboration-tools-for-teams/
- Hernández-Sellés, N., Muñoz-Carril, P. C., & González-Sanmamed, M. (2019). Computer-supported collaborative learning: An analysis of the relationship between interaction, emotional support and online collaborative tools. *Computers & Education*, *138*, 1-12
- Vasilijevic, M., Atanasijevic, S., Zahar, M. & Atanasijevic, T. (2021). Selecting Effective Collaborative Tools For Creating New Social Network in Company For Raising Employee's Engagement During COVID-19 Pandemic. *Information Systems And Security Session*, 238-245.
- Venkatesh, V., Morris, M., Davis, G., & Davis F. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*. 27 (3), 425–478.

## Design of a Pneumatics System Learning Material with AR Technology for Vocational Education Students

#### Tanit PHETCHAKAN <sup>a,b\*</sup>, Suppachai HOWIMANPORN <sup>a</sup> & Sasithorn CHOOKAEW <sup>a</sup>

<sup>a</sup>Department of Teacher Training in Mechanical Engineering, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Thailand <sup>b</sup>Rayong Technical College, Thailand \* datatanit@gmail.com

Abstract: Nowadays, education in industrial engineering places a significant emphasis on pneumatic systems due to their widespread application in conveyor belt systems and various automated industrial processes. Components within pneumatic systems include compressors, air preparation units, directional control valves, actuators, and gripping devices. Therefore, understanding pneumatic systems necessitates having a solid understanding of how to employ these components efficiently. In educational management, the development of learning materials holds paramount importance. These materials play a pivotal role in facilitating enhanced comprehension among students. Integrating augmented reality technology further enriches the learning experience, specifically in understanding pneumatic systems. This technology provides students with a comprehensive system overview before engaging with actual equipment in practical tasks. This study proposed the design of a pneumatics system learning material with AR technology for vocational education students. It encourages the student's motivation in the concept of pneumatic systems. We created a model using 3D and AR application programs to generate learning material to display on mobile devices. The circuits of the function of the cylinder work in various conditions. It has been used to simulate the use of training skills for students or interested people to apply such knowledge to work in the workplace in the future.

Keywords: Augmented reality technology, pneumatics system

#### 1. Introduction

The world is moving towards a new industry of the future. Especially Thailand Industry 4.0. policy focused on the technicians or workers is a critical force that will dramatically change the pattern of the industry. This group of technicians' development is significant and must come from knowledge, experience, and technology. There are fewer workers in Thailand in the field of professional work with skills that meet the industry's needs. Thus, many vocational colleges attempt to increase the workforce with such produce technical skills to reach a high level, and it is necessary to work together to accelerate the development of vocational students who are professional workers in the system. Innovative technology is becoming increasingly vital in the education system. Mainly engineering education necessitates using technology to assist students in understanding abstract concepts and principles. Augmented reality (AR) is a technology that can be used to create effective and engaging technology-based solutions and instructional materials.

AR is a technology for learning that has significantly increased over a wide range of applications, primarily used in engineering education (Singh et al., 2019; Enzai et al., 2021; Takrouri et al., 2022; Koparan et al., 2023). Many studies presented an augmented reality-based learning experience to teach students in many engineering majors, such as electronics engineering majors identifying AR intervention on students' academic achievement levels, positive learning attitudes toward the subject, and individual attitudes toward AR (Tuli et al., 2022). It provides significant benefits, such as increased engagement and interactivity, and can help to minimize the adverse effects of the disruption of face-to-face education (Criollo-C et al., 2021).

AR technology is categorized into two primary types: those reliant on pictograms and those based on coordinate systems for data analysis to generate virtual world information. Commonly, the symbols employed in these systems are referred to as "Markers" or sometimes AR Codes. These markers are detected using a webcam, and when the software processing the image identifies a specified symbol, it renders the 3D image information programmed for that marker. Users have the capability to freely rotate the displayed image in all directions, a feature often described as a 360-degree rotation.

Pneumatic systems are used extensively in industrial applications, regulating machinery operations across a spectrum ranging from small to large industries. Employing pneumatic systems for control aims to streamline work processes with simple circuit design and control mechanisms. Additionally, pneumatic systems boast a high level of safety (Wang et al.,2021; Montalvo et al.,2021). Now commonly utilized, these systems are automatic control systems and are controlled electrically. Presently, numerous factories have established production bases in Thailand, leading to the emergence of multiple industrial estates. This trend is attributed to government policies promoting the growth of small and medium-sized enterprises. Consequently, there is intense competition in the marketing and pricing of products. This study designed a pneumatics system learning material with AR technology for vocational education students. It encourages the student's motivation in the concept of pneumatic systems. We created a model using 3D and AR application programs to generate learning material to display on mobile devices. The circuits of the function of the cylinder work in various conditions. It has been used to simulate the use of training skills for students or interested people to apply such knowledge to work in the workplace in the future.

## 2. Related work

#### 2.1 Mobile Application for Learning

Technological innovation will transform the teaching and learning process with new types to enhance students' development and contextual understanding. Most students are focused on mobile screens in the classroom, and they will all be viewing holographic 3D objects surfaced from a table while the teacher explains the visuals (Hadgraft & Kolmos, 2020). The modern engineering educational environment ensures an extensive application of AR on mobile device technology to memorize graphic material successfully (Zhylenko et al., 2021). Arulanand et al. (2020) presented an AR-based application with Android studio on a topic in the Engineering Graphics course for engineering drawing. Several steps are involved in creating the application, and the below block diagram shows how to create and deploy the application. Wongwatkit et al. (2018) presented the context of an educational institution focusing on developing student learning outcomes throughout the curriculum. The learning outcome is and determines how students will be after the teaching and learning process.

## 3. Conceptual Framework

#### 3.1 Development AR mobile application

This study proposed the development of a pneumatics system learning material using AR to promote vocational students based on the learning factory. Pneumatics system concepts should be learned based on the subject's learning outcomes so that the student can understand the principle of the pneumatics system. Thus, the conceptual framework of the study is shown in Figure 1.



Figure 1. Augmented Reality technology conceptual framework

## 3.2 ADDIE instructional development model

ADDIE model is a well-established framework for creating and delivering educational media, backed by published research. It serves as an excellent guideline for instructional designers, lesson planners, and training course developers in designing effective and engaging learning experiences.



Figure 2. ADDIE model

In this development process, we use ADDIE instructional development model (Chookaew et al., 2014) to create effective AR learning material as follows:

- *Step 1 Analysis:* The analysis clarifies the instructional problems and objectives and identifies the learning environment and student's existing knowledge and skills. We analyzed the content and concept of control circuits in pneumatics systems in the diploma program.
- *Step 2 Design:* This step is designing the learning material based on the learning objectives of control circuits in pneumatics systems. It is guidelines for presenting the lesson as content details of the pneumatics systems. In this study, the student can scan the QR code as a maker for a selection of cylinder control circuits in the worksheet, as shown in Figure 3.



Figure 3. Scan the QR code in the worksheet to select the cylinder control circuits.

- Step 3 Development: The conceptual development of the learning material on the mobile application. In this step, the design of the application screen is to match the content and be appropriate for the students. Figure 4 shows the content of the cylinder control circuit. We created an AR application for the pneumatics systems course, including four sections as follows:
  - 1. The entrance to the instruction is to scan the QR Code to select the cylinder control circuit worksheet. The student can select the One-way cylinder operation control circuit or the Two-way cylinder operation control circuit.
  - 2. The entrance to the AR application are content pages, and the students can click details of pneumatic equipment.
  - 3. The student can see symbols and images of pneumatic equipment.
  - 4. The students click the RUN button; they can see a video of the working process of the pneumatic system.



Figure 4. Examples of instrumental material of cylinder control circuit on mobile device

- Step 4 Implementation: After developing the learning material is complete. This step can be implemented with students in target groups that are vocational education to achieve based on learning outcomes and prepare evaluation procedures.
- Step 5 Evaluation: This step consists of two sections. First, we design the evaluation of students in the cognitive domain, which is students' understanding of pneumatics system concepts. Second, we designed to evaluate student motivation for AR learning materials.

## 4. Conclusion and Future work

This paper presents designing a pneumatics system learning material using augmented reality technology for vocational students. We focus on simulated circuits of the function of the cylinder work in various conditions. It has been used to motivate the use of training skills for students or interested people to apply such knowledge to work in the workplace. This proposed AR application of learning a pneumatics system is used to motivate and promote the students' understanding of pneumatic circuits before implementing them in actual situations. We plan to implement it with the students in future studies to ensure this learning material can encourage students to improve their learning performance and motivation.

#### Acknowledgements

This research was supported the Rayong Technical College and the Division of Mechatronics and Robotics, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok.

#### References

- Arulanand, N., Babu, A. R., & Rajesh, P. K. (2020). Enriched learning experience using Augmented Reality Framework in Engineering Education. *Procedia Computer Science*, *172*, 937–942.
- Chookaew, S., Howimanporn, S., Sootkaneung, W., Pradubsri, W., & Yoothai, P. (2014). Computer assisted learning based on ADDIE instructional development model for visual impaired students. *Proceeding of the 22nd International Conference on Computers in Education, (ICCE)* (pp. 349–354).
- Criollo-C, S., Abad-Vásquez, D., Martic-Nieto, M., Velásquez-G, F. A., Pérez-Medina, J. L., & Luján-Mora, S. (2021). Towards a new learning experience through a mobile application with augmented reality in engineering education. *Applied Sciences*, 11(11), 4921.
- Enzai, N. I. M., Ahmad, N., Ghani, M. A. H. A., Rais, S. S., & Mohamed, S. (2021). Development of augmented reality (AR) for innovative teaching and learning in engineering education. Asian Journal of University Education, 16(4), 99-108.
- Hadgraft, R. G., & Kolmos, A. (2020). Emerging learning environments in engineering education. *Australasian Journal of Engineering Education*, 25(1), 3-16.
- Koparan, T., Dinar, H., Koparan, E. T., & Haldan, Z. S. (2023). Integrating augmented reality into mathematics teaching and learning and examining its effectiveness. *Thinking Skills and Creativity*, 47, 101245.
- Montalvo-Lopez, W., Catota, P., Garcia, C. A., & Garcia, M. V. (2021, September). Development of a virtual reality environment based on the CoAP protocol for teaching pneumatic systems. *Proceeding of the International Conference on Augmented Reality, Virtual Reality and Computer Graphics* (pp. 528-543). Cham: Springer International Publishing.
- Singh, G., Mantri, A., Sharma, O., Dutta, R., & Kaur, R. (2019). Evaluating the impact of the augmented reality learning environment on electronics laboratory skills of engineering students. *Computer Applications in Engineering Education*, 27(6), 1361-1375.
- Tuli, N., Singh, G., Mantri, A., & Sharma, S. (2022). Augmented reality learning environment to aid engineering students in performing practical laboratory experiments in electronics engineering. *Smart Learning Environments*, 9(1), 1-20.
- Takrouri, K., Causton, E., & Simpson, B. (2022). AR Technologies in Engineering Education: *Applications, Potential, and Limitations*. Digital, 2(2), 171-190.
- Wang, H., Tong, Z., Dong, X., Xiong, W., Ting, D. S. K., Carriveau, R., & Wang, Z. (2021). Design and energy saving analysis of a novel isobaric compressed air storage device in pneumatic systems. *Journal of Energy Storage*, 38, 102614.
- Wongwatkit, C., Prommool, P., Chookaew, S., & Mee-inta, A. (2018). A Mobile Web Application for Learning Outcome Evaluation: Analysis and Design of Teacher and Student Interfaces. *Proceedings of the Global Wireless Summit (GWS)* (pp. 151-155). IEEE.
- Zhylenko, T., Ivanov, V., Pavlenko, I., Martynova, N., Zuban, Y., & Samokhvalov, D. (2021, June). Mobile Applications in Engineering Based on the Technology of Augmented Reality. *Proceeding of the International Conference Innovation in Engineering (ICIE)* (pp. 366-376). Springer, Cham.

## Using of Augmented Reality Technology in the Learning Process of Calculus2 for Higher Education Students

Anek PUTTHIDECH<sup>a\*</sup>, Amnaj SOOKJAM<sup>a</sup>, Suwit SOMSUPHAPRUNGYOS<sup>b</sup> & Sangtong BOONYING<sup>b</sup>

<sup>a</sup>Department of Science and Technology, Rajamangala University of Technology Suvarnabhumi, Thailand <sup>b</sup>Department of Business Administration and Information Technology, Rajamangala University of Technology Suvarnabhumi, Thailand \*anek@rmutsb.ac.th

**Abstract:** This paper presents the use of augmented reality techniques to enhance learning experience of studying Calculus2. One of the main problems in learning Calculus is the lack of spatial recognition in objects, especially 3-dimensional objects. To improve students' understanding and immersive experience, Augmented Reality (AR) is applied to overlay 3D objects onto the student's view of the physical world. The chosen 3-dimensional objects include sphere, hyperbolic paraboloid, elliptic paraboloid, paraboloid, cone, ellipsoid, cylindrical, elliptic hyperboloid of one sheet, and elliptic hyperboloid of two sheet. From testing, 30 students found the augmented reality technology effectively helps in learning shapes and calculation for volume and surface area.

Keywords: Augmented Reality Technology, Geogebra 3D, Geometry, Calculus2

#### 1. Introduction

Information and communication technology has invented and developed to provide new mediums for communication such as 3D illustration, video and interactive graphical object. These new technologies help to improve understanding more than just words communicating that can be ambiguous. In education, a textbook is the main teaching materials to provide structural and comprehensive content, reference material, and visual aids. Textbooks generally include diagrams, illustrations, charts, graphs, and other visual aids that help students visualize and understand complex concepts. Illustrations of real-world objects give students a rough concept of thing that is mentioned in a descriptive content. However, illustrations have major limitation as two-dimensional presentation (2D). Furthermore, the illustration is a still image which does not instigate sufficient interest of students (McQuail, 2005). In the higher education such as engineer and biology, the complexity of the content requires students to clearly understand the content and visualize the complex structure of things to learn further in the subject. As one of the mandatory subjects, Calculus2 is to study and analyze three-dimensional objects, particularly when it comes to finding volumes, surface areas, and other properties of such objects. Students may confuse or not fully understand 3D shapes such as hyperbolic paraboloid, elliptic paraboloid, ellipsoid, cylindrical, elliptic hyperboloid of one sheet, and elliptic hyperboloid of two sheet.

In this paper, we present the use of augmented reality to visualize 3D shapes for students who learn in Calculus2. This research raises the question whether teaching Calculus2 with 3D images will help to improve interest and understanding for students or not. The scope of this research is 3D shapes of sphere, hyperbolic paraboloid, elliptic paraboloid, paraboloid, cone, ellipsoid, cylindrical, elliptic hyperboloid of one sheet, elliptic hyperboloid of 2 sheet. The applied augmented reality is developed as a mobile application (android operating system only) and use in a classroom of 30 students in a subject of Calculus2.

## 2. Analysis of recent research and publications

Education 4.0 has modern digital technologies such as augmented reality (AR), virtual reality (VR), and 3D printers applied to different levels of the teaching and learning process. Each of these technologies show a positive effect on learning. The interest in these technologies is growing among teachers and learners.

Augmented reality (AR) is a technology that can combine the real world with the virtual world through a mobile smart device in conjunction with various application software to make the image seen on the screen. This can be an object with a 3D perspective or a 360-degree panoramic view. Users at different levels are increasingly willing to interact with subjects in cyberspace, which can lead to the expansion of educational activities, improving existing and new organizational forms, types and methods of teaching, improving student interaction and educational space (Monfared et al., 2022). AR is expecting to change education and displace the game as an educational technology (Ashley-Welbeck & Vlachopoulos, 2020). Features of the use of AR technology, its impact and importance in the educational academic articles (Akçayır et al., 2016; Dinis et al., 2018; Huttar & BrintzenhofeSzoc, 2020; Frolli et al., 2021; Yusuf & Ichsan, 2021; Trust et al., 2021; Volynets, 2021)

Researchers believe AR will make teaching and learning more interesting. It facilitates the learning process and use frequently in an e-learning system. However, there is a remark as it may reduce direct dialogue between teachers and students. Therefore, the use of elearning technology in education should be used in short-term classes or as a simulation device (Murodillayevich, Eshpulatovich, & Pardaboyevich, 2019). Moreover, AR is a popular technology in mathematics education (Melnyk, Nefedova, & Zadyrey, 2018; Osypova, & Tatochenko, 2021). It is noted that while studying mathematics, students found it is difficult to follow and lost interest due to the complexity of the subject. The use of AR technology in teaching can help students learn some concepts or theories that are difficult to understand in the classroom by creating instructional materials with appropriate 3D models to develop students' spatial and imagination abilities (Gargrish et ai., 2021). This assists to deepen the understanding of processes, properties, theorem proofs, etc. In addition, AR can radically change the way subjects are taught by emphasizing student participation and interaction with learning activities through practice (Melnyk, Nefedova, & Zadyrey, 2018). Some articles (Ancochea & Cárdenas, 2020; Babkin et al., 2021; Osypova & Tatochenko, 2021) applied GeoGebra 3D Graphing Calculator, a mathematical aid program, for solving 3D mathematical problems. The provided graphing 3D functions and 3D geometry surfaces is highly efficient with augmented reality (AR) technology in the study of geometry. The article (Babkin et al., 2021) discusses cross-platform products that should be used to develop AR technologies including Unreal Development, Kit, Unity, Godot, Engine, Cocos2D, MonoGame, Unreal Engine, Marmalade, and the capabilities of a new AR software development kit (or SDK) called ARCore that will be brought to Android smartphones (Wikitude, Vuforia, Kudan, Maxst, Xzimg, NyARToolkit, Metaio SDK) (Vakaliuk & Pochtoviuk, 2021).

## 3. Research design

This research is a research and development (R&D). Research objective is to develop a learning assistant tool for Calculus2 subject using augmented reality (AR) technology for representing three-dimensional space with various shapes including hyperbolic paraboloid, elliptic paraboloid, ellipsoid, cylindrical, elliptic hyperboloid of one sheet, and elliptic hyperboloid of two sheet. Procedures in this research applied several stages as (1) observing research location, (2) creating learning materials, (3) expert validation, (4) testing into small group, (5) setting up into large group (field test).

To develop the learning assistant tool for Calculus2 subject, we chose GeoGebra 3D with AR (GeoGebra Augmented Reality) extension as the development platform to construct the three-dimensional space with various shapes. The image processing application is

performed using the Augment application. This application is a mobile application downloadable on android OS.

For evaluation, there are three experts who are tasked to examine learning materials, namely two experts in mathematics and computer technology. The issues for evaluating the quality of augmented reality (AR) learning materials include graphics and design, fast access to information, animations, clarity and technicality. After the expert commented and consented the tool, adjustments were made according to the instructions, and the tool was brought to trial (Pilot) with thirty students. The learning materials and tools were assessed with the developed questionnaire and improved again before being experimented in a large group (for field test). This research collected both qualitative and quantitative data. The qualitative data such as answers, reviews and suggestions are obtained from interviews and the quantitative data consisted of a quality check form of experts and opinions of students from questionnaires. The data was analyzed using descriptive analysis consisting of frequency, percentage, mean and standard deviation.

Consider examples of the use of GeoGebra Augmented Reality to solve problems of the course "Calculus2" on the topics "Application of the integral" (functions of one and more variables) and "Area of definition of the function of several variables". Geogebra Mathematical Package can also be used to calculate geometric quantities (areas, volumes) and physical quantities (masses, coordinates of the center of mass, work, pressure forces). However, we focus on the ability of the GeoGebra Augmented Reality tool to complement physical reality with digital model-objects that are specified in the task condition. In solving such problems, the greatest difficulties arise when performing geometric constructions and visualization of mathematical objects. Qualitatively performed geometric constructions are one of the conditions for the correct solution of the problem.

For usage, students initially open up GeoGebra 3D Calculator application and create any 3D solid model (see Figure 1) to create a 3D Graphics in an augmented reality application. Next, they are asked to press the AR button (lower right within 3D Graphics pane). Students thus will see environment on the screen. After following the usage instruction, the tool will virtually place the object you were viewing within your environment. Students can also use fingers to reposition the objects and allow to use the touch screen to change the size and color of the object.



Figure 1. Demonstration of 3D creating sphere shape.



Figure 2. Demonstration of triangulated and square appear over the flat surface.

From the example given in Figure 3, the calculation of the area of the part of the paraboloid is  $z = x^2 + y^2$ . Students can see the visualization of the mathematical object given by the condition of the problem. The volume of the body of rotation can be calculated using the tool "Integral".



Figure 3. Demonstration of 3D creating paraboloid shape.

#### 4. Conclusions

The usage of augmented reality technologies promotes visualization in the educational process, the formation of cognitive interest and motivation of students. It has significant methodological potential, which should be used in the study of various mathematical disciplines, particularly in the study of mathematical analysis. The use of such technologies can probably have a great effect, but the constant use in the standard practical training of 90 minutes leads to a significant violation of the program. However, the use of augmented reality technologies is the most adequate as an additional means of visualizing calculus2 objects in solving problems.

The results of student surveys and trainings confirm the positive impact of augmented reality (AR) technologies on the learning process and the quality of learning outcomes. From interview results, a significant number of higher education students note that augmented reality (AR) technology helps to better understand theoretical material and enhances the visual component. Some students also mentioned that it gives more immersive experience than traditional lecture and self-reading. Furthermore, it keeps students' interest in class and focus on studying as using the tool is more practical and require students' interactivity.

#### References

- Akçayır, M., Akçayır, G., Pektaş, H. M., & Ocak, M. A. (2016). Augmented reality in science laboratories: The effects of augmented reality on university students' laboratory skills and attitudes toward science laboratories. *Computers in Human Behavior, 57*, 334-342.
- Ancochea, B., & Cárdenas, M. I. (2020). Exploring real world environments using potential of Geogebra AR. In L. Matthias, J. Simone, C. Amélia, & M. Ana (Eds.), *Research on Outdoor STEM Education in the digiTal Age. Proceedings of the ROSETA Online Conference in June* (pp. 41-46). https://doi.org/10.37626/GA9783959871440.0
- Ashley-Welbeck, A., & Vlachopoulos, D. (2020). Teachers' perceptions on using augmented reality for language learning in primary years programme (PYP) education. *International Journal of Emerging Technologies in Learning (iJET), 15*(12), 116-135.
- Babkin, V. V., Sharavara, V. V., Sharavara, V. V., Bilous, V. V., Voznyak, A. V., & Kharchenko, S. Y. (2021). Using augmented reality in university education for future IT specialists: educational process and student research work. In S. H. Lytvynova, & S. O. Semerikov (Eds.), Augmented Reality in Education 2021: Vol. 2898. 4th International Workshop on Augmented Reality in Education (pp. 255-268). https://doi.org/ 10.31812/123456789/4628
- Dinis, F. M., Martins, J. P., Carvalho, B. R., & Guimarães, A. S. (2018). Disseminating civil engineering through virtual reality: An immersive interface. *i-joe International Journal of Online and Biomedical*, 14(5), 225-232.
- Frolli, A., Bosco, A., Lombardi, A., Di Carmine, F., Marzo, S., Rega, A., & Ricci, M. C. (2021). Asperger's and virtual reality. In P. Limone, & R. D. Fuccio (Eds.), *Technology Enhanced Learning Environments for Blended Education: Vol. 2817. Procceedings of the First Workshop on Technology Enhanced Learning Environments for Blended Education* (pp. 212-222). https://ceur-ws.org/Vol-2817/
- Gargrish, S., Kaur, D. P., Mantri, A., Singh, G., & Sharma, B. (2021). Measuring effectiveness of augmented reality-based geometry learning assistant on memory retention abilities of the students in 3D geometry. *Computer Applications in Engineering Education, 29*(6), 1811-1824.
- Huttar, C. M., & BrintzenhofeSzoc, K. (2020). Virtual reality and computer simulation in social work education: A systematic review. *Journal of Social Work Education, 56*(1), 131-141.
- McQuail, D. (2005). Book Review: Comparing media systems: Three models of media and politics. *European Journal of Communication*, 20(2), 266-268.
- Melnyk, I. Y., Nefedova, H. D. & Zadyrey, N. M. (2018). Augmented and virtual reality as a resource for students educational activities Information technologies and computer modeling. *Proceedings of the Third Workshop on Technology Enhanced Learning Environments for Blended Education* (pp. 61-64). https://doi.org/10.1201/9781003196686-9
- Monfared, M., Shukla, V. K., Dutta, S., & Chaubey, A. (2022). Reshaping Education Through Augmented Reality and Virtual Reality. In J. M. R. S. Tavares, P. Dutta, S. Dutta, & D. Samanta, (eds). *Cyber Intelligence and Information Retrieval: vol. 291. Recture Notes in Networks and Systems* (pp.619-629). Springer, Singapore. https://doi.org/10.1007/978-981-16-4284-5\_55
- Murodillayevich, N. F., Eshpulatovich, U. G., & Pardaboyevich, J. O. (2019). Integration of virtual reality and 3D modeling use of environments in education. *In 2019 International conference on information science and communications technologies (ICISCT)* (pp. 1-6). IEEE. https://doi.org/ 10.1109/ICISCT47635.2019.9011899
- Osypova, N. V., & Tatochenko, V. I. (2021). Improving the learning environment for future mathematics teachers with the use application of the dynamic mathematics system GeoGebra AR. In S. H. Lytvynova, & S. O. Semerikov (Eds.), Augmented Reality in Education 2021: Vol. 2898. 4th International Workshop on Augmented Reality in Education (pp.178-196). https://doi.org/ 10.31812/123456789/4628
- Trust, T., Woodruff, N., Checrallah, M., & Whalen, J. (2021). Educators' interests, prior knowledge and questions regarding augmented reality, virtual reality and 3D printing and modeling. *TechTrends*, *65*, 548-561.
- Volynets, K. (2021). The Use of Virtual Reality Technologies in Education. *Continuing professional education: Theory and practice, 2,* 40-47.
- Vakaliuk, T. A., & Pochtoviuk, S. I. (2021). Analysis of tools for the development of augmented reality technologies. In S. H. Lytvynova, & S. O. Semerikov (Eds.), *Augmented Reality in Education 2021: Vol. 2898. 4th International Workshop on Augmented Reality in Education* (pp. 119-130). https://doi.org/ 10.31812/123456789/4628
- Yusuf, M., & Ichsan, R. N. (2021). Analysis of banking performance in the aftermath of the merger of bank syariah indonesia in Covid 19. *International Journal of Science, Technology & Management,* 2(2), 472-478.

## Encouraging Primary Students' Environmental Awareness by Using STEM Inquiry-based Learning

Sasivimol PREMTHAISONG<sup>a</sup>, & Pawat CHAIPIDECH<sup>b\*</sup>

<sup>a</sup>Khon Kaen University Demonstation School International Division, Khon Kaen University, Thailand

<sup>b</sup>Faculty of Education, Khon Kaen University, Thailand \*pawach@kku.ac.th

Abstract: The Circular Economy (CE) has emerged as an essential idea in the worldwide response to unsustainable resource consumption and waste management. and educators have incorporated this scenario into the theme in classroom. At the same time, STEM (Science, Technology, Engineering, and Mathematics) education is an essential approach for educating students with the necessary skills for a competitive workforce in the twenty-first century. Several academics are currently attempting to integrate inquiry-based learning, which is one such technique whose efficacy is higher in science-based education. Furthermore, technological integration has been used in scientific courses. This research examined the implementation of combining inquirybased learning with STEM education using ICT technology in citizen science activities to promote the circular economy for fourth-grade students in a university-based demonstration school located in the northeastern part of Thailand. The results of this study include aspects of students' conceptual understanding about types of plastic and their awareness of the environment. We found that students had a significantly increased conceptual understanding regarding types of plastic. Moreover, the students expressed positive attitude towards plastic waste management awareness. These findings support the notion that ICT technology could assist students in better understanding science concepts and attitudes, especially when integrated with a STEM inquiry-based learning method in the citizen science activity.

Keywords: STEM, Technology, Circular economy, Citizen inquiry

## 1. Introduction

In recent years, the discipline of science, technology, engineering, and mathematics (STEM) has experienced an important rise in popularity and importance as an academic study and research area. STEM education has become a significant focal point of national importance in several countries worldwide. Several educators have adopted an integrated approach to STEM (science, technology, engineering, and mathematics) education to promote essential skills among students, such as creativity, problem-solving, and teamwork. This strategy aims to cultivate a highly skilled workforce that can effectively compete in the 2 1st century (Freeman, Marginson, & Tytler, 2019). Moreover, STEM education could improve students' motivation, interests, and persistence in learning, fostering their development as self-directed learners (Abouhashem et al., 2021; Syed Abdullah, & Md Khambari, 2022). Additionally, students' effective domain toward STEM educational work is important for successful STEM education implementation in a sustainability context (Sellami et al., 2023). Consequently, to promote a nation's sustainable economic and social well-being, education must increasingly prepare students for STEM-related disciplines of study and future careers (Mohr-Schroeder et al., 2020).

Citizen Inquiry is an innovative approach to informal science education that attempts to facilitate the involvement of citizens in online scientific investigations (Aristeidou, Scanlon,

& Sharples, 2013). This approach integrates aspects from both Citizen Science and Inquirybased Learning, including practices such as sharing information and evaluation by others (Citizen Science), as well as experimentation, exploration, critical analysis, and reflection (Inquiry-based Learning). In order to enhance environmental awareness among students may be achieved through formal and informal educational approaches. As informal learning, citizen Inquiry arises as the learning method, motivated by people's personal interests in science and its basic rationale. Environmental awareness can be developed by the adoption of attitudes and actions that are focused on conserving the environment, while the obtaining of knowledge is essential for effectively managing the environment, Nizaar et al. (2020) found that by teaching students to process garbage using the 3R principle, it is possible to reduce excessive household waste production.

During the emerging technologies era, mobile technology, one of the technologies that contribute to citizen inquiry, is implemented to enhance students' effective domain in different subjects and levels. For example, Plastic Island Game (Adita & Srisawasdi, 2022), Weatherit platform (Aristeidou et al., 2015), 360-degree Virtual Learning Environment (Chaipidech & Srisawasdi, 2023), and the Sense-it app (Sharples et al., 2017). Consequently, this research investigates the effect of integrating STEM learning into guided inquiry with ICT technology on the conceptual understanding and awareness of students. The outcome might have implications for teaching and learning with technology in the citizen science activity.

## 2. Literature Review

## 2.1 The Role of Integrated STEM with Circular Economy

Nowadays, education is a fundamental requirement for attaining Sustainable Development Goals (SDGs) as it provides individuals with essential skills, values, and knowledge necessary for personal and societal progress. To effectively incorporate sustainable development into educational institutions, it is important to begin an extensive redesign of the learning environment and instructional methodologies. In addition, the topic of waste management and recycling is presently the most important challenge in the advancement of sustainable development. The circular economy (CE) has become an essential concept in response to waste management and recycling. Moreover, citizens of the world are expected to use the scientific and technological knowledge they obtain in school to resolve real-world issues, such as environmental degradation, unpredictable climate change, and resource depletion (Nguyen et al., 2020). STEM education is a one of learning that could facilitates the integration of complex academic concepts with practical applications by employing an interdisciplinary approach to instruction and learning (Bybee, 2013). Consequently, several educators endeavor to integrate instructional methods and learning strategies. As integrate STEM education with the circular economy, Nguyen (2023) demonstrated the integration of Circular Economy principles into Science, Technology, Engineering, and Mathematics (STEM) education to promote sustainable development and reduce waste generation. This paper shows that the incorporation of CE into STEM education is part of a "breadth and depth" educational strategy that encourages teachers to cross disciplinary lines while maintaining their expertise in the field.

#### 2.2 Citizen Inquiry Integrating with Technologies

Citizen science is a rapidly expanding field of study and practice that generates new knowledge and understanding through the participation of citizens in scientific research. It is becoming increasingly essential for educators to consider the field's potential to promote educational and learning possibilities as it develops. Citizen science has been recognized as having enormous potential in the field of scientific education and learning for some time (Bonney et al., 2009). As for the potential of citizen science, numerous researchers challenge integrating of citizen science into mainstream education systems, especially inquiry learning.

Citizen inquiry is one of innovative method that incorporates components from citizen science and inquiry-based Learning, including sharing of information and review by others (Citizen Science) and experimentation, discovery, critique, and reflection (Inquiry-based Learning). The purpose of citizen inquiry is to enable students or anyone to engage citizens in online scientific investigation (Aristeidou et al., 2013). For example, Aristeidou et al. (2013) explore the creation and evolution of an online community called Weather-it enabling people to create and maintain their weather investigations to which others can contribute. The result showed that the sustained engagement and evolution of a citizen inquiry community will require active engagement of a person. The study also found that feeling a part of the community increases the likelihood of remaining active, but there are concerns about the members who did not feel like a part of the community, even though they had many contributions. Additionally, Herodotou (2014) used technology that is the nQuire Missions framework is a web-based platform that provides the development and administration of personal inquiry missions by young people and a sensor-based mobile application that facilitates the collection of data on mobile phones to support citizen inquiry. Moreover, Chaipidech and Srisawasdi (2023) developed and implemented a 360-degree virtual learning environment to support out-of-class inquiry activity for preservice teachers into the citizen inquiry activity. The study found that the learners had positive perceptions and self-reflection regarding their learning experiences using the 360-degree virtual learning environment. Furthermore, using simulation and serious games as learning tools for the circular economy concept, the teaching approaches could effectively promote active student learning (Torre et al., 2021).

## 2.3 Mobile Technology Facilitates Science Learning

Mobile technologies, such as smartphones and tablet computers, have been used in educational settings increasingly over the past ten years in attempts to enhance educational quality and the learning process. Mobile devices, also known as m-devices, have been recognized as a new educational instrument that has the ability to support teaching and learning techniques that take use of the context of each student. As such, Chaipidech and Srisawasdi (2018) used integrating mobile technology with personalized flipped open inquirybased approach that is a simulation and hands-on laboratory via mobile for secondary school students. This result showed that a personalized flipped open inquiry-based approach using mobile technology is effective in enhancing students' conceptual understanding and intrinsic motivation towards learning science. In addition, Thumtathong, Premthaisong and Srisawasdi (2019) examines the effectiveness of teaching English with science using Content and Language Integrated Learning (CLIL) approach and mobile-assisted inquiry pedagogy on students' learning motivations and perceptions. They participated with interactive video from Lifesaver application to learn a choking situation. After that, students used iStyle science note to collect data and Nearpod application to discuss their idea via smartphone about respiratory system concepts in science classroom. Moreover, many researchers integrated STEM modules with various technology, such as computer-based board games, online game. This study demonstrated that connecting classrooms with online interclass tournaments could be possible to implement in schools. Besides, integrating the sustainability concept into Virtual Reality (VR) system-aided STEAM education could enhance the learning satisfaction and outcomes of students and to arouse their learning motivation (Hsiao & Su, 2021).

## 3. An Example of STEM integrated Technology-enhanced Inquiry learning

STEM discipline integrated technology-enhance guided-inquiry science learning is a pedagogical approach aimed at enhancing students' learning in science within the context of circular economy. This research integrated the four disciplines of science, technology, engineering, and mathematics (STEM) in one class, unit, or lesson, based on STEM principles and situated within an actual context or real-world scenario, serves to enhance student learning (Kelly & Knowles, 2016). Moreover, this approach involves engaging students in

processes of scientific inquiry, which enable them to actively work with the scientific method and foster their oriented guided inquiry-based learning, Buck, Bretz, and Towns (2000) describes six steps; the first step is Problem or Question; the teacher will start the session by presenting a problem scenario related to the management of plastic waste, specifically focusing on the amount of plastic waste that has become to the plastic trash mountain. Second is Theory or Background; students will be provided an interactive 360-degree video to learn about plastic varieties in the area of university. The third step is Procedures or Design related to the phenomenon, using "Plastic collector" on the Appsheet, students will be required to collect data by taking photos regarding plastic debris in daily life. In addition, students must select the types and locations of plastic. Fourth is Result analysis; students will be tasked with analyzing data from the "Plastic collector" using Google Sheet to transform the quantity of each type of plastic into a graph. The fifth step is Result Communication; students utilized the Padlet platform to share information and discuss the quantity of plastic. Finally, the conclusion of the session, the teacher will conclude the lesson on plastic types. Figure 1 illustrates the pedagogy that was developed.



Design prototype of the plastic management machine

Figure 1. A STEM inquiry-based Learning with ICT Technology.

Furthermore, the teacher will provide a case study of the procedure for managing plastic waste, which complies with the principles of the 3R approach: reduce, reuse, and recycle. Finally, students will be tasked with the responsibility of designing a prototype for the plastic management machine. Therefore, this study comprises the four disciplines of science, technology, engineering, and mathematics (STEM). The scientific aspect focuses on the various types of plastic, including Polyethylene (PE) and Polypropylene (PP). The technological component involves the 360-degree interactive video, as well as the utilization of platforms such as the application of plastic collectors, Padlet, and other relevant tools. The mathematical dimension includes the analysis of plastic data, while the engineering aspect relates to the design of a prototype for the management of plastic.

## 4. Methodology

## 4.1 Study Participants

The participants of this experiment included 25 fourth-grade students from primary school at a university-based demonstration school located in the northeastern part of Thailand. They were enrolled in a basic education science course and were invited to participate in this study. Participants were between 9 and 10 years old. All of them had adequate fundamental computer and information and communication technology skills, but none had ever used a computer for science instruction before.

#### 4.2 Research Instruments and Data Analysis

The research instruments comprised two instruments. For the first instrument, a conceptual pre-test and post-test for measuring students' learning achievement of the STEM learning activity about types of plastic. The conceptual pre-test was designed to determine whether students had a foundational understanding of all types of plastic. It consisted of ten multiple-choice questions, and one point was given for each correct response, resulting in a total of 10 points. The purpose of the conceptual post-test was to evaluate students' understanding of the main characteristics of various plastic varieties. The pre-test and post-test were both created and evaluated by three experienced science instructors. The second instrument also included a post-awareness interviewing to measuring students' environmental awareness after completing the learning activity. The post-awareness interviewing was adapted from Ozden (2008). An interview protocol has been developed for this study regarding the attitudes toward environmental problems to investigate students' awareness after interacting with a STEM inquiry-based learning with ICT technology.

For the activity, students were given 15 minutes to fill out the conceptual understanding test toward types of plastic concepts. Afterward, over a period of 3 days (a total of 200 min), the students learned about the types of plastic (i.e., Polyethylene Terephthalate, Polyvinyl Chloride, High-Density Polyethylene) in STEM lessons. After completing the instruments, they were explored to interact with the assigned of the technology-integrated STEM inquiry-based learning in the citizen science activity. After finishing the learning process, students were administered by the same conceptual test again for 15 minutes to determine their conceptual understanding. Moreover, students were asked to complete post-interviewing to determine students' awareness of the environment. Figure 2 shows students working on the assignment using various technologies.



*Figure 2.* An illustration of students' interaction with ICT technology integrated STEM inquiry-based learning about types of plastic concepts.

Following the completion of the learning activities, all students participated in a post-test to determine their conceptual understanding. Additionally, they were asked to complete an interviewing to measure their awareness of plastic-related issues. The duration of the interviewing was around 15-20 minutes. Furthermore, researchers focused their attention on the study of the data obtained from the post-interviewing. The statistical data for evaluating students' conceptual knowledge were analyzed using IBM SPSS Statistics 23.

## 4.3 Learning Material

According to previous studies, several researchers have developed a 360-degree virtual tour as an educational tool (Chaipidech & Srisawasdi, 2023; Rupp et al., 2019; Abidin, Suryani, & Sariyatun, 2020; Garcia, Nadelson, & Yeh, 2023). Nevertheless, using 360-degree virtual tours when combined with STEM inquiry-based learning was not employed. As a result, the present study demonstrated the use of 360-degree learning media within the context of STEM education in a science inquiry-based learning module focused on understanding various concepts related to varieties of plastic.

The 360-degree learning media on types of plastic was designed with three components in each location. In the first section, information about varieties of plastic will be displayed, including the names of objects, the type of plastic, the characteristics of plastic, and examples of items made from this type of plastic. After interacting with the second section, the information section, students were given multiple plastic-related questions that create by using "H5P" to assess their understanding. After the user selects an answer, the system will display the correct response and respond in real-time. Students can select the running man icon to navigate to another location and study additional plastic varieties in the last section. The 360-degree media is shown in Figure 3.



Figure 3. Illustrates a 360-degree virtual tour for learning about types of plastic.

According to the citizen inquiry activity, the learner is given the essential inquiry, "What types of plastic are most commonly found in our city?". In order to enhance students' awareness and knowledge of plastic types, we provided the "Plastic Collector" Appsheet platform for collecting plastic waste in everyday life, shown in Figure 4.

ราชชื่อ	guunin Q 🛛 C	- C 🖸	User ID F	orm				← 🔛 User ID For	n
	Alex	The Collect	or*				- 1	0	
e	babus	Pinnny						Photo	
	Bell	Dute					- 1	-	
(8)	Benz(g)		12	Jul 2023		r –	- 1	-	100 <sup>-10</sup>
	0.000	Time						27 29 25	
		19.38.07			٢			R AN	A19104
ny	KETER	Location*							and and a
	Keter				٠			201	
	Keter	• Type of I	Nastic					- 314	
w	mAcHI?	△ PET.	PETE	& HOPE	& PVC				
	MLN	& LDPE	≜PI	e ers	6 liv 1			1.10	Con the second
	Meni 🛨	e Quantity						-	
	II +				P				

*Figure 4.* Illustrate of "Plastic Collector" on Google Appsheet to collect plastic around the city.

As part of this activity's citizen science component, students collected objects produced from plastic by taking at least three daily photographs. After taking photographs, students will choose the types of plastics related to the picture. Then, they have to choose the location of plastics. The advantage of the Appsheet's "Plastic Collector" is that students can share all information about plastic at their location, and others are able to view their information. In addition, students will be aware of the amount of each form of plastic in the city. Based on the aforementioned, this research presented various technologies to support STEM inquiry-based learning in the context of the circular economy. Additionally, in an effort to promote citizen science, students were given the opportunity to participate in citizen science by using their own devices on the Plastic collector application.

## 5. Result and Discussion

## 5.1 Types of Plastic Conceptual

The Wilcoxon signed rank test findings indicate a significant statistical difference (Z = -4.073, p = .000) between the pre-test and post-test scores following the implementation of STEM-integrated technology-based inquiry learning. In addition, the average score of the pre-test understanding was 3.84, with a standard deviation of 1.864. Table 1 displays the mean score of post-test conceptual comprehension, which was found to be 6.28 with a standard deviation of 1.542.

Table 1.	Statistical	Results of	on	Wilcoxon	Signed-Rank	Test	for	The	Students'	Conceptual
Understa	nding Of T	ypes of P	las	tic						

Time		Ν	Sum of Ranks	Z	P value
Pre-test	Negative rank	1 <sup>a</sup>	5.50	-4.073 <sup>b</sup>	.000*
Post-test	Positive rank	22 <sup>b</sup>	270.50		
	Equal	2 <sup>c</sup>			
	Total	25			

\*p<0.05

As shown in Table 1.1, the conceptual understanding of students who learn with a STEM-integrated inquiry-based approach performs better after learning. This evidence demonstrated that integrating a STEM inquiry-based learning approach with various technologies could have greater effects on fostering students' scientific understanding of plastic types in the context of a circular economy. Similar to the prior research that using technology with STEM education, Lindner et al. (2022) found technology study that involved the creation of an augmented reality (AR) app, as well as a worksheet and lesson plan, to incorporate hyperspectral remote sensing data into STEM education. The findings revealed that when students interacted with this strategy, their understanding increased. Similar research was found when using mobile learning the STEM education, the implementation of STEM practical kit using free mobile apps could improve students' understanding on the topic of sound (Chu, et al., 2023).

#### 5.2 Student's Awareness

As mentioned in the section on methodology, we conduct research by qualitatively analyzing data collected through semi-structured interviews. In this part, we present a sample of the collected data based on the recordings of interviews conducted with thirteen students participating in STEM inquiry-based learning within the context of the circular economy. The quotations in this section are organized in accordance with the five dimensions of the survey (Ozden, 2008). Cronbach alpha determined that the internal consistency of the scale was 0.88.

#### 5.2.1 Awareness of environmental problems-AEI

In this dimension, awareness of environmental problems focuses on students' awareness of environmental challenges and their understanding of the impact of these issues on everyday life. Students commented in respect of the problem of environmental. Student 01 mentioned that "If there is an issue with plastic garbage near the school, it may harm us". In the meantime, student 04 told that: "It is dependent on whether people care about the environment. If people do not care for the environment, it can lead to issues in the future". Moreover, student 09 said that: "It takes a long time for plastic trash to break down. The more people use it, the worse it is for the earth".

Based on the aforementioned comments, it is evident children possess a heightened awareness about environmental concerns. Furthermore, it is important for students to understand the implications of these challenges on their daily lives.

#### 5.2.2 Awareness of individual responsibility-AIR

In this part, awareness of individual responsibility focus on awareness of individual environmental responsibilities among students. In this part, the awareness of individual responsibility focuses on students' awareness of their individual environmental responsibilities. All students have indicated that the collective responsibility for the preservation of the environment extends beyond the authority of the government or industrial entities. Student 08 told that: "Every country should work to solve environmental problems, not just one. This is because the effects of environmental problems may not only affect the country in issue, but also the countries around it". Additionally, student 10 told that: "If we don't take care of the environment, our hometown will be full of trash, which will cause problems". Furthermore, student 13 said that: "People use a lot of plastic every day, which means that there is a lot more trash. We are the ones who should use less".

From the mentioned quotations, students' awareness of their environmental responsibilities is emphasized in their awareness of individual responsibility. In addition, they presented ideas for environmental management.

#### 5.2.3 General attitudes towards environmental solution-GAES

In this dimension, the general attitudes towards environmental solutions focus on the perspectives of students on potential solutions to environmental issues. The opinions of students about possible solutions to environmental concerns were discussed. Student 08 said that: "People's lack of care for conservation may result in an around the world trash overflow". Moreover, student 04 told that: "Someone does not throw trash into the garbage container, that it might be a problem of plastic waste issue. So, we should start doing it by ourselves". In addition, student 09 said that: "We can start by doing simple things, like making toys or other helpful things out of plastic trash instead of just throwing it away."

The students' quotation highlighted their views on potential solutions for addressing environmental challenges. From the perspective of students, it is the method through which they can initiate initiatives inside their own environment and individually.

#### 5.2.4 General attitudes towards environmental problems-GAEP

For this part, general attitudes towards environmental problems focus on students' perspectives on the impact of environmental concerns on daily life. Students expressed their opinions about the influence of environmental issues. Student 07 told that: "If you throw the trash into the water, fish and other marine animals might eat it or get caught in it, which could kill them." Student 01 told that: "People shouldn't destroy nature because making too much trash is bad for us. For example, different pollutants that are released and breathed in can hurt our health". Student 13 concludes and points out: "There is a lot of plastic in the world as several companies produce many types of plastic. Factories also make a lot of other kinds of hazardous waste."

In these quotations, students emphasize the factors that have an adverse impact on the environment and the country's inhabitants. Meanwhile, students continue to be oblivious to how technology can be environmentally damaging.

#### 6. Conclusion and Limitation

This study examined the impact of a technology-integrated pedagogy of STEM learning in the citizen science activity through inquiry on fourth-grade students, and the findings revealed that students improved their conceptual understanding of different types of plastic and developed a positive attitude toward the environment after participating in the citizen inquiry activity. Consequently, it implied that the possibility of promoting primary students' integration of various technologies into STEM inquiry-based learning in a circular economy framework might be an effective strategy to increase their environmental awareness. However, the number of participants in this study was limited, as they were chosen from a science class at a university-based demonstration school located in the northeastern part of Thailand. This limits the generalizability of the findings for students in primary schools in Thailand. Furthermore, in order for this strategy to be worth future study, the control and experimental groups of participants must be improved in order to compare treatment in each approach. Nevertheless, this study provides evidence of a positive starting point for designing ICT technology-integrated STEM inquiry-based learning of citizen science activity in primary students.

#### Acknowledgements

This contribution was financially supported by Khon Kaen university, Thailand. The author would like to express gratefully thanks to all participants. This study also supported by Khon Kaen University demonstration school international division, Khon Kaen university, Thailand.

#### References

- Abidin, R., Suryani, N., & Sariyatun. (2020). Students' Perceptions of 360 Degree Virtual Tour-Based Historical Learning About the Cultural Heritage Area of the Kapitan and Al-Munawar Villages in Palembang City. *International Journal of Social Sciences and Management*, 7(3), 105-112.
- Abouhashem, A., Abdou, R. M., Bhadra, J., Siby, N., Ahmad, Z., & Al-Thani, N. J. (2021). COVID-19 inspired a STEM-based virtual learning model for middle schools—a case study of Qatar. *Sustainability*, *13*, 2799.
- Adita, A., & Srisawasdi, N. (2022). Plastic island game: A digital game for facilitating citizen inquiry pedagogy in school science education. In S. Iyer et al. (Ed.), *Proceedings of the 30th International Conference on Computers in Education* (pp. 374-384). Malaysia: Asia-Pacific Society for Computers in Education.
- Alburaiki, A. A., Syed Abdullah, S. I., & Md Khambari, M. N. (2022). Enhancing students' interest in STEM-related subject at Omani post-basic schools through application of augmented reality. In S. Iyer et al. (Ed.), *Proceedings of the 30th International Conference on Computers in Education* (pp. 318-324). Malaysia: Asia-Pacific Society for Computers in Education.
- Araya, R. (2023). Connecting classrooms with online interclass tournaments: A strategy to imitate, recombine and innovate teaching practices. *Sustainability*, *15*, 8047.
- Aristeidou, M., Scanlon, E., & Sharples, M. (2013). A design-based study of Citizen Inquiry for geology. EC-TEL 2013 Eighth European Conference on Technology Enhanced Learning. Cyprus.
- Aristeidou, M., Scanlon, E., & Sharples, M. (2015). Weather-it: Evolution of an Online Community for Citizen Inquiry. *Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business*, (pp. 1-8). Austria.
- Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience*, *59*, 977-984.
- Buck, L. B., Bretz, S. L., & Towns, M. H. (2008). Characterizing the level of inquiry in the undergraduate laboratory. *Journal of College Science Teaching*, 52-58.

- Bybee, R. W. (2013). The case for STEM education: Challenges and opportunities. Arlington, Virginia: NSTA Press.
- Chaipidech, P., & Srisawasdi, N. (2018). A proposal for personalized inquiry-based flipped learning with mobile technology. In Y. -T. Wu et al. (Ed.), *Proceeding of the 26th International Conference on Computers in Education* (pp. 344-353). Philippines: Asia-Pacific Society for Computers in Education.
- Chaipidech, P., & Srisawasdi, N. (2023). Integrating 360° virtual learning environment to support outof-class inquiry activity for preservice teachers: A preliminary study. *The 6th Eurasian Conference on Educational Innovation (ECEI)* (pp. 163-166). Singapore: IEEE.
- Chen, N. -S., & Hwang, G. -J. (2014). Transforming the classrooms: innovative digital game-based learning designs and applications. *Educational Technology Research and Development*, 62, 125-128.
- Chu, W. W., Ong, E. T., Ayop, S. K., Azmi, M. S. M., Abdullah, A. S., Karim, N. S. A., & Tho, S. W. (2023). The innovative use of smartphone for sound STEM practical kit: a pilot implementation for secondary classroom. *Research in Science and Technological Education*, *41*(3), 1008-1030.
- Freeman, B., Marginson, S., & Tytler, R. (2019). An International View of STEM Education. *In STEM Education 2.0.* Brill.
- Garcia, M. B., Nadelson, L. S., & Yeh, A. (2023). "We're going on a virtual trip!": a switching-replications experiment of 3 6 0 degree videos as a physical field trip alternative in primary education. *International Journal of Child Care and Education Policy*, *17*(4).
- Herodotou, C., Fernández, E. V., & Sharples, M. (2014). Scaffolding citizen inquiry science learning through the nQuire toolkit. *EARLI SIG 20,* (pp. 1-3). Sweden.
- Hsiao, P. -W., & Su, C. -H. (2021). A study on the impact of STEAM education for sustainable development courses and its effects on student motivation and learning. *Sustainability*, 13, 3772.
- Kelley, T. R., & Knowles, J. G. (2016). A Conceptual framework for integrated STEM education. International Journal of STEM Education, 3(11).
- Lindner, et al. (2022). Development of an app and teaching concept for implementation of hyperspectral remote sensing data into school lessons using augmented reality. *Remote sensing, 14, 791.*
- Meinhold, J., & Malkus, A. (2005). Adolescent environmental behaviors: can knowledge, attitudes, and self-efficacy make a difference? *Environment and Behavior*, 37(4), 511-532.
- Mohr-Schroeder, M. J., Bush, S. B., Maiorca, C., & Nickels, M. (2020). Moving toward an equity-based approach for STEM literacy. *In Handbook of Research on STEM Education*. Routledge.
- Nguyen, T. P. (2023). Integrating circular economy into STEM education: A promising pathway toward circular citizenship development. *Frontier in Education*, 1-9.
- Nguyen, T. P., Nguyen, T. H., & Tran, T. K. (2020). STEM education in secondary schools: Teachers' perspective towards sustainable development. *Sustainability, 12,* 8865.
- Nizaar, M., Sukirno, M., Djukri, M., Muhardini, S., & Mas'ad, S. (2020). Improving students' environmental awareness using 3R principles. *Universal Journal of Educational Research*, 8(11), 6146-6151.
- Ozden, M. (2008). Environmental Awareness and Attitudes of Student Teachers: An Empirical Research. *International Research in Geographical and Environmental Education, 17*(1), 40-55.
- Rupp, M. A., Odette, K. L., Kozachuk, J., Michaelis, J. R., Smither, J. A., & McConnell, D. S. (2019). Investigating learning outcomes and subjective experiences in 360- degree videos. *Computers & Education*, 128, 256-268.
- Sellami, A. L., Al-Ali, A., Allouh, A., & Alhazbi, S. (2023). Student attitudes and interests in STEM in Qatar through the lens of the social cognitive theory. *Sustainability*, *15*, 7504.
- Sharples, M., Aristeidou, M., Fernández, E. V., Herodotou, C., & Scanlon, E. (2017). The Sense-it app: A smartphone sensor toolkit for citizen inquiry learning. *International Journal of Mobile and Blended Learning*, 9(2), 16-38.
- Thumtathong, A., Premthaisong, S., & Srisawasdi, N. (2019). Teaching english with science: A result of content and language integrated learning approach and mobile-assisted inquiry pedagogy. In J. L. Shih (Ed.), *Proceedings of the 27th International Conference on Computers in Education* (pp. 286-293). Tiwan: Asia-Pacific Society for Computers in Education.
- Torre, R. d., Onggo, B. S., Corlu, C. G., Nogal, M., & Juan, A. A. (2021). The role of simulation and serious games in teaching concepts on circular economy and sustainable energy. *Energies*, *14*, 1138.

## Enhancing Understanding of Complex Systems through Analogy-Based Video Scenarios

Meera PAWAR\*, Sheeja VASUDEVAN & Sahana MURTHY IDP in Educational Technology, Indian Institute of Technology Bombay, India \*214380002@iitb.ac.in

**Abstract:** Biological education grapples with the challenge of teaching intricate systems, particularly the complex human immune system. Existing teaching methods often fall short of conveying these complexities effectively. Analogies and metaphors have proven successful in teaching complex biological systems across educational levels. We developed video scenarios connecting a wound-healing phenomenon with a real-life analogy, incorporating reflection spot activities for students to map elements based on structure, function, and behavior. Our study with high school students revealed significant improvements in their understanding of the wound-healing phenomenon through analogy mapping and reasoning. The study's contribution lies in an effective pedagogical and technological design for complex biological systems.

Keywords: Complex biology system, SBF framework, Analogical reasoning

#### 1. Introduction

Biological education involves mastering complex, interconnected systems characterized by multilevel organization, diverse components, and dynamic processes, often posing challenges in understanding the relationships among facts (Hmelo-Silver et al., 2007). To comprehend such systems, it's crucial to learn their components, functions, and processes. The Human Immune system, a prime example of complexity, combines humoral and cellular interactions, frequently causing confusion among students (Su, Cheng, & Lin, 2014). While various teaching strategies have been explored, an educational card game using combat metaphors improved interest but lacked focus on cellular mechanisms (Su et al., 2014; Steinman & Blastos, 2002). Similarly, the Humanology Game integrated game features with immunology concepts, enhancing procedural knowledge (Cheng et al., 2014). These games employ metaphors or analogies, aiding students in connecting new concepts to familiar ideas.

Analogies and metaphors are valuable tools for teaching complex biological systems across various educational levels, including high school and undergraduate settings. Brown (2003) emphasized the importance of analogy as a universal learning mechanism at all ages. The structure-behavior-function theory posits that understanding any complex system involves analyzing its structures, functions, and how it achieves its purpose (Goel et al., 1996). Research has shown that even elementary school children can enhance their understanding of function and mechanism in complex systems through analogical modeling exercises, such as mapping respiratory system structures to the mechanism of human breathing (Moonhyun Han1 & Heui-baik Kim, 2017). This underscores the potential of analogies and analogous models for fostering learners' knowledge construction and reasoning skills related to the structure, mechanism, and function of complex systems. While strategies like card games and VR-based games use analogies or metaphors, they often focus on a superficial level rather than delving into the structure, behavior, and function aspects.

This study proposes a design of video scenarios based on the SBF theory and implements it in developing a wound-healing phenomenon from the human immune system and its corresponding analogy from a real-life situation, the water leakage process. The aim of this research was to investigate high school learners' ability to map the base analogy of "Water leakage at home" to the target concept of "Wound healing" as this is a new complex concept for high school students and understand their analogical reasoning through the SBF theory perspective.

## 2. Theoretical Framework

## 2.1 Analogy and Analogical Reasoning

Analogy and Analogical Reasoning Analogies involve comparing dissimilar concepts to uncover similarities, aiding comprehension of unfamiliar subjects. Gentner and Gentner (1983) argue that analogies enhance cognitive development by connecting existing knowledge to new concepts. This process involves relating the known (analog) to the unknown (target). Analogical reasoning is a multifaceted cognitive process involving steps such as identifying relevant information, extracting relationships, and mapping dissimilar domains for inferences and common principles (Trey & Khan, 2008). Mapping eases knowledge transfer from a source domain to a destination domain (Gentner 1983). Analogies assist in structuring concepts (Celik, Kirindi, & KOTAMAN, 2020). For instance, a bank robbery analogy helped students grasp hormone synthesis, transport, and mechanism concepts (Cardinali et al., 2022).

## 2.2 Structure-Behavior-Function Framework

Structure-behavior-function (SBF) theory models complex systems, considering components, their functions, and enabling mechanisms. This representation aids comprehension and causal reasoning (Goel, 1996). Researchers used SBF theory and augmented reality (AR) to teach the circulatory system, yielding a better understanding (Gregorčič & Torkar, 2022). SBF is a valuable tool for comprehending complex systems (Hmelo-Silver & Pfeffer, 2004). We employ this framework to design video scenarios promoting analogical reasoning in students regarding the human immune system."

## 3. Design and Development of Scenarios

#### 3.1 Designing video Scenarios

A wound-healing phenomenon from the human immune system was considered for the development of the scenario as it is a complex phenomenon that includes different types of blood cells, skin layers, proteins, and other factors that come together to perform an action. Its corresponding analogy was drawn from the day-to-day life routine i.e. water leaking at home. When water starts leaking, there is a mechanism to stop the leakage where different factors such as human factors, and other materials come into play. While designing videos for the analogy and scientific phenomena, mappings at structure, behavior, and function levels were considered. Table 1 shows elements from the water leakage scenario mapped to elements from the wound healing scenario. For example, Broken Skin from the wound healing phenomenon is mapped to plastic blocks stopping the leakage, this mapping is at the function level. Mechanism level mapping is done in the process of healing and stopping water leakage.



 Table 1: Design and Development of mapped elements in video scenarios

We created the wound healing video scenario in two ways: one without reflection spot activity and another with reflection spot activity using Edpuzzle's interactivity feature. The reflection spot activity includes multiple-choice questions in that the student has to map elements from the water leakage analogy to the wound healing concept, followed by reasoning for that mapping. There are a total of 5 mapping and reasoning questions. The analogy video scenario and the wound healing scenario with reflection spot activity were uploaded to the newly created 'Science Classroom' on the Edpuzzle platform (Mischel, L. J.,2019) - https://edpuzzle.com/open/cepbowi

## 4. Methodology

The purpose of this study is to determine whether students are able to identify and map the elements from the analogy water leakage scenario to the wound healing phenomenon scenario. Their analogical reasoning process will also shed light on how they grasp the target concept at the level of structure, function, or mechanism. This study is based on the following research question: "In what ways do analogy-based video scenarios help learners in their understanding of a complex system at three different levels (structure, function, and mechanism)?"

## 4.1 Participants and Procedure

In this study, a total of 12 students from grades 6 to 8 participated. They undertook two tasks: The initial task involved viewing a wound-healing scenario without reflection spot activity, followed by a pre-test questionnaire. In the subsequent task, students received instructions on using the Edpuzzle platform and were guided through the steps to complete the activity. During this phase, participants watched an analogy video scenario (Water leakage story) followed by a wound-healing phenomenon scenario. They engaged in a reflection spot activity to map elements and provided reasoning for each mapping. Following this, a post-test, similar to the pre-test, was administered, and a semi-structured interview concluded the study.

## 4.2 Data collection and analysis

Out of a total of 12 students, two students did not finish the activities entirely, while one student completed every activity except the pre-test. Thus, the data analysis of the reflection spot activity includes data from 10 students, whereas pre- and post-data include information from 9 participants. Throughout this study, four data sources were employed: pre-test, reflection spot activity questions, post-test, and interviews. The pre-test and post-test gathered data through a Google form, having open-ended questions like "Identify elements from the wound healing story" and "Describe or explain the wound healing process in proper sequence".

Reflection spot activity data was collected in the form of MCQ and open-ended questions, for example, "Which element from the analogy scenario is similar to the skin in this phenomenon scenario" was the MCQ question followed by a question for reasoning "In what aspects both the elements are similar?". A semi-structured interview was taken after students completed the activity. The questions for the interview were based on the responses students gave during the activity.

The SBF theory served as the framework for the analysis of the data. In order to determine if a student's thinking was at the level of structure, behavior, or function, a grading rubric based on the SBF theory was developed, and students' responses were examined using this criterion. Table 2. shows a scoring matrix for the reflection spot activity, Mappings are graded based on answers to multiple-choice questions, and reasoning is examined to see whether any points from the structure, function, and behavior level descriptors are present.

	Skin and Roof	Mentioned at least one of the phrases or words from the description	Did not mention even a single word/phrase from the descriptors
	Identification and Mapping	If done correctly (1)	If not done correctly(0)
Structure	<ol> <li>Outer layer or a boundary</li> <li>Damage or the broken area is similar in the skin and roof.</li> <li>The roof is like skin and the pipe is like a blood vessel</li> <li>Roof had a leakage similarly Skin was bleeding.</li> </ol>		
Function	<ol> <li>1) Skin Covers/ protects us similarly roof protects the house.</li> <li>2) Roof had a pipe that had water similarly skin has blood vessels that have blood flowing in it. (carrying water and blood)</li> </ol>		
Behavior/ Mechanism	1) Skin <b>prevents entry</b> of external bacteria from coming inside our body similarly roof is preventing rain or insects from coming inside the house.		

Table 2. Scoring rubric for reflection spot activity

A score of "1" was assigned if these features were identified, while a score of "0" was given if they were entirely absent from the reasoning. For pre-test and post-test responses similar scoring rubric was developed on 3 scales. If the responses included most of the structures, their function, and the proper sequence of the wound healing process then it was scored "3". If half of the structures, functions, or sequence were correct then it was scored "2" and if only a few things mentioned were a correct or completely incorrect sequence of the process then it was scored "1".

## 5. Result and Findings

## 5.1 Understanding students' analogy mapping and reasoning

Analysis of individual students' analogy mapping and reasoning of reflection spots showed that all of the students were able to map the base analogy to the target concept with a high degree of accuracy. However, the reasoning varied for the structure-function behavior. Figure 1 below represents the reflection spot activity data of 10 students.



Figure 1: Reflection spot activity Scores

## 5.2 Analyzing Pre and Post-test Scores

A paired t-test was performed to evaluate if there was a difference between the learning gain from the pre-test to the post-test. The data was checked for normality assumption. The results showed that there was a statistically significant difference between the average of the pre-test (*Mean-4.11, SD-1.27*) and post-test scores (*Mean –6.78, SD-1.92*) shown in Table 3., with a p-value of 0.0014 (< 0.01). The results of the paired t-test suggest that the use of analogy was effective in improving learning. The students who used analogy showed a significant improvement in their scores from the pre-test to the post-test.

Scores	Mean	N	SD	SEM					
Pre-test	4.11	9	1.27	0.42					
Post-test	6.78	9	1.92	0.64					

Table 3: Paired t-test Statistics

# 5.3 Correlation between reflection spot activity score and students' post-test performance

To check if the reflection spots and reasoning activity helped learners in answering the posttest questions, we performed a Pearson correlation test between scores on reflection spot activity and post-test scores. The variable reflection spot activity score and post-test score were found to be weakly correlated r (0.276) but statistically not significant p (0.44).

## 6. Discussion

In this article, we explored high school students' ability to connect the "Water leakage at home" analogy with the concept of "Wound healing" using video scenarios based on the SBF theory. Our findings indicate that these analogy video scenarios effectively enhance students' understanding of complex systems at three different levels: structure, function, and behavior.

Specifically, 70% of students successfully mapped elements from the analogy to wound healing, while 70% understood the structural aspects with 40% accuracy, 80% grasped the functional aspects with over 60% accuracy, and 60% comprehended the mechanisms with 40% accuracy. The significant improvement in learning from the pre-test to the post-test underscores the pedagogical value of using analogies (Cardinali et al., 2022). However, we noted a statistically non-significant low correlation between the reflection spot activity scores and post-test scores, possibly due to the small sample size, prompting further exploration of the causal relationship between activity performance and subsequent test results.

This study contributes by using the SBF framework to map real-world scenarios to scientific phenomena through video-based scenarios and reflection spot activities. While our approach showed promise, its generalizability is limited by the study's focus on a single analogy and concept. Future research should involve a more diverse range of analogies and concepts and include larger sample sizes for improved statistical power and result reliability. We intend to continue designing analogical scenarios based on the SBF theory, providing scaffolding for analogical reasoning and investigating the impact of different reflection questions on students' learning with analogies to gain further insights.

#### Acknowledgments

We would like to thank our colleagues Rajashri Priyadarshini and Ishika for providing their time in advising and reviewing the study materials.

#### References

- Cardinali, C. A., Martins, Y. A., Prates, R. P., de Araújo, E. V., Costa Viana, F. J., de Sousa, M. E., ...
   & Nunes, M. T. (2022). The bank robbery analogy as a first approach for understanding basic concepts of synthesis, transport, and mechanism of action of hormones. *Advances in Physiology Education*, 46(4), 724-727.
- Celik, H., Kirindi, T., & KOTAMAN, Y. A. (2020). The effect of the computer-based analogy used in science teaching on learning outcomes. *Journal of Turkish Science Education*, *17*(1), 73-93.
- Cheng, M. T., Su, T., Huang, W. Y., & Chen, J. H. (2014). An educational game for learning human immunology: What do students learn and how do they perceive?. *British Journal of Educational Technology*, *45*(5), 820-833.
- Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. *Mental models*, *99*, 129.
- Goel, A. K., Gómez de Silva Garza, A., Grué, N., Murdock, J. W., Recker, M. M., & Govindaraj, T. (1996). Towards design learning environments—I: Exploring how devices work. In *Intelligent Tutoring Systems: Third International Conference, ITS'96 Montréal, Canada, June 12–14, 1996 Proceedings 3* (pp. 493-501). Springer Berlin Heidelberg..
- Gregorčič, T., & Torkar, G. (2022). Using the structure-behavior-function model in conjunction with augmented reality helps students understand the complexity of the circulatory system. *Advances in Physiology Education*, *46*(3), 367-374.
- Han, M., & Kim, H. B. (2019). Elementary students' modeling using analogy models to reveal the hidden mechanism of the human respiratory system. International Journal of Science and Mathematics Education, 17, 923-942.
- Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive science*, 28(1), 127-138.
- Mischel, L. J. (2019). Watch and learn? Using EDpuzzle to enhance the use of online videos. *Management Teaching Review*, 4(3), 283-289.
- Steinman, R. A., & Blastos, M. T. (2002). A trading-card game teaching about host defence. *Medical education*, *36*(12), 1201-1208.
- Su, T., Cheng, M. T., & Lin, S. H. (2014). Investigating the effectiveness of an educational card game for learning how human immunology is regulated. *CBE—Life Sciences Education*, 13(3), 504-515..
- Trey, L., & Khan, S. (2008). How science students can learn about unobservable phenomena using computer-based analogies. *Computers & Education*, *51*(2), 519-529.
# Effectiveness of Learning Seminar Course to Promoting Research Skills : Observe-Plan-Organize-Present in Science Project

Chitphon YACHULAWETKUNAKORN<sup>a\*</sup>, Witsanu SUPANDEE<sup>b</sup> & RATTHAKARN NA PHATTHALUNG<sup>c</sup>

<sup>a,b,c</sup>Engineering Science Classroom, King Mongkut's University of Technology Thonburi, Thailand \*chitphon.yac@kmutt.ac.th

**Abstract:** This paper attempted to study the effectiveness of learning approaches between Seminars Course, and Science Project Course in the context of research skill development; Observation, Planning, Organizing, and Presentation skills. Besides, each process of learning could increase the essential research skills, This process required learner's effort to work in group within two semester. The learners received and continued to this CRSS approach the high level of learning achievement in the context of research skills; Observation, Planning, Organizing, and Presentation. Furthermore, the finding of this study could be expressed to the correlative soft skills could be used more than one course, which can be designed others curriculum involved.

Keywords: Seminars Course, Research Skills, Science project

# 1. Introduction

Science Project's considered as a significant learning approach in every level. Exploring Science, Technology, Engineering, and Science Experiment are common supporting Science Project in High School. There were various definition, Generally, Science Project were defined as an educational activity for students involving experiments or construction of models in one of the science disciplines. Students could present their science project, So it could be called a science fair project. In this present, Science projects were classified into four main types. For instance, Experimental projects, Engineering projects, Display projects, and Theoretical projects.

Science Project development were basically developed by research methodology. For instance, research questions, hypothesis, research design, etc. Which are significant elements of research development inevitably. In addition, soft skill supporting Science project development considered as the other necessary skill called "Research skill". They also be defined as the ability to search for, locate, extract, organize, evaluate, and present information that is relevant to a particular topic. Furthermore, there are various skill promoting research skill such as, Time management, Communication skill, Information gathers.

Research skill are continuously significant for Science Project development. Sanchez, Marcos, Rivero, and Ruiz (2017) founded that learners could develop soft skills during exploring Science Project (Sanchez, Marcos, Rivero, & Ruiz, 2017). According to the Engineering Science Classroom Program, Students were required to study the Seminars Course supposed to train research skills, research methodology in order to preparing student for Science Project development in Grade 11<sup>th</sup>. Hence, various soft skills in the seminars course were used to create and develop Science Project.

According to the correlation between Seminars course, and Science Project course, authors assumed to study the effectiveness of learning Seminar Course to Promoting Research Skills in Science Project Course by using research skill as the criteria to assess.

# 2. Related Study

#### 2.1 Seminars Course and Science Project

The Word "Seminar" come from Latin, and it means "seed". In general, the seminar method encouraged active participants, a group of participants were guided to interact with each other in class. In addition, the seminar defined as the applicable to settings from conferences to other meetings or training. Most of Seminar's topic are the research seminars. Nevertheless, this course was popularized in the senior high school and undergraduate level to discuss, exchange, and construct knowledge too. The students were provided with the opportunities to present their papers which supported presentation and public-speaking skills. McMullin (2014) also defined that this learning approach often accompanied by a visual presentation and even hands-on (McMullen, 2014). In educational institutions of all levels, there are various objectives. The seminar promoted the basis learning of students. It focused on instruction that was designed for the student to gain specific knowledge within the academic area. In the same time, seminars were an approach for people to increase their educational background through lectures, training, and possible experimentation.

Science Project's considered as the educational activities supporting experiment or construction of model in one of science discipline. This learning approach continuously popularized over decades. In particular STEM. In addition, the elementary level increasingly taught Science course through science project. Initially, Science project trained and encourage learner the concept of Scientific inquiry. By the way, there were some disruptive physical learning methods after the Covid-19 pandemics, traditional learning were transformed into virtual learning, include laboratory experiment. Science Projects is another fundamental activity creating and constructing knowledge (Jumaat, Dayana, Tasir, & Ashari, 2017). Moreover, Science Project was the other method support self–efficacy in Secondary school in STEM education. Nowadays, Science Project was scoped in the form of Project-based learning. Haatainen, and Aksela (2021) called this approach as "Project-based science learning" (Haatainen, & Aksela, 2021).

According to the definition and significance of Seminar course and Science Project, there are some correlations both in the context of learning methodology. Seminars course initially encouraged learner exchanged knowledge through discussion. Hence the necessary skills are Critical thinking, Communication, and Reading literacy. Similar to science project, learners were encouraged by the intention and hypothesis that curious for various type of phenomenon.

#### 2.2 Research Skills

Research skills defined as the ability to search for extract, organize, evaluate and present information that is relevant to a particular topic. Academic research was a specific type of research. It involved intensive search, investigation, and critical analysis response to a research question or hypothesis. In addition, Mydin, Rahman, and Mohammad (2021) defined as research skills are the ability to find an answer to a question or a solution to a problem (Mydin, Rahman, & Mohammad, 2021). For instance, gathering information, analyzing, and interpreting.

Research skills also considered as a significant in education, it started from observation, learners could create research question, and research topic which's relevant. Secondly, planning, a skill helped to determine research design. This process comprised of data collection, hypothesis, design experiment, data analysis. Meerah, Osman, Zakaria, Ikhsan, Krish, Lian, and Mahmod (2012) studied the RSD, Research Skill Development framework in

higher education around the development of skills associated with research skills, and interdisciplinary projects (Meerah, Osman, Zakaria, Ikhsan, Krish, Lian, & Mahmod, 2012).

The definition of research skills was defined in various perspectives. In particular, the educational sector. Moreover, somehow, those skills were used to improve and develop teaching, learning method (Attakorn, Tayut, Pisitthawat, & Kanokorn, 2014).

According to the significance of research skills, in particular term of education, they were essential for creating and developing Science project. Similar to Salybekova, Issayev, Abdrassulova, Bostanova, Dairabaev, and Erdenov (2021), Project-based science learning's also considered as an approach consorting Science Project development in every process of Science Project (Salybekova, Issayev, Abdrassulova, Bostanova, Dairabaev, & Erdenov, 2021).

#### 3. Description of Correlation of Research Skills between Seminars and Science Project

#### 3.1 Background and Overall Structure

Research skills was an essential skill, they could encourage learners to get more efficient learning. The research skills consisted of various skills: formulating questions, observing, planning, collecting data, organizing data, and presenting research. Both of courses are the required course in the Engineering Science Classroom, King Mongkut's University of Technology Thonburi for Grade 10<sup>th</sup> and 11<sup>th</sup> students. Student could practice the research skills in those programs.



Figure 1. Correlation of Research Skills between Seminars and Science Project (CRSS).

Seminar course is the course for Grade 10<sup>th</sup> students. This course is provided into three main processes, (Figure 1) the first one is search and find journal article. Then, the learners would read the selected article clearly by discuss with the group members and their advisor both in virtual world, and reality. After that, each group prepared a presentation. In the last process of this course, each group would present the article. According to the learning process, learners were required to develop the skills as follow; observation, planning, organization and presentation that categorized as a part of research skills. Science project course is a course for Grade 11<sup>th</sup> students. The process consists of four steps. First, choose the research topic and do a literature review. Second, learners wrote a research proposal and make a presentation material. After that, they could start the experiment. In this process the learners were guided by research advisor both in virtual and reality. For the last process, learners analyzed the result, and also made a presentation for the results. Hence, in the end of this course, learners obtained observe skill, planning skill, organize skill and present skills that are significant research skill like the seminar course. (Figure 1)

# 3.2 Overall of Correlation between Seminars and Science Project Course

From the process of seminar course, the students developed the research skills such as observation skill, planning skill, organize skill as well as presentation skill that encouraged the science project course activity. (Table 1)

Research	Seminar Course and Technology	Science Project and Technology
Observation skill Planning Skill	Learners obtained observation skill through searching an interested journal article and analyzing the credibility of source. In context of Technology-supported, leaners mostly used the credible search engine to find the article. For instance, Google scholar, Research gate, Science direct, etc. Learners could develop Planning skill through a groupwork planning. They had to work together in ordered to create their comprehensive abstract. In the context of Technology-supported, learner used	Learners obtained observation skill through research topic's creation process: Finding research topic, doing literature review as well as the article finding in the Seminars course. In context of Technology-supported, leaners also used the search engine to find the credible data source for doing the literature review in this process. Learners developed Planning skill through same as in the seminars in term of planning the overall project development. According to learning process, leaners also used various application. For instance, Zoom for
Organizing Skill	various application. For instance, Zoom for consulting with advisor, Chat GPT for writing guideline. Learners could gain organizing skill through preparing to present process. It started from presentation making. Leaners had to prepare slide and practice to present so they needed to design and organize whole data in Presentation Slide. In the context of Technology- supported, learners popularized to use Canva for making presentation, and Zoom for virtual consulting with advisor.	application. For instance, 200m for consulting with research's advisor, Miro, Chat GPT, etc. for making presentation. The organizing skill obtained from seminar course could support the experiment design, doing experiment as well. They could organize and design the experiment follow by research proposal. According to doing experiment process, learners rarely used a virtual conference except collecting data.
Presentation Skill	Learners could gain presentation skill. In this process, learners had to answer and discuss in class. Hence leaners developed the presentation skills. In the context of Technology- supported, learners rarely used the technology-supported except their own gadgets usage.	Presentation is the last process of learning both in the Seminars, Project course. Learners could develop the presentation skill through answering, discussing in the class. According to those process, learners rarely used the technology-supported except their own gadgets usage.

Table 1. The Correlation between Research Skill in Seminar and Science Project Course Activity.

# 4. Research Design

#### 4.1 Participant

The participant in this study were 81 Grade 11<sup>th</sup> students in Engineering Science Classroom, King Mongkut's University of Technology Thonburi. All students passed Seminar course in Grade 10<sup>th</sup> and Science project course in Grade 11<sup>th</sup>.

# 4.2 Procedure

The procedure of CRSS were divided into 2 parts, the first one was beginning with Seminar course (figure 2) that consists of 4 main sections including search and find journal article, reading comprehended and discussion, preparing to presentation. The second section was a Science project course that consists of 4 main sections too, including chose a research topic and do a literature review, write a proposal, and prepare to presentation, do the experiment and data analyze as well as presentation.



Figure 2. The process of Seminar Course and Science Project Course

# 4.3 Data Collection

The Questionnaire was designed into 4 items that participant gained. For instance, Observation, Planning, Organizing, and Presentation skill variables in ordered to evaluate the effectiveness of each research skill that learner developed from the seminars course which's continue to the Science Project course in Grade 11<sup>th</sup>.

# 4.4 Data Analysis

Research skills performance was conducted from 81 students who attended the Seminars Course in Grade 10<sup>th</sup>, and Science Project Course in Grade 11<sup>th</sup>. All quantitative data were analyzed using mean and standard deviation. For example of question, According to seminars course learning, Which is the most effective skill that learners develop from?., Which is the most effective skill from the seminars course that learners use in Science Project Course?.

# 5. Result

# 5.1 Learning Achievement

The self-evaluation score of learners in the development of research skills; observation, planning, organizing, and presentation skills from seminar course was ranging from  $3.47\pm0.78$  to  $3.91\pm0.85$  (table 2) and the self-evaluation score of learners in the development of research skills from Science Project Course was ranging from  $3.44\pm0.80$  to  $3.89\pm0.75$  (table 2).

Table 2. Self-evaluation in The Development of Research Skill in the Seminars and Science *Project* 

Research skill	Mean ± SD	
	Seminars Course	Science Project
Observation	3.75±0.82	3.44±0.80
Planning	3.54±0.76	3.89±0.75

Organizing	3.91±0.85	3.78±0.83
Presentation	3.47±0.78	3.55±0.79

# 5.2 The Correlation between Research skills development from Seminar Course and Science Project Course

This development showed different levels in each dimension (Table 2). The result showed that the participants were able to develop research skills (Observation, Planning, Organizing and Presentation) derived from the Seminar Course to work on the project in Science Project Course (Figure 3).



Figure 3. Correlation of Research skills development from Seminars and Science Project

#### 6. Conclusion

In this context, there can be no doubt that the Seminars, and Science Project are significantly develop and improve the research skills in the criteria as follow: Observation, Planning, Organizing, and Presentation. One of the consequences was an increase in the number of students' capabilities acquiring both in an academics, and soft skill supporting research skills. According to the correlation between research skills in the Seminars, and Science Project Course, there are significantly concluded that students derived Planning, and Organizing skills as the most effective skill in both two course. Finally, the development of research skill both in two course can be consort with Meerah, Osman, Zakaria, Ikhsan, Krish, Lian, and Mahmod (2012), Attakorn, Tayut, Pisitthawat, and Kanokorn (2014) and Salybekova, Issayev, Abdrassulova, Bostanova, Dairabaev, and Erdenov (2021) that Research Skill Development could support interdisciplinary projects, and improved teaching, and learning approaches.

# References

- Attakorn, Tayut, Pisitthawat & Kanokorn. (2014). Soft Skills of New Teachers in the Secondary Schools of Khon Kaen Secondary Educational Service Area 25, Thailand. *Procedia Social and Behavioral Sciences*, *112*, 1010–1013.
- Sanchez, Marcos, Rivero, Ruiz. (2017). Project Management Training: An Integrative Approach for Strengthening the Soft Skills of Engineering Students. *International Journal of Engineering Education.* Vol.33, pp.1912-1926
- Haatainen, O., & Aksela, M. (2021). Project-based learning in integrated science education: Active teachers' perceptions and practices. *LUMAT*, *9*(1), 149–173.
- Jumaat, N. F., Tasir, Z., Halim, N. D. A., & Ashari, Z. M. (2017). Project-based learning from constructivism point of view. *Advanced Science Letters*, 23(8), 7904–7906.
- McMullen, V. B. (2014). Using Student-Led Seminars and Conceptual Workshops to Increase Student Participation. College Teaching, 62(2), 62–67.
- Meerah, T. S. M., Osman, K., Zakaria, E., Ikhsan, Z. H., Krish, P., Lian, D. K. C., & Mahmod, D. (2012). Developing an Instrument to Measure Research Skills. *Procedia - Social and Behavioral Sciences*, 60, 630–636.
- Mydin, F., Rahman, R. S. A. R. A., & Mohammad, W. M. R. W. (2021). Research Collaboration: Enhancing the Research Skills and SelfConfidence of Early Career Academics. Asian Journal of University Education, 17(3), 142–153.
- Salybekova, N., Issayev, G., Abdrassulova, Z., Bostanova, A., Dairabaev, R., & Erdenov, M. (2021). Pupils' research skills development through project-based learning in biology. *Cypriot Journal of Educational Sciences*, *16*(3), 1106–1121.

# The Collaborative Learning of Science Project to Supporting 4C's Skills for Learning in 21<sup>st</sup> Century.

Witsanu SUPANDEE<sup>a\*</sup>& Chitphon YACHULAWETKUNAKORN<sup>b</sup>

<sup>a,b</sup>Engineering Science Classroom, King Mongkut's University of Technology Thonburi, Thailand

\*witsanu.sup@mail.kmutt.ac.th

**Abstract:** The 4C's skills considered as a significant learning skills in 21<sup>st</sup> Century. These skills can be practiced in various learning approach, one of them is Science Project. In this study aimed to present the correlation between collaborative learning in Science Project course and development of 4C's skills in participants. The learning process required students work in group for a semester. To evaluate the effectiveness, participants were assessed 4C's skills including creative thinking, critical thinking, communication, and collaboration as well as integrative knowledge skill after finished Science Project course. Finally, the finding of this study ensured that Science Project could develop all of 4C's skills by themselves. Furthermore, It can be another alternative tools for design learning to develop student soft skill in the future.

Keywords: Science Project, 4C's skill, Collaborative Learning, 21<sup>st</sup> Century Skill

# 1. Introduction

In the latest decade, the educational trends continuously developed and transform in various perspectives toward technological development. In addition, learners could access knowledge and education easily through online databases. Due to its convenience, there are various soft skills that support learning and working called "21<sup>st</sup> Century skills, the skill, abilities that being required for success in 21<sup>st</sup> century workplace. Academic contents and knowledges considered as a core component of this skill, there are 3 main components. For instance, Learning and Innovation skill (4C's), Digital literacy, and Career and Life.

4C's skills considered as the soft skill can support both in academic and workplace. It comprises of Collaboration, Communication, Critical Thinking, and Creativity. The P21 Organization defined that 4C's skills as deeper learning competency (Sudjatmoko, Sofya, & Ritonga, 2021). According to 4C's skill definition, author on behalf of instructor and facilitator who managed the Science Project development in Grade 11<sup>th</sup> students, perceived the correlation between 4C's skills and skill need for Science Project development. Owing to the learning outcome of Science Project course anticipated learners to create innovation by analyze and criticize from real human demand. In addition, they were anticipated to work as teamwork.

4C's Skill's recognized as an important skill need for learning and working in 21<sup>st</sup> century inevitably. Many academic attempted to apply those skills into curriculum. According to the Engineering Science Classroom, the senior high school program which expertise in Science, and Engineering affiliate with King Mongkut's University of Technology Thonburi, Thailand. This Program offered Science Project course for Grade 11<sup>th</sup>-12<sup>th</sup> students in order to learn, explore, do project in laboratory experiments, documentary. From those courses, learners were anticipated to use various soft skills for working in group. Hence 4C's skills, For instance, Collaboration, Communication, Critical Thinking, and Creativity were significance for developing every process in Science Project.

# 2. Related Studies

# 2.1 Science Project

The concept of Science Project initially defined in 1973, defined as the educational activities for laboratory experiment in Chemistry, Biology course. Their definition continuously develops. Generally, Science Project were classified into four type. For instance, Experimental project, Engineering project, Display project, and Theoretical project. In Thailand, Government have attached great importance to science education since 1982. In addition, there are many Sciences Projects competition. For instance, SCiUS Forum, MWIT, Princess Chulabhorn Science school, and others domestic and international Science Project competition.

Science Project have played a central role in school for many decades. It's considered as a latest current course that continuously evolve. In addition, it could engage students in authentic science experience. Clark, and Dickson (2003) opined that Project can encourage sustained reasoning, connect classroom to personal problems, prepare students for lifelong learning. Amelia, Tegariyani, and Santoso (2021) founded that Science project could encouraged student's enthusiasm in various dimension; Hands-on learning, the chance to use scientific method, and scholarship and prize.

The evolution of Science learning have continuously developed, particularly last decade. Its approach based on the Project-based learning, and Problem-based learning. Nevertheless, their learning approach integrated with STEM. Muzana, Jumadi, Wilujeng, Yanto, and Mustamin (2021) opined that there're necessary for improving and developing Science learning approach continuously. According to the evolution of Science Project learning, 21<sup>st</sup> century skills were also used to apply in Science Project learning in Nowadays.

# 2.2 The 4C's Skills

21<sup>st</sup> Century's defined as an essential skill support learning in this century among the innovative change. 4C's categorized as a part of learning and innovative skill. It comprised of Critical thinking, Communication, Collaboration, and Creativity. Those concepts popularized to use in various part of educational sector, Critical thinking was focused, careful analysis of something to better understand it. Second, Communication skill defined as analyzing the situation means thinking about the subject, purpose, sender, receiver, medium. Third, Collaboration skill meant to allocate resources and responsibilities ensures that all members of a team can work optimally. Fourth, Creative thinking was expansive, open-ended invention and discovery of possibilities.

The importance of learning aimed to develop in each skill. Firstly, Critical Thinking's considered as a core skill both in the context of learning. In particular, learning activity. Weng, Cui, Ng, Jong, and Chiu (2022) opined that 4C's are necessary for learning in STEM, Education Maker to develop constructionism in digital education (Weng, Cui, Ng, Jong, & Chiu, 2022). 4C's Skill's considered as a main component of 21<sup>st</sup> Century skill. Putri, Bukit, and Simanjuntak (2021) studied each 4C skill that affect to Project-based learning, founded that the most effective skill that support PjBL is Critical thinking.

PjBL could improve the competency of 4C's skill's in Economics course. In addition to the individual project-based learning, 4C's skill's also optimized learning competency through Team based project. According to the correlation between 4C's skill's and Science Project learning, those skills also optimized the learning competency in various approaches. For instance, Problem-based learning, and Interdisciplinary learning.

# 3. Description of Science Project to Supporting 4C's Skills

# 3.1 Background and Overall Structure

The Science Project is a course for Grade 11<sup>th</sup> student of the Engineering Science Classroom, King Mongkut's University of Technology Thonburi. According to the course, learners have an

opportunity to do research with University's Professors. They used their creativity and observation skill to create the research topic. Moreover, the leaners were required to determine the scope of project, searching for information to do literature review, planning the experiment, recording the results, and analyzing a data and presenting. Learners used various technology to support their work for all process, such as searching for information on the internet (Science Direct, PubMed, ChatGPT, etc.), planning experiments on the Miro board, Padlet platform as well as make a presentation slide via Canva or MS power point. So, this course could be practiced the different learning skills in 21<sup>st</sup> Century. Sometimes, it can be called the 4C's skills. For instance, Creative thinking, Critical thinking, Communication, and Collaboration as well as integrative knowledge skill.



Figure 1. The Overall Structure of Science Project to Supporting 4C's Skills.

The learning process of Science Project course is divided into 4 parts. Part 1 start with group division, 3 person/group and choose the research topic. Part 2 start from research advisor's meeting to design and make a scope of research agreement and do a literature review. Then, each group would write a research proposal, and present to the research committee. Part 3 start with doing the experiment which's followed by their research proposal for collect the data. Finally, Part 4, analyzing the data to conclude. Then the learners have to write research report, and present to research committee which are presented in figure 1.

# 3.2 Design of Learning activity process

In the learning process design, all of learning activity that effected to 4C's skills practice would occur in the classroom and laboratory for 1 semester (approximately 18 weeks). The first activity until the last activity that effected to each skill would be explained in the table 1.

Period	Activity	Process	Practical Skill
Week	Part 1:	Learners created group according to their	Communication,
1-2	Choose	interested topic. After that, selected research	Collaboration
	research topic.	topic obtained from 2 method; University's	skills, and
	-	lecturer, their own interest.	Critical thinking
Week	Part 2:	Learners met the research advisor to talk	Collaboration,
3-8	Design scope of	about the research topic for understand	Communication,
	research and do	clearly, made design research scope's	Critical thinking,
	the literature	agreement, did the literature review for	and
	review.	proposal writing. Finally, the learners would	Creativity skills
		present a proposal to committee and peers.	

Table 1. The Process of Learning Activity Management that Effect to 4C's Skills in Science Project Course.

l able 1. (	Continue)		
Period	Activity	Process	Practical Skill
Week	Part 3:	After completely presented the proposal,	Collaboration,
9-15	Do the	they would start the experiment for collect	Communication,
	experiment.	research data. When they completed	and
		experimental, they would analyze the data to	Critical thinking
-		do research conclusion.	
Week	Part 4:	After the experiment was done, the learners	Collaboration,
16-18	Data analysis	would analyze data to conclusion and write	Communication,
	and	the research report. Then, sent the report to	Critical thinking,
	Presentation.	research advisor to edit some mistake.	and
		Finally, all of learners would present their	Creativity skills
		research results to committee and peers for	
		receive some suggestion or idea for research	
		develop in the future.	

In this study, 4C's skills, for instance creative thinking, critical thinking, communication, and collaboration as well as integrative knowledge would be evaluated during before and after finished the Science Project course. Then, the data would be analyzed by statistic program.

# 4. Research Design

# 4.1 Participant

. .

The participants of this study were all Grade 11<sup>th</sup> students in Engineering Science Classroom, King Mongkut's University of Technology Thonburi. All participants enrolled a Science Project course that consists of 81 participants.

# 4.2 Procedure

Science Project course processes were consisted of 4 part which were presented in figure 2.



Figure 2. The Procedure of Science Project Course

In figure 2 showed all process of Science Project course. First, learners chose research topic (A). Then, learners met and discuss with research advisor in ordered to make research scope, did the literature review in ordered to make a research proposal (B). After that, learners made the presentation for research proposal (C). Then, learners did the experiment following the research proposal for collect data (D). Finally, learners analyzed the research data obtaining from an experiment and concluded it for write research report and make a presentation to committee.

# 4.3 Data Collection & Data Analysis

The technique used for collecting data in this study was student self-assessment. The questionnaires used are self-rating on a 5 scale (1-5). The data were collected during before and after the Science Project course. Then, the results were compared and analyze by statistic

program. The questionnaires were designed into 5 items that include the creative skill, collaboration skill, communication skills, critical thinking skill as well as the ability to integrate knowledge to problem solving that obtained from the Science Project course. For example, of question; before and after you did the Science Project, what is the level score of your creative skills? Additionally, 4C's skill development of learners was assessed by advisor and research committee ratings.

# 5. Result and Discussion

# 5.1 The 4C's Skill Development of Science Project

From the results showed that the students thought Science Project can develop their 4C's skills. Moreover, after the student did the experiment in Science Project, the ability to integrate knowledge to problem solving increased too which score  $4.01\pm0.78$  from 5. Finally, the skills that students thought they have developed the most after do the Science Project are collaboration skills is 35%, creative skill is 34%, communication skill is 20% as well as critical thinking skill is 11% (figure 3).



Figure 3. The 4C's Skills Development in Science Project Course

# 5.2 Self-Assessment Score about 4C's Skills Obtained from Science Project.

After the learners answered the questionnaire completely, the score of each skill before and after do the Science Project are presented in the table 2 and figure 4.

Table 2. The Score of Each 4C's Skills Obtained from Self-Assessment of Student Before and After Done The Science Project.

ltomo	Before	After
liems	(Mean±SD)	(Mean±SD)
1. Creative skill The learners usually use creative skill for solve a problem that occur while their do the research, and creative skill was used to asking question during do the experiment as well as create the presentation through various platforms such as Canva and PPT.	3.33±0.77	4.14±0.65
2. Collaboration skill The learners thought that science project can develop their collaboration skill. Because while they were doing the research, they contacted to many people, such as research advisor and group member for work cooperate include brainstorm via online platforms.	3.57±0.91	4.13±0.77
3. Communication skill The learners developed a collaboration in group member during did the experiment, they had to increase communication skill. The learner used many applications for the communicate, such as Line, Facebook, Discord as well as zoom.	3.26±0.89	4.13±0.61

	Itomo	Before	After
Items	(Mean±SD)	(Mean±SD)	

#### 4. Critical thinking skill

The learners developed their critical thinking skill because when they had a problem during do the experiment, they had to use a critical thinking for analyze the problem and solve it. Sometimes, it also used in data analysis or search for reliable information through various online platform including Google scholar, Science Direct, etc.



*Figure 4.* The Score of Each 4C's Skills Obtained from Self-Assessment of Student Before and After Done The Science Project.

# 6. Conclusion

From the results, author concluded that the learners could develop their 4C's skills, creative skills, collaboration skill, communication skill as well as critical thinking skill from do the Science Project. Moreover, the other significant skill that learners developed from doing the Science Project is an ability to integrate knowledge to solve other problems based on the reasonably. Therefore, learning design for student's 4C's skills development is an important role for learning in 21<sup>st</sup> century. Teachers should let them do some activities that allowed they think and solve problem with their classmate, rather than only focus on lecture-based learning.

# References

- Amelia, R., Tegariyani, S., & Santoso, P. (2021). 21<sup>st</sup> Century Skills in Project Based Learning Integrated STEM on Science Subject: A Systematic Literature Review. Advances in Social Science, Education and Humanities Research, (Vol.529, pp. 583-590).
- Clark, W. C., & Dickson, N. M. (2003). Sustainability science: The emerging research program. *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 100, Issue 14, pp. 8059–8061).
- Muzana, S. R., Jumadi, Wilujeng, I., Yanto, B. E., & Mustamin, A. A. (2021). E-STEM project-based learning in teaching science to increase ICT literacy and problem solving. *International Journal of Evaluation and Research in Education (IJERE), 10*(4), 1386-1394.
- Putri, R. K., Bukit, N., & Simanjuntak, M. P. (2021). The Effect of Project Based Learning Model's on Critical Thinking Skills, Creative Thinking Skills, Collaboration Skills, & Communication Skills (4C) Physics in Senior High School. Advance in Social Science, Education, and Humanities Research, (Vol.591).
- Sudjatmoko, A., Sofya, R., & Ritonga, M. (2021). *Project-Based Learning Model Towards Students'* P21 Skills in the 21st Century: In a Systematic Literature Review.
- Weng, X., Cui, Z., Ng, O. L., Jong, M. S. Y., & Chiu, T. K. F. (2022). Characterizing Students' 4C Skills Development During Problem-based Digital Making. *Journal of Science Education and Technology*, *31*(3), 372–385.

# Implementation of Professional Development Training for Industrial Employees on Artificial Intelligence of Things

Sasithorn CHOOKAEW<sup>a\*</sup>, Suppachai HOWIMANPORN<sup>a</sup>, & Warin SOOTKANEUNG <sup>b</sup>

<sup>a</sup>Department of Teacher Training in Mechanical Engineering, Faculty of Technical Education, King Mongkut's University of Technology North Bangkok, Thailand <sup>b</sup>Faculty of Engineering, Rajamangala University of Technology Phra Nakhon, Bangkok, Thailand \*sasithorn.c@fte.kmutnb.ac.th

Abstract: Artificial Intelligence of Things (AloT) is the most emerging era of a combination of Artificial Intelligence (AI) and the Internet of Things (IoT), which is exponentially gaining researchers' attention with every passing day because of its broad applicability in various sectors. Technology 4.0 in the industry focuses on effective object sorting of conveyors and energy saving. There is an increase in the need for industrial employee training. However, the new technology of conveyor systems to enhance performance has yet to be widespread. However, the new technology of conveyor systems to enhance performance has yet to be widespread. Owing to the learning material for training is expensive. It is also not possible to simulate the operation of a comprehensive system. Industrial employees need to understand the big picture, causing problems in learning and lacking practical skills. Especially new industrial employees may need more knowledge and experience when operating errors; they can have problem-solving skills to seek solutions. Therefore, we proposed the implementation of an object-sorting training kit based on AloT technology. In the experiment, nineteen industrial employees from many industrial enterprises participated in three days of training activities. The results show increased employee performance with an AloT training kit that can simulate object sorting with color and size and then offer a dashboard. They can understand how to save energy in manufacturing. Mainly, they have practical skills and positive perceptions toward this study.

Keywords: AloT, professional development, learning tool, engineering education

# 1. Background and Motivation

Artificial Intelligence of Things (AloT) is the most emerging era of a combination of Artificial Intelligence (AI) and the Internet of Things (IoT), which is exponentially gaining researchers' attention with every passing day because of its broad applicability in various sectors (Ishengoma et al. 2022). Especially in the industrial sector many researchers' effort to present new technologies based on AloT (Matin et al., 2023; Xian et al., 2023; Wang et al., 2022). For example, Nishimura et al. (2021) presented AloT industrial controllers programming of an industrial automation controller, which features capabilities, virtualization, and open-source software integration while creating the possibility to integrate machine learning models into the automation process. Cheng et al. (2021) proposed an AloT module that uses Al algorithms to identify objects, calculate locations for items in automated production, and then send the information to the controller in the automation process. Technology 4.0 in the industry focuses on effective object sorting of conveyors and energy saving. There is an increase in the need for professional training for industrial employees (Chookaew & Howimanporn, 2022). However, the new technology of conveyor systems to enhance performance has yet to be widespread. Owing to the learning material for training is expensive. It is also not possible to simulate the operation of a comprehensive system. Industrial employees need to understand the big picture, causing problems in learning and lacking practical skills. Especially new

industrial employees may need more knowledge and experience when operating errors; they can have problem-solving skills to seek solutions. In addition,

Therefore, we proposed the implementation of professional development training for industrial employees with an object-sorting training kit based on an AloT platform. The significant contributions of the research studies are as follows:

- To develop an AloT platform for proper AloT training in the industrial sector.
- To evaluate industrial employee performance after participating in AloT training.
- To investigate industrial employees' perception toward AloT training.

# 2. Development of Artificial Intelligence of Things platform

# 2.1 AloT platform

Initially, we designed the prototype of the AloT platform for use in training activities based on the architecture of AloT, which is composed of the following four layers:

- Laver 1: The device layer includes various hardware for input and output devices such as load cell, photoelectric sensor, encoder, and DC motor.
- Laver 2: The connectivity layer includes hardware and software that links cloud storage to controllers and sensors such as Omron PLC: NX1P2 that can connect the network via TCT/IP.
- Laver 3: The cloud layer includes data storage and access via Node-Red, an application programming interface (API).
- Laver 4: The user communication layer includes web portals on the monitor and mobile applications.



Figure 1. Architecture of AloT

After, we developed the prototype of the AloT training kit, as shown in Figure 2 (a), which is the part of the training kit, and Figure 2 (b) is the part of the PLC for the controller. In addition, fuzzy inference is used for energy efficiency experiments to make decisions on wind energy saving with MATLAB (Howimanporn et al.,2021), as shown in Figures 3 (a) and 3 (b).

In addition, the data for processing in this system are three inputs: speed, weight, and distance. We use all input data, the conditions obtained from the rules of data in the MATLAB program, to be written into the PLC device, as shown in Figure 4. (a-c).



(b) *Figure 2.* (a) structure of AloT training kit (b) PLC kit



Figure 3. (a) and (b) the screenshots of MATLAB



Figure 4. (a) - (c) The PLC algorithm programming based on input data

# 2.2 Training course

In this study, we have designed the learning activity in three days. These learning activities focus on hands-on learning processes that can support student development of learning outcomes (see Figure 5). We designed the learning outcomes based on four concepts in three days consisting of:

- Day 1: Able to use the computer to connect PLC and to program to control hardware and machine.
- Day 2: Able to display data on the dashboard.
- Day 3: Able to use AI technology for data decision-making to save energy.



*Figure 5.* (a) and (b) the AloT hands-on learning activities for industrial employees.

## 3. Results

#### 3.1 Results of industrial employees' performance after participating in AloT training.

Figure 6 shows the results of industrial employees' performance scores after the training experiment. We found that industrial employees have high performance (HP) of more than 80 percent (42 %), middle performance (MP) between 79-60 percent (42%), and low performance (LP) between 69-40 percent (16%). We found that the industrial employees who are low performance need a background in AI concepts, so that is a barrier for them in application.



*Figure 6.* The performance score of industrial employees.

#### 3.2 Results of industrial employees' perception toward AloT training.

Table 1 shows the results of industrial employees' perception of AloT training, which shows three dimensions of satisfaction at high levels. In addition, Table 2 showed a positive response to activity. Most employees mention that they like participating in training activities through learning opportunities in new technology for their industrial development.

	Table 1.	The industrial	employees'	satisfaction
--	----------	----------------	------------	--------------

Satisfactions	Mean	SD	Level
Training process	3.88	0.89	High
Training activities	3.67	1.12	High
Training usefulness	3.94	1.04	High
	1 0 5 ( 0 50 )		

Note: 1.00-1.50 = very low, 1.51-2.00 = low, 2.51-3.50 = moderate, 3.51-4.50 = high, 4.51-5.00 = very high.

Table 2.	The industrial	employees	response
----------	----------------	-----------	----------

Employees	Responses
No.1	"I wish I had more time to train in new technology. Because I want to have
	the ability to do my work more, I would like to attend the training again if
	there is an opportunity."
No.2	"The training kit is very good and can be used for real experiments. Able to
	apply knowledge to practical use in industrial plants."
No.3	"This training is a new experience that can be applied, and beneficial to work
	admire the speaker and a good team paying attention to operators."
No.4	"The training set is interesting because it is something new in the industry to
	reduce energy consumption and use appropriate energy."
No.5	"The training set is diverse and interested in further application of shadows. It
	is a way to learn more."
No.6	"It is interested in bringing the learning package to develop industrial workers
	that improve the country's industry further is great."

# 4. Conclusion

The overview of the proposed professional development training for industrial employees on AloT is essential to motivate them to improve technology skills it has achieved implementation. The training course proposed an object-sorting training kit based on AloT technology to promote industrial employees' understanding and practical skills. After training, their improved performance, and positive responses. Thus, this course can be deployed to guideline professional training in the industrial sector. Nevertheless, the limitation of this study is that we explored the industrial employees' performance levels of AloT and their perception of 19 employees in a few industrial enterprises. This finding may not be generalized to other contexts. Thus, future studies should aim to expand the sample size and include employees from various industrial enterprises.

# Acknowledgements

This research was funded by King Mongkut's University of Technology North Bangkok. Contract no. KMUTNB-67-BASIC-25.

# References

- Cheng, E. S., Yang, J. Y., Lee, J. D., Chen, K. Y., Hu, N. Z., & Chen, L. Y. (2021, September). AloT module development for automated production. In 2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW) (pp. 1-2). IEEE.
- Chookaew, S., & Howimanporn, S. (2022). Upskilling and reskilling for engineering workforce: Implementing an automated manufacturing 4.0 technology training course. Global Journal of Engineering Education, 24(1), 34-39.
- Ishengoma, F. R., Shao, D., Alexopoulos, C., Saxena, S., & Nikiforova, A. (2022). Integration of artificial intelligence of things (AloT) in the public sector: Drivers, barriers and future research agenda. Digital Policy, Regulation and Governance, 24(5), 449-462.
- Nishimura, E. H., Iano, Y., de Oliveira, G. G., & Vaz, G. C. (2021, October). Application and requirements of AIoT-enabled industrial control units. In Brazilian Technology Symposium (pp. 724-733). Cham: Springer International Publishing.
- Matin, A., Islam, M. R., Wang, X., Huo, H., & Xu, G. (2023). AloT for sustainable manufacturing: Overview, challenges, and opportunities. Internet of Things, 100901.
- Howimanporn, S., Chookaew, S., & Silawatchananai, C. (2021). Real-Time evaluation position control of directional wheel conveyor using fuzzy embedded PLC and SCADA. International Journal of Mechanical Engineering and Robotics Research, 10(6), 328-336.
- Wang, Y., Tan, Y. A., Baker, T., Kumar, N., & Zhang, Q. (2022). Deep fusion: crafting transferable adversarial examples and improving robustness of industrial artificial intelligence of things. IEEE Transactions on Industrial Informatics.
- Xian, W., Yu, K., Han, F., Fang, L., He, D., & Han, Q. L. (2023). Advanced Manufacturing in Industry 5.0: A Survey of Key Enabling Technologies and Future Trends. IEEE Transactions on Industrial Informatics.

# STEM Education in Sustainability: A Brief Literature Review (2019 - 2023)

Pawat CHAIPIDECH<sup>a\*</sup>, Sasivimol PREMTHAISONG<sup>b</sup>, & Niwat SRISAWASDI<sup>a</sup>

<sup>a</sup>Division of Science, Mathematics, and Technology Education, Faculty of Education, Khon Kaen University, Thailand

<sup>b</sup>Khon Kaen University Demonstation School International Division, Khon Kaen University, Thailand

\*pawach@kku.ac.th

**Abstract:** During the half past of a decade, STEM education has been adapted worldwide as an effective way to promote leaners and teachers' application of knowledge regarding the interdisciplinary on their own filed. Meanwhile, the learning about sustainability concept has been increasing focus within educational activities in schools. This study aims to provide an overview of implementation of STEM education with sustainability concept in school level during 2019 to 2023. Based on a systematic process to retrieved from SCOPUS database, 44 research articles were analyzed. The coding scheme to examine the papers including learning strategies, research issues, and research method. The findings of this study could shed the light on the provision for educators and future research on STEM education in sustainability studies.

**Keywords:** Research trend, STEM, sustainable development, school, systematic review

# 1. Introduction

In the 21<sup>st</sup> century, STEM education has been recognized by teacher and educator around the world. For education, it became the central focus of many countries' national curriculums in order to fulfill the goal of building and preserving success through the development of their human capital into productive workers. Meanwhile, there are diverse perspectives about the disciplinarity of STEM and STEM education (Erduran, 2020; Tytler, 2020). Despite being at its early stages, the area of STEM education has shown significant expansion in the last decade. Focusing on the development of STEM education, it has been supported by various factors. For example, Li et al. (2022) report that research funding and the growth of research publications are the keys to develop STEM education research. As STEM education evolves and gains significance across the world, its integration with sustainable practices becomes essential. Incorporating sustainability into STEM education not only corresponds with worldwide educational reforms, but it also educates the next generation to tackle and enhance critical environmental concerns.

There is an increasing focus on addressing sustainability concerns within educational activities in schools, with the aim of promoting Education for Sustainable Development (ESD) (Leicht et al., 2018). This recognition has led to calls for educational reforms that consideration of the challenges in science education, particularly in terms of preparing students for the future society (Schulz, 2009; Van Laar et al., 2017). The incorporation of STEM Education with sustainable development has become a significant educational framework within the context of science teaching in schools (Pitt, 2009). As a result, educators can put more of their focus on helping students develop a broad understanding of STEM so that they can effectively contribute as well scientific knowledge citizens in the society (Christenson, Chang Rundgren, & Zeidler, 2014). There are a variety of article categories among high-impact research publications, including empirical studies, research review, and theoretical discussion papers. Review articles and conceptual/theoretical discussion papers are viewed as an assistance

and valuable because they synthesis of the existing studies on a particular topic and/or provide new perspectives and directions.

Typically, Review articles provide the purpose of offering a comprehensive summary of the existing research within a specific field or on a particular topic. They play a crucial role in enabling readers to obtain a broad understanding of a subject matter, including its important aspects and relevant publications (Palmatier, Houston, & Hulland, 2018). Therefore, this study aims to gain insight about research development in STEM education in sustainability, we thus decided to focus on empirical research regarding to the field. Moreover, systematic examining and reviewing those studies were utilized which provide a better comprehension to the emerging trend in terms of research issues, methods, and learning strategies.

# 2. Research Methodology

#### 2.1 Data Collection

This research examined papers from the SCOPUS database from 2019 to 2023 on 12 July 2023. The searching was for the publications whose article title, abstract, or keywords met the logical condition. There were 146 papers were found by using ("STEM" or "STEAM" or "science, technology, engineering, and mathematics" or "Science Learning") AND ("Sustainability" or "circular economy" or "green economy" or "bio economy") AND ("school" or "k-12") as the searching condition. By limiting the search to social science and article and journal and english, there were 77 papers. By remove papers which were not research article and not school level, a total of 44 papers met the condition (see Figure 1).



Figure 1. Data Collection Procedure.

After gathering data, two experienced researchers then read and divided the papers based on the coding scheme, if any incompatibility between the coding process was found, the researchers need to discuss until agreement was reached.

#### 2.2 Data Distribution

Figure 2 demonstrates the situation of STEM education in the field of sustainability context papers from 2019 to 2023. It was found that the research issues have paid more attention in this field since 2019. There were 2, 12, 10, 10, and 10 papers published in each year, indicating that the researchers started to focus on this research issue increasingly. Moreover, since 2021 to 2023, it indicated that there is remaining focus to this research context.



Figure 2. STEM Education in Sustainability Published in 2019 to 2023

#### 2.3 Coding Scheme

This study used a coding scheme to analyze information from the published papers which obtained from SCOPUS database, including basic information from the paper (i.e., nationalities, journals, and authors), learning strategies, research methodology, and research issue. The following items describe the coding schemes:

- Nationalities, authors, and journals: The standard information of those published papers is discussed, including nationalities, and journals. The purposes to comprehend which countries have more frequently published papers about STEM Education in the field of sustainability.
- Learning strategies: This study classified the learning strategies into 16 categories that is, learning program, training workshop, case study, project-based learning, citizen science, place-based learning, combination of approaches, exhibition, experiential learning, game interaction, guided learning, inquiry-based engineering learning, online learning, science camp, and thinking-based STEM learning.
- Research methods: The classified research methodology based on the 3 common research design presented by (Creswell, 2014), including mixed method, quantitative methods, qualitative method, and others (i.e., system development).
- Research issue: The research issues were investigated among the 44 papers related to STEM education in sustainability, including cognitive domain, affective domain, and psychomotor domain.

# 3. Results

#### 3.1 Journals

This study examined the distribution of the published journals, as presented in Figure 4. It was found that many studies were published in Sustainability (65%), followed by the International Journal of Technology and Design Education (18%), Frontiers in Education (12%), and the others various journal (4%). The sustainability is an open access which welcomes research papers that discuss the challenge related to sustainability. Apart from Sustainability, the other journals also focus on the application of STEM Education in learning environments, and investigate meaningful educational settings.



Figure 4. Distribution of Published Journals

# 3.2 Learning Strategies

The learning strategies among the 44 papers which applied in STEM education from 2019 to 2023 were analyzed as shown in Figure 6. The results illustrated that there are 12 papers not mentioned about learning strategy. However, the second greatest is learning program (n = 7), followed by training workshop (n = 5). It can be found that the most explicitly learning strategy which the research designed to emphasize participant's learning on sustainability within STEM education context are learning program and training workshop. For example, Schultz and Delaney (2021) developed chemistry outreach activities related to the periodic table and sustainability in order to increase learner's perception toward chemistry with circular economy.



Figure 6. Distribution of Learning Strategy Mentioned Among the Published Papers

#### 3.3 Research Issues

Among the selected papers, the research issues in terms of cognitive, affect, and psychomotor domain of each publication were analyzed as shown in Figure 7 and Figure 8. According to the participants who were recruited in the 44 selected papers are students and teachers, this study divided into 2 group of research issues based on the participants. Figure 7 was presented the result of frequency of the issue for student as the participant. It revealed that the most popular issue regarding affective domain including, interest (n = 4) and attitude (n = 4), followed by perception (n = 3), motivation (n = 3), and environmental awareness (n = 3) as shown in Figure 7 (A). Moreover, Figure 7 (B) represented the research issues on cognitive domain including science knowledge (n = 5), learning performance (n = 2), and the understanding of nature of science (n = 1), respectively. Meanwhile in psychomotor domain, the result illustrated that learning behavior (n = 3) was the greatest number of research issues in this domain, following by creativity (n = 2).



Figure 7. Distribution of Student's Research Issues within STEM Education in Sustainability

In addition, the analysis results shown that the most frequency research issue on teacher in terms of STEM education in sustainability is attitude (n = 4). Moreover, figure 8 illustrated that the following issues including Interest (n = 3), Perspective (n = 3), Motivation (n = 2), teaching confidence (n = 2), and teaching experience (n = 2).



Figure 8. Distribution of Teacher's Research Issues within STEM Education in Sustainability

# 3.4 Research Methodology

Figure 9 shows the most often utilized research methodologies among the 44 publications selected. In STEM education in sustainability, qualitative research (n=20, 48%) was the most commonly employed research, with interviews or discussions used to obtain data for analysis. A quarter of the research (n=11, 26%) employed mixed methods, which combined quantitative and qualitative methodologies. Similarly, the quantitative method (n=10, 24%) obtained data mostly through questionnaires. A relatively small number of the studies (2%) studied the design and development of the system on STEM education in sustainability.



Figure 9. Distribution of Research Methodology

#### 4. Conclusion and Discussion

This study analyzed the publications of STEM education in sustainability in the SCOPUS database to determine the most important or influential keywords, authors, countries, hot topics, learning strategies, and research methodology. The results showed that the number of papers was constantly increasing since 2019, indicating that more research on the field was published with the focusing of sustainability. In addition, the findings also revealed that the most published journal among the samples was Sustainability (65%). This might be due to the aims and scopes of the journal related to the aims or subject area in the articles. In addition, learning strategies' result shown that learning program is the most strategies to conduct research among the published articles. This suggested that the researchers concentrated on learning programs when STEM education in sustainability concept was implemented, using qualitative method (48%) which illustrated the shifting of research method reported by Li et al. (2022), that quantitative methods were used as the dominant approach from 2005 to 2019 among the high impact empirical studies in STEM education. In terms of research issues, it was categorized into two types based on the participant among the publications. The research issues on students revealed that science knowledge, interest, and learning behavior were the most study issue regarding to cognitive, affective, and psychomotor domain respectively. For

teachers, the most focusing issues was attitude, interest, perspective, motivation, teaching confidence, and teaching experience. This implied that investigating and enhancing teacher's teaching competency in STEM education in sustainability remain a challenging and important issue. Nevertheless, it is crucial to acknowledge several limitations when interpreting the aforementioned conclusions. The search terms employed, the journals chosen, and the publication dates utilized in this study may have unintentionally ignored other relevant literature and may not fully capture the entirety of research trends.

#### Acknowledgements

This study is funded by National Research Council of Thailand (NRCT) under contract number N42A661011. This study also supported by Faculty of Education, Khon Kaen University, Thailand.

#### References

- Christenson, N., Chang Rundgren, S.-N., & Zeidler, D. L. (2014). The relationship of discipline background to upper secondary students' argumentation on socioscientific issues. *Research in Science Education*, *44*(4), pp. 581–601.
- Creswell, J. W. (2014). *Research design: qualitative, quantitative, and mixed methods approaches.* 4th ed. Thousand Oaks, California, SAGE Publications.
- Erduran, S. (2020). Nature of "STEM"? Science & Education, 29, 781–784. https:// doi.org/10.1007/s11191-020-00150-6
- Leicht, A., Combes, B., Byun, W. J., & Agbedahin, A. V. (2018). From agenda 21 to target 4.7: the development of education for sustainable development. In: A. Leicht, Julia, H, V. A Adesuwa and J. B. E Won. (ed.) *Issues and Trends in Education for Sustainable Development*. Paris: UNESCO.
- Li, Y., Xiao, Y., Wang, K., Zhang, N., Pang, Y., Wang, R., Qi, C., Yuan, Z., Xu, J., Nite, S. B., & Star, J. R. (2022). A systematic review of high impact empirical studies in STEM education. *International Journal of STEM Education*, *9*, 72. https://doi.org/10.1186/s40594-022-00389-1
- Palmatier, R. W., Houston, M. B., & Hulland, J. (2018). Review articles: purpose, process, and structure. *Journal of the Academy of Marketing Science, 46*, pp. 1-5.
- Pitt, J. (2009). Blurring the boundaries—STEM education and education for sustainable development. *International Journal of Technology and Design Education*, *14*, pp. 37–48.
- Schulz, R. M. (2009). Reforming science education: Part I. The search for a philosophy of science education. *Science & Education*, 18, 225–249.
- Schultz, M. & Delaney, S. (2021). Development, use, and evaluation of chemistry outreach activities related to the periodic table and sustainability. *Journal of Chemical Education*, *98*(12), pp. 3921–3929.
- Tytler, R. (2020). STEM education for the twenty-first century. In J. Anderson & Y. Li (Eds.), Integrated approaches to STEM education: An international perspec- tive (pp. 21–43). Springer.
- van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, 72, 577–588. https://doi.org/10.1016/j. chb.2017.03.010

# A Proposal for Mobile-assisted Citizen Inquiry Learning Approach in Learning of Plastic Pollution

#### Chawadol SRIBOONPIMSUAY<sup>a\*</sup>, Niwat SRISAWASDI<sup>a,b,c\*</sup>

<sup>a</sup>Faculty of Education, Khon Kaen University, Thailand <sup>b</sup>Smart Learning Innovation Research Center, Khon Kaen University, Thailand <sup>c</sup>Digital Education and Learning Engineering Association of Thailand, Thailand \*niwsri@kku.ac.th

**Abstract:** Engaging young people in citizen inquiry could facilitate the process of constructing learning outcomes through investigation and reflection on their experiences. This study aimed to investigate the effects of using the citizen inquiry approach on students' learning outcomes regarding environmental awareness, scientific explanation, and scientific communication. An intensive science camp regarding the citizen inquiry approach to plastic pollution has been developed and implemented for elementary school students. The results showed considerable improvement in their learning outcomes after interacting with the intervention. However, there is a challenge and opportunity to redesign the implemented approach with the support of mobile technology to improve their learning outcomes. This paper proposes how we could transform the citizen inquiry approach with digital devices to promote elementary school students' environmental awareness, scientific explanation, and scientific communication.

**Keywords:** Citizen science, inquiry-based learning, environmental awareness, scientific communication, scientific explanation

# 1. Introduction

Currently, there is a growing global trend in citizen science initiatives to reduce the amount of plastic pollution in the environment. Citizen science is an approach that involves the public in scientific research related to scientists' interests (Silvertown, 2009). This approach allows people to participate in scientific projects and collect and analyze data, regardless of their educational background (Aristeidou et al., 2013). By engaging the public in citizen science, individuals can learn about science in a real-life context and bridge the gap between scientific communities and the general public (Bonney et al., 2014; Herodotou et al., 2014). For example, Ballard et al. (2017) studied and analyzed two youth-focused programs, called Environmental Science Agency. It aimed to link conservation learning and action. They found the program enhanced various aspects of environmental science agency. Despite the growth of youth participation in citizen science, studies on achieving rigorous environmental education and conservation science are scarce.

Addressing plastic pollution through citizen science is effective for youth education. Despite the integration of environmental issues into school curricula increasing (Rickinson & Lundholm, 2008), current methods still have limitations in fostering youth environmental awareness (Dalu et al., 2020; Robertson & Krugly-Smolska, 1997). To promote the application of citizen science in the school education context, Herodotou et al. (2018) proposed citizen inquiry, a transformative approach combining inquiry-based learning and citizen science. It aims to involve citizens in all scientific inquiry stages including the conception of a project, the definition of research objectives, the selection of methods for data collection and analysis, and the implementation of research. Moreover, Aristeidou et al. (2020) used online citizen inquiry

to scaffold students in scientific processes during citizen science activities, reporting meaningful involvement and practice of inquiry skills.

In the process of citizen inquiry, mobile or smartphone are beneficial leaning tools. It can provide learners with access to a wide range of resources, facilitating data collection and analysis, and promoting collaboration and engagement (Sharples et al., 2015; Suárez et al., 2018). Thus, there is a need to develop better version of citizen inquiry approach to promote their environmental awareness and motivation among individuals (Henderson & Green, 2020; Williams et al., 2019). Furthermore, scientific competencies, such as scientific explanation and communication, are also crucial outcome for learning science by citizen inquiry.

In conclusion, promoting young people to learn science through citizen inquiry addressing environmental issue, such as plastic pollution, requires not only scientific inquiry skills, but also environmental awareness, scientific explanation, and effective scientific communication. This study aims to investigate an effect of traditional citizen inquiry learning integrated open inquiry approach on students' environmental awareness, scientific explanation, and scientific communication, and introduce a proposal for a mobile-assisted citizen inquiry learning approach in learning of plastic pollution for elementary school students.

#### 2. Review of Literature

#### 2.1 Citizen Science and Citizen Inquiry

Citizen science, a growing field involving the public in scientific projects, benefits both the public and scientific community through quality data collection and increased scientific literacy (Bonney et al., 2009). However, its traditional model has been criticized for limiting research questions and participant diversity (Haklay, 2013; Jordan et al., 2011). The citizen inquiry approach addresses these limitations by encouraging active knowledge construction (Dewey, 1916; Hmelo-Silver et al., 2007), empowering citizens to make informed decisions (Sharples et al., 2013) and merging elements of citizen science and inquiry-based learning (Aristeidou et al., 2013).

In 2014, citizen inquiry was defined as a research collaboration involving scientists and the public (Herodotou et al., 2014). it combines principles of inquiry-based learning and citizen science for relevant, authentic investigations (Aristeidou et al., 2017). It also bridges the gap between citizen science and inquiry-based learning in informal education (Aristeidou et al., 2017).

#### 2.2 Citizen Inquiry of Plastic Pollution

Citizen inquiry into plastic pollution has grown since the 1990s, with initiatives like the International Coastal Cleanup (ICC) launched in 1986, involving volunteers in beach cleanups and data collection on plastic debris, informing policy decisions (*International Coastal Cleanup*, 2017) Recent projects include the Plastic Citizen project, where young people collected data on plastic pollution (Burden et al., 2021) and the "Plastic Pirates" program, allowing students to collect and share data about riverbank trash (Gudrun, 2022).

These initiatives show citizen inquiry's effectiveness in addressing environmental concerns and its role in scientific processes. Thus, citizen inquiry has been instrumental in tackling plastic pollution, leading to policy and behavioral changes.

#### 2.3 Inquiry-Based Learning and Guided-inquiry Process

Scholars have indicated inquiry-based learning strategies can support students in constructing and developing content knowledge and enhancing comprehension of the scientific investigation (Srisawasdi & Panjaburee, 2019). Inquiry-based learning involves students actively participating in scientific investigations, exploring concepts, and fostering inquiry skills (National Research Council, 2000). On the other hand, guided inquiry learning offers students guidance throughout the process. It includes a series of steps that steer students through the inquiry process which would help them gain necessary knowledge and skills (Bell et al., 2009). It can be seen as a part of inquiry-based learning with more teacher involvement.

Guided inquiry engages students in investigations and simulations related to environmental issues, helping them understand the impact of human activities on the environment and explore the causes, effects, and solutions to these problems (Sholahuddin et al., 2020). It also provides opportunities for students to ask questions, design investigations, analyze data, and construct scientific explanations based on evidence (Stone, 2014). Furthermore, it enhances scientific communication skills through activities such as presenting findings, writing lab reports, and participating in scientific discussions, enabling students to communicate scientific information clearly and effectively to different audiences (Sarwi et al., 2018).

#### 2.4 Mobile-Assisted Learning in Citizen Inquiry Approach

Mobile-assisted learning is using mobile devices to support learning. While, citizen inquiry is a way of engaging citizens in scientific research. Those can be combined for effective learning experiences. Mobile devices provide easy access to information, real-time data recording, and facilitate collaboration among citizen scientists. They also offer interactive learning through educational apps and gamified activities (Herodotou et al., 2019; Kukulska-Hulme, 2011).

Research indicates that mobile-assisted learning enhances engagement, motivation, and understanding in citizen inquiry. Kukulska-Hulme (2011) found that mobile devices improved engagement, motivation, and understanding of scientific concepts related to air pollution. Similarly, Herodotou et al. (2019) found that mobile-assisted learning facilitated data collection, analysis, and collaboration among participants.

# 3. Methodology

#### 3.1 Participants

The study was conducted as an intensive science camp with 29 students aged 10-12 years old from Demonstration School of Khon Kaen University, international division Thailand. The camp took place both within and outside the school premises, all within the university area.

# 3.2 The Intensive Science Camp

A -day science camp was designed in the plastic pollution theme to help students recognize the effect of plastic on the surrounding environment through the plastic problem and experiment. All activities were created using citizen inquiry and guided inquiry. All activities in the camp were conducted mainly in English with some Thai in the section where students seemed struggle to understand.

The camp was divided into three stages: Ignite, Explore, and Evaluate, based on the work of Srisawasdi & Kroothkeaw, (2014). How the experimental procedure of the intensive science camp was conducted was shown in Figure 1.

In the ignite stage, it consists of two sub-processes and there was a total of 25 minutes. The first sub-process is an inquiry question. In the process, students were presented with an question entailing their investigative inquiry in a real context as citizen (15 minutes). They will use the inquiry question as a main compass to explore the rest of the learning missions in the camp. In the second sub-process, a series of scientific background and information related to the question has been provided to students for engaging their existing ideas and information about plastic pollution and microplastic (10 minutes).

In the explore stage, this stage also contains two sub-process and there was a total of 195 minutes. The first sub-process is investigative procedures, and they were described all the learning missions for the camp (20 minutes).



Figure 1. Experimental Procedure of The Intensive Science Camp

For the investigative procedures, there were three citizen inquiry-related practices consisting of plastic waste investigation (65 minutes), microplastic lab investigation (30 minutes), and plastic circular economy investigation (30 minutes). Figure 2 illustrates the three learning missions in the camp intervention.



*Figure 2.* Illustrative examples of learning mission 1 (left), learning mission 2 (middle), and learning mission 3 (right)

In Figure 2, the three inquiry learning missions could be described as follows.

Mission 1 - What can you tell me about these plastics? This mission was conducted using a field-based collaborative inquiry approach using quadrat sampling methods. Students will count the amount of plastic, living things and vegetation in the designated 4-square meters plotted area. It was adapted from Ghaffari et al., (2019).

Mission 2 - Can you trust what you see? Students will try to find microplastics in the given water sample. In this mission was adapted from "Microplastics in the Classroom" mission of Blue Ocean Society, (2022). Each group explored the microplastics in the given water sample with portable microscope kit which was attached to their mobile devices.

Mission 3 - The 4 Rs: Students will –explore the feasible ways of reducing plastic pollution by using 4Rs from the 9R Framework of Kirchherr et al., (2017), which are – Reduce, Reuse, Recycle and Repurpose. Students worked in groups to study stations set up by researchers and were guided by a facilitator.

The second sub-process of the explore stage is data analysis. students gather and analyze data from all missions Then, combining quantitative and qualitative data to form a conclusive explanation.

In the evaluation stage, it consists of two sub-processes with 55 minutes. The first subprocess, result communication (45 minutes), is where students were asked to explain and communicate their understanding. Then, they exchanged ideas on plastic pollution solutions with peers and teachers.

In the last sub-process of the evaluation stage (10 minutes). This is when researchers and students try to conclude the whole key concept of the camp and answer question together. This would also depict the whole system of plastic pollution from plastic production to soil settlement and water contamination which, eventually, come back to harm humans.

# 3.3 Data Collection

Students' environmental awareness was assessed before and after the camp through videorecorded responses to questions about plastic pollution. It then were scored using framework of Lederman et al. (2014). Scientific explanation was assessed through three worksheets designed to assess their scientific explanation according to the framework of Mcneill (2008). Scientific communication was evaluated through a final video-recorded response. Then scored based on work of Lederman et al. (2014). The transcriptions were also analyzed to depict the level of circularity in the chosen Rs of students and to interpret the effect of the treated intervention on their understanding. Scores were given based on whether they stated, mentioned, or offered information that was previously presented in the intervention.

#### 3.4 Analysis of data

#### 3.4.1 Environmental Awareness

Students individually answered three questions related to Missions 1-3 via video. Transcriptions were evaluated by two experts for environmental awareness at three levels - naïve, mixed, and informed. Any disagreements were resolved through discussion until consensus was reached. Then, scores before and after each mission were compared and depicted in a line graph to compare changes in students' environmental awareness due to the intervention.

#### 3.4.2 Scientific Explanation

The worksheets from mission 1-3 were given to students. After the camp is finished, each worksheet was examined for the level of each component – claim, evidence, and reasoning – using the criteria of McNeil (2008). Then it went through the process of grouping. If the students had a similar score in each component, they fell into same group e.g., if student has level 0 cross claim, evidence, and reasoning, they will fall into group 000 – where each number describes the level of each element respectively. After that, the frequency of the students in each group was counted to determine the abundance of the students in each group.

#### 3.4.3 Scientific Communication

All students were presented with a final question. They answered by recording video. Then the video transcriptions were scored within three levels – naïve, mixed, and informed. The frequency was counted and calculated as percentage. Also, the level of circularity was analyzed based on the work of Kirchherr et al., (2017). The scores -2, -1, +1, and +2 were assigned to recycle, repurpose, reuse, and reduce respectively. If the multiple Rs were chosen, it would be assigned to the highest score. This was assigned to the X axis.

The comprehension in the camp intervention was evaluated and given score -1 and 1 were given to each criterion where 1 was given when the answer was true and vice versa. The sum calculation was made to determine the total number of scores. This will be the Y-axis. The data from the individual was plotted to create bubble graph.

# 4. Result

#### 4.1 Environmental Awareness

To depict the comparison of change of the score, the line graph was created in Figure 3. It is important to note the data collection occurred at the end of the camp day. Some students may have been absent due to early pick-up, which should be considered when interpreting the results.



*Figure 3.* Result of Environmental Awareness for Individual Student. Comparing Between Before and After the Camp of 3 Missions

In mission 1, most students showed noticeable improvement from 'Naïve' to 'Mixed' or 'Mixed' to 'Informed' levels after the camp (M1A), compared to before (M1B). However, some students' scores remained unchanged. Mission 2 results were similar, with some students maintaining the same 'Naïve' level in both pre- and post-mission assessments (M2B and M2A). Mission 3 (M3B and M3A) showed similar trends, but interestingly, some students regressed from 'Mixed' to 'Naïve'. The rest of the details was depicted in figure 3.

# 4.2 Scientific Explanation

Student scores were categorized into groups based on the three components of scientific explanation - claim, evidence, and reasoning. This led to further grouping into 7 categories as shown in Table 1.

Group	Description	Score
I	Score from all each component of claim, evidence, and reasoning were	000, 111,
	found to be correlated	222
	Students could form claims and conduct experiments for evidence, but	121
	struggled to create the link between claim and evidence	
	Students could create somewhat level of claims and conducted the	110, 120
	experiment to extract the evidence on somewhat or strong level but	
	unable to create the link between claim and evidence.	
IV	Students could make only claim.	100
V	Students could only conduct the experiment to bring out the evidence	010
	but unable to create both claim and reasoning.	
VI	Students trended to use their prior knowledge to create the evidence	101
	then link it with claim to create reasoning	
VII	Students trended to make a guess when making the reasoning and	021
	ignored the stated claim.	

 Table 1 Group Classification of Students' Level of Explanation

Figure 4 presents the number of students in each group from Missions 1, 2, and 3. Group I had 3, 1, and 6 students for each mission respectively, but most students were in

Group III. This suggests that the intervention fostered some scientific reasoning elements, but students may have struggled to link claim and evidence to form valid scientific reasoning.



The second largest - Group III- suggests that students could make claims and draw evidence from Missions 1 and 2 but struggled with reasoning. Also, group VI is notable as students seemed to use prior knowledge for reasoning, despite being unable to extract evidence from the experiment. While in group VII, students appeared to attempt reasoning despite a lack of claim.

#### 4.3 Scientific Communication

Figure 5 shows the number of students at each level. The majority, 50% or 13 students, are at the 'naïve' level, while only 11.54% or 3 students reached the 'informed' level.



Figure 5. Number of Students from Each Level of Awareness

The correlation between the selected circularity level and their comprehension in the intervention is depicted in Figure 6. Most students fall into quadrant 2 of the graph, indicating they chose low-level circularity to tackle plastic pollution, but their statements were based on intervention information. The second largest group is in quadrant 1, indicating that they chose higher-level circularity, and their statements were based on the intervention. Few students stated information not from the intervention.



Figure 6. Result of Students' Circularity Level and Understanding

# 5. Discussion

The study conducted a camp with three missions to assess students' environmental awareness, scientific explanation skills, and communication abilities. The results provide valuable insights into the effectiveness of the intervention and areas for potential improvement. The camp appears to have had a positive impact on students' environmental awareness. However, some students showed no change or even regression, possibly due to a lack of exposure to environmental concepts (Hungerford & Volk, 2013). In terms of scientific explanation, most students understood parts of it but struggled to connect claims and evidence to create reasoning. This suggests that there is room for improvement in helping students form valid scientific reasoning. This could be due to insufficient topic knowledge or difficulty transferring skills to new contexts (Choowong & Worapun, 2021; Seifert et al., 2022). A lack of interest or motivation in science could also hinder the development of evidence gathering and reasoning skills (Anderhag et al., 2016).

In scientific communication, few students achieved the highest level, indicating a need for more targeted practice. Also, varying levels of prior scientific knowledge or communication abilities among students could be a contributing factor (Weigold, 2001). Most students chose low-level circularity but based their statements on intervention information, suggesting they grasped some camp concepts. However, few students stated information not from the intervention, indicating a potential gap in their understanding. Future interventions should consider strategies for differentiated instruction to foster each skill of students.

# 6. A Proposal for Activity Incorporated Within Technology Learning.

Although there was progress in increasing environmental awareness, explaining science, and communicating science of students, there's still space to do better in these areas. To do that, this section puts forth an idea to improve this intervention by making better use of technology. By bringing in more advanced tools during these camps, we could make the outcomes even better. To do that, the incorporation of technology and mobile learning in creative ways during the intervention should be made. For instance, instead of just reading from paper, interactive videos could be used to teach and provide information to students. Also, rather than writing on paper, students could use mobile apps to complete their assignments, which would give them quick feedback.





Moreover, when doing an interview with students, it could be done online as a group instead of recording each person separately. The propose is that by using these technology tools, it can make learning more exciting and effective for students attending these camps, making them more interested in the environment and better at communicating scientific ideas. In the end, hopefully, this could potentially lead to creating young children to be more resilient and become greater global citizens. Figure 7 displays a conceptual idea of mobile-assisted citizen inquiry approach for the learning of plastic pollution.

#### 7. Conclusion

The study found that an intervention improved students' environmental awareness and scientific reasoning, but some students responded less, and challenges remained in linking claims with evidence. Despite using guided inquiry and citizen inquiry, scientific communication skills varied among students. The intervention's success in fostering environmental awareness and scientific reasoning was noticeable, but enhancing scientific communication skills needs more comprehensive strategies.

While there were improvements, there's still room to do better. The study suggests using advanced technology tools to make learning more engaging and effective. For example, using interactive videos for teaching, mobile apps for assignments, and online group interviews. These tools could make learning more exciting, increase interest in the environment, and improve communication of scientific ideas. This could lead to the development of resilient young global citizens. A concept of a mobile-assisted citizen inquiry learning approach for learning about plastic pollution is also presented.

# References

- Anderhag, P., Wickman, P.-O., Bergqvist, K., Jakobson, B., Hamza, K. M., & Säljö, R. (2016). Why Do Secondary School Students Lose Their Interest in Science? Or Does it Never Emerge? A Possible and Overlooked Explanation. *Science Education*, *100*(5), 791–813.
- Aristeidou, M., Scanlon, E., & Sharples, M. (2013). *A design-based study of Citizen Inquiry for geology*. 7.
- Aristeidou, M., Scanlon, E., & Sharples, M. (2017). Design processes of a citizen inquiry community. In C. Herodotou, M. Sharples, & E. Scanlon (Eds.), *Citizen Inquiry* (1st ed., pp. 210–229). Routledge.
- Aristeidou, M., Scanlon, E., & Sharples, M. (2020). Learning outcomes in online citizen science communities designed for inquiry. *International Journal of Science Education, Part B*, 10(4), 277– 294.
- Ballard, H. L., Dixon, C. G. H., & Harris, E. M. (2017). Youth-focused citizen science: Examining the role of environmental science learning and agency for conservation. *Biological Conservation*, *208*, 65–75.
- Blue Ocean Society. (2022, February 22). *Microplastics in the Classroom*.
- Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience*, 59(11), 977–984.
- Bonney, R., Shirk, J. L., Phillips, T. B., Wiggins, A., Ballard, H. L., Miller-Rushing, A. J., & Parrish, J. K. (2014). Next Steps for Citizen Science. *Science*, *343*(6178), 1436–1437.
- Burden, K., Wurzel, R., & Dean, C. (2021). *Plastic Citizen*. https://www.hull.ac.uk/work-with-us/research/case-studies/plastic-citizen.aspx
- Choowong, K., & Worapun, W. (2021). The Development of Scientific Reasoning Ability on Concept of Light and Image of Grade 9 Students by Using Inquiry-Based Learning 5E with Prediction Observation and Explanation Strategy. *Journal of Education and Learning*, *10*(5), Article 5.
- Dalu, M. T. B., Cuthbert, R. N., Muhali, H., Chari, L. D., Manyani, A., Masunungure, C., & Dalu, T. (2020). Is Awareness on Plastic Pollution Being Raised in Schools? Understanding Perceptions of Primary and Secondary School Educators. *Sustainability*, *12*(17).
- Dewey, J. (1916). *Democracy and Education: An Introduction to the Philosophy of Education*. Macmillan.
- Ghaffari, S., Bakhtiari, A. R., Ghasempouri, S. M., & Nasrolahi, A. (2019). The influence of human activity and morphological characteristics of beaches on plastic debris distribution along the Caspian Sea as a closed water body. *Environmental Science and Pollution Research*, *26*(25), 25712–25724.
- Gudrun, O. (2022). About the campaign | Plastic Pirates. https://www.plastic-pirates.eu/en/about
- Haklay, M. (2013). Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In D. Sui, S. Elwood, & M. Goodchild (Eds.), *Crowdsourcing Geographic*

*Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice* (pp. 105–122). Springer Netherlands.

- Henderson, L., & Green, C. (2020). Making sense of microplastics? Public understandings of plastic pollution. *Marine Pollution Bulletin*, *152*, 110908.
- Herodotou, C., Aristeidou, M., Sharples, M., & Scanlon, E. (2018). Designing citizen science tools for learning: Lessons learnt from the iterative development of nQuire. *Research and Practice in Technology Enhanced Learning*, 13(1), 4.
- Herodotou, C., Sharples, M., Gaved, M., Kukulska-Hulme, A., Rienties, B., Scanlon, E., & Whitelock, D. (2019). Innovative Pedagogies of the Future: An Evidence-Based Selection. *Frontiers in Education*, *4*.
- Herodotou, C., Villasclaras-Fernandez, E., & Sharples, M. (2014). *Scaffolding citizen inquiry science learning through the nQuire toolkit.* 5.
- Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), 99–107.
- International Coastal Cleanup®. (2017, March 7). Ocean Conservancy.
- Jordan, R. C., Gray, S. A., Howe, D. V., Brooks, W. R., & Ehrenfeld, J. G. (2011). Knowledge Gain and Behavioral Change in Citizen-Science Programs. *Conservation Biology*, *25*(6), 1148–1154.
- Kirchherr, J., Reike, D., & Hekkert, M. P. (2017). Conceptualizing the Circular Economy: An Analysis of 114 Definitions. SSRN Electronic Journal, 127.
- Kukulska-Hulme, A. (2011). Smart learning with mobile devices? NEW MEDIA.
- Lederman, J. S., Lederman, N. G., Bartos, S. A., Bartels, S. L., Meyer, A. A., & Schwartz, R. S. (2014). Meaningful assessment of learners' understandings about scientific inquiry—The views about scientific inquiry (VASI) questionnaire. *Journal of Research in Science Teaching*, *51*(1), 65–83.
- Mcneill, K. (2008). Inquiry and Scientific Explanations: Helping Students Use Evidence and Reasoning. Science as Inquiry in the Secondary Setting.
- Rickinson, M., & Lundholm, C. (2008). Exploring students' challenges in environmental learning. *Cambridge Journal of Education*, *38*, 341–353.
- Robertson, C. L., & Krugly-Smolska, E. (1997). Gaps between Advocated Practices and Teaching Realities in Environmental Education. *Environmental Education Research*, *3*(3), 311–326.
- Sarwi, S., Fauziah, N., & Astuti, B. (2018). The analysis of scientific communications and students' character development through guided inquiry learning. *Journal of Physics: Conference Series*, 983, 012031.
- Seifert, C. M., Harrington, M., Michal, A. L., & Shah, P. (2022). Causal theory error in college students' understanding of science studies. *Cognitive Research: Principles and Implications*, 7(1), 4.
- Sharples, M., Aristeidou, M., Villasclaras-Fern´andez, E., Herodotou, C., & Scanlon, E. (2015, October 17). A Smartphone Toolkit for Citizen Inquiry Learning. https://doi.org/10.1007/978-3-319-25684-9\_27
- Sharples, M., Mcandrew, P., Weller, M., Ferguson, R., FitzGerald, E., Hirst, T., & Gaved, M. (2013). *Innovating Pedagogy 2013: Exploring new forms of teaching, learning and assessment, to guide educators and policy makers.*
- Sholahuddin, A., Sholihah, A., Mahdian, & Susilowati, E. (2020). Can the guided inquiry with environment learning resources increase conceptual understanding and scientific literacy? *Journal of Physics: Conference Series*, *1422*(1), 012038. https://doi.org/10.1088/1742-6596/1422/1/012038
- Silvertown, J. (2009). A new dawn for citizen science. Trends in Ecology & Evolution, 24(9), 467-471.
- Srisawasdi, N., & Kroothkeaw, S. (2014). Supporting students' conceptual development of light refraction by simulation-based open inquiry with dual-situated learning model. *Journal of Computers in Education*, 1(1), 49–79.
- Srisawasdi, N., Panjaburee, P. (2019). Implementation of game-transformed inquiry-based learning to promote the understanding of and motivation to learn chemistry. *Journal of Science Education and Technology*, *28*, 152–164. https://doi.org/10.1007/s10956-018-9754-0
- Stone, E. M. (2014). Guiding Students to Develop an Understanding of Scientific Inquiry: A Science Skills Approach to Instruction and Assessment. *CBE Life Sciences Education*, *13*(1), 90–101.
- Suárez, Á., Specht, M., Prinsen, F., Kalz, M., & Ternier, S. (2018). A review of the types of mobile activities in mobile inquiry-based learning. *Computers & Education*, *118*, 38–55.
- Weigold, M. (2001). Communicating Science: A Review of the Literature. *Science Communication SCI COMMUN*, 23, 164–193.
- Williams, M., Gower, R., Green, J., Whitebread, E., Lenkiewicz, Z., & Schröder, P. (2019). *No Time to Waste: Tackling the Plastic Pollution Crisis Before it's Too Late*.

# W09

The First International workshop on Ethics of AIED: Challenges and Opportunities for Achieving Learner Centricity

# Ethical Challenges of Artificial Intelligence in Education: Achieving Learner Centricity with Respect to Learner Autonomy

Meera PAWAR<sup>\*</sup>, Ram Das RAI, Ramkumar RAJENDRAN IDP in Educational Technology, Indian Institute of Technology Bombay, India \*214380002@iitb.ac.in

**Abstract:** In the realm of education, the integration of Artificial Intelligence (AI) has emerged as a transformative force, offering the potential to reshape conventional learning paradigms through the provision of personalized and adaptive educational experiences that cater to the unique requirements of each learner. While the promise of AI in education is undeniably compelling, its implementation is accompanied by a nuanced tapestry of ethical challenges, notably in the context of learner autonomy and agency. The connection between learner autonomy, ethical considerations, and AI in education is carefully examined in this research. A thorough literature analysis reveals what is learner autonomy and its potential effects on it. It examines different ethical issues such as algorithmic biases and decreased active involvement that arise from balancing AI's educational augmentation with preserving learner autonomy. In order to strengthen student autonomy in AI-integrated learning settings, the study suggests best practices to address challenges faced for achieving learner autonomy.

**Keywords:** AIED, Leaner centricity, learner autonomy, learner agency, challenges, practices

# 1. Introduction

A rapidly growing topic called artificial intelligence in education (AIED) promises to provide each student with individualized learning opportunities that cater to their unique needs and learning preferences. The use of AIED as a technology engine for educational innovation can also be used to support the development and reform of instructional strategies and delivery systems. Additionally, AIED can be utilized as a technical instrument to unlock the "learning black box," allowing for the precise quantification and clear expression of social, psychological, and academic knowledge. However, the use of AIED has also raised more ethical questions and hazards about a number of issues, including learner autonomy and personal data (Nguyen et al., 2023). Additionally, AIED raises ethical issues such as the possibility that AI may be used to violate students' privacy, discriminate against members of marginalized groups, and foster a culture of surveillance in the classroom. Therefore, it is crucial to create moral standards for using AIED that guarantee that this technology is utilized in a way that is beneficial to all.

The integration of Artificial Intelligence in Education (AIED) introduces a complex ethical landscape concerning the pursuit of learner-centricity in relation to autonomy. Learner autonomy, a cornerstone of effective education, allows individuals the agency to shape their learning paths according to their preferences. However, the extensive capabilities of AIED systems, while aiming to optimize learning experiences, can inadvertently compromise learner autonomy if they become excessively prescriptive or make decisions on behalf of learners without their involvement. Such instances could curtail learners' freedom of choice, inhibiting their ability to explore diverse topics or adopt alternative learning strategies. Scholarly
discussions have highlighted concerns over the potential erosion of learner autonomy and agency arising from the integration of AI into education (Shum, S. J. B., & Luckin, R. (2019). . Moreover, the convergence of privacy and autonomy issues emerges with the concept of affective capture, wherein behavioral nudges informed by AI-driven insights could potentially infringe upon learners' privacy and autonomy (Nemorin, 2017). To comprehensively address these challenges, this study embarks on a systematic literature review to unravel the intricate dimensions of learner autonomy and agency in the context of AIED and learner-centricity.

In pursuit of a holistic understanding, this research seeks to answer pivotal research questions:

RQ1. Definition of "learner autonomy and agency": Delving into the conceptual underpinnings of learner autonomy and agency, this question aims to provide a clear and nuanced comprehension of these crucial aspects.

RQ2. Challenges in Achieving Learner-Centricity with Respect to Learner Autonomy and Agency: This question investigates the hurdles and complexities that emerge in the quest for learner-centricity, particularly concerning the preservation of learner autonomy and agency.

RQ3. Best Practices from the Literature to Address Challenges: By surveying existing literature, this question seeks to identify and elucidate best practices that have been proposed as effective strategies to mitigate the ethical challenges linked to learner autonomy and agency within the AIED framework.

Through this systematic exploration, we aim to illuminate the multifaceted terrain of learner autonomy and agency within AIED, providing insights into the challenges that arise and proposing potential solutions to foster a balanced integration of AI that upholds learner-centricity while respecting their autonomy and agency

# 2. Methodology

#### 2.1 Systematic literature review

The literature review commenced by selecting appropriate keywords and constructing a focused search string. To comprehensively assess the impact of AIED across various dimensions including transparency, fairness, privacy, and the integration of autonomy within this context, a targeted exploration was undertaken. The initial search string employed was "(Artificial Intelligence OR AI) AND (Education) AND (Ethics) AND (Transparency OR Accountability OR Privacy OR Fairness OR Explainability OR Autonomy)," which guided the retrieval of relevant literature. This inclusive combination of keywords facilitated a comprehensive search across 8 distinct databases, including ERIC, JSTOR, Taylor & Francis, Scopus, Science Direct, IEEE, and ACM. The scope of analysis was refined to encompass peer-reviewed articles published in the English language from the years 2013 to 2023. Subsequently, articles were subject to stringent screening, as illustrated in the PRISMA analysis in Figure 1. The figure depicts the application of diverse inclusion criteria and the resultant count of articles that met the requisites for thorough scrutiny and detailed analysis.

On the basis of the finalized publications on autonomy that were analyzed, we have provided answers to our study questions. The findings about autonomy are discussed in the section after this.



Figure. 1: Systematic Literature Review PRISMA Diagram

# 3. Findings

# 3.1 Concept of Autonomy and Agency

Ethical frameworks often link agency with autonomy, which means the ability to live by one's own rules. Agency refers to a person's ability to take action and make choices based on their personal goals and values. It determines the range of options and achievements they can pursue in a meaningful way (Ilkka Tuomi , 2023). In the ethical frameworks, it is considered that agency subsumes autonomy and it is connected to freedom, free will, and oppression. Positive freedom is the capacity to act on one's free will, while negative freedom is freedom from external control. Autonomy allows individuals to choose from these imagined alternatives based on their own evaluations and values. Autonomous actions reflect personal value choices, and the individual can be held morally responsible for their decisions. (Ilkka Tuomi , 2023). According to Right to Autonomy, learners should be provided the choices mentioned in figure 2 (Holmes, W., Persson, J., Chounta, I. A., Wasson, B., & Dimitrova, V. (2022).



Figure 2: Choices to be given to leaner

- 1. Individuals have a right not to be subject to a decision that is made solely through automated processing if these decisions have legal consequences or significantly impact individuals.
- 2. The ability to effectively question and challenge judgments informed and/or made by an AI system, as well as to request that such decisions be reviewed by a human.
- 3. The freedom to opt out of AI-enabled manipulation, personalized profiling, and projections.

Currently, AI assistants have become versatile in their functions, aiming to offer guidance and support to users. They fulfil various roles, from providing information and answering queries to managing tasks and offering recommendations. However, their presence and influence can impact how individuals perceive and process information. When AI assistants provide advice or recommendations, they act as external or extrinsic influences on a person's decision-making process. This can alter the way individuals perceive certain information, leading to changes in their cognitive biases. Cognitive biases are mental shortcuts or patterns that can influence how people interpret and make judgments about information and manipulate one's intrinsic motivation by encouraging and providing support (Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023).

Al systems, including personalized learning, intelligent tutoring, chatbots, emotional support, automated grading, learning analytics, language learning, and proctoring, are widely used in education. However, when these systems restrict learners' choices and excessively regulate and monitor their actions without their involvement or permission, it undermines learners' autonomy and agency (Akgun & Greenhow, 2021; Adams et al., 2023). Al systems often rely on vast amounts of data to make predictions, limiting learners' options to prevent perceived faults (Nguyen et al., 2023). Online proctoring during exams can also impede learners' natural behaviors, such as muttering to themselves or taking breaks (Coghlan et al., 2021).Learner autonomy is essential for empowering individuals and giving them control over their destinies. Neglecting learner autonomy poses challenges, emphasizing the importance of respecting learner choices. The following section discusses these challenges and the reasons why Al in education should uphold learner autonomy

#### 3.2 Ethical Challenges and the Importance of Learner Autonomy in AIED

Learner autonomy is a very crucial thing while a child is learning and when it is ignored learner gets affected deeply and it impacts their future learnings. According to Piaget, at the age of seven or eight, a child's moral thinking shifts from heteronomy to autonomy. Heteronomy is when a child uses authoritative rules for his/her moral judgments. Whereas autonomy is when a child takes moral judgments on their own will. Piaget stressed that autonomy grows through social cooperation. In child psychology, the development of identity, self-image, self-efficacy, and growth mindset is closely connected to the advancement of autonomy. Consequently, there are assertions that children are social actors who require comparable freedoms and rights as other members of society (Ilkka Tuomi, 2023). Worries about autonomy arise when leaners' freedom to act according to their own values and interests is restricted or reduced. For instance, predictive algorithms that predict future actions based on people's metadata can intensify existing biases and prejudices in society. This raises significant concerns about social inequalities (Akgun, S., & Greenhow, C., 2022).

One of the key ways this impact is seen is through the vast amount of data that AIED systems use. These systems gather data from various sources, including learner interactions, performance metrics, and historical patterns. However, this data might carry inherent biases, reflecting the societal and cultural contexts from which it was collected. When AIED systems rely heavily on biased data, they can inadvertently perpetuate and amplify existing inequalities and stereotypes. Additionally, the prescriptive nature of AIED systems can impact learners' sense of autonomy. When students receive automated recommendations and instructions, they might feel compelled to follow them without fully understanding the rationale behind those

choices. This could diminish their ownership and agency over their learning journey and weaken their critical thinking skills.

The challenges relating to learners' autonomy are highlighted by the recently expanded use of AI in education. A few number of scholars have discussed these difficulties and provided a few practices that deal with learner autonomy in a leaner-centric approach.

#### 3.2 Best practices for ensuring learner autonomy and agency

The previous sections discussed what is learner autonomy, and what are the challenges. This section, Table 1 lists many strategies for overcoming and challenges related to sustaining to learner autonomy.

Practices	What it addresses		
Obtaining learner assent	Respects learners' capacity for choice and autonomy		
Clearly communicating that learners have the freedom to opt out of sharing their data	Respect learners' privacy and autonomy		
Leveraging AI systems to enhance learners' Agency	Fosters autonomy across cognitive, social, and cultural dimensions		
Prioritizing accuracy in AI development and deployment	Avoids false information dissemination and safeguards users' capacity for autonomous cognition, emotional well- being, and social development		
Avoiding the use of words or algorithms promoting computational propaganda in AIED systems	Protects learners from manipulation		
Providing comprehensive training programs for educators	Equips educators with the requisite skills to implement AIED effectively and reduce potential manipulation of learners' thought processes		

Table 1: Practices to address challenges related to learner autonomy

In K-12 educational settings, where legal consent might not be applicable to children, a proactive approach to empower learners is through obtaining their assent. It is imperative to clearly communicate, using age-appropriate language, that learners have the freedom to opt out of sharing their data, thereby respecting their capacity for choice (Adams et al., 2023), (Nguyen et al., 2023). Leveraging AI systems to enhance learners' agency fosters autonomy across cognitive, social, and cultural dimensions, enabling them to realize their full potential (Adams et al., 2023).

However, while interacting with AI systems, challenges arise, potentially limiting learner autonomy. For example, the provision of hints and suggestions by the system can impact decision-making. To mitigate this, it is essential that AI development and deployment prioritize accuracy, avoiding false information dissemination, and safeguarding users' capacity for autonomous cognition, emotional well-being, and social development (Nguyen et al., 2023). Ethical considerations extend to the linguistic and algorithmic elements employed; the use of words or algorithms promoting computational propaganda should be avoided within AIED systems (European Parliament, 2021).

In facilitating this transition, comprehensive training programs for educators are crucial. Equipping educators with the requisite skills to implement AIED effectively enables them to tailor, filter, or minimize the automation of AI systems, thereby reducing potential manipulation of learners' thought processes and promoting their identity development. This multifaceted approach, encompassing informed assent, learner agency, ethical system design, and educator training, serves as a foundation for the ethical and responsible integration of AI in educational contexts

#### 4. Discussion and Conclusion

Personalized learning through artificial intelligence in education (AIED) has the potential to transform education by tailoring content to individual students. However, ethical concerns arise, particularly concerning learner autonomy and agency. Learner autonomy refers to a student's ability to choose what they learn, while learner agency involves taking control of their learning. AIED systems that are overly prescriptive can unintentionally limit learner autonomy and agency. Addressing these ethical issues requires a comprehensive strategy to preserve and respect learner autonomy and agency. Key steps include providing transparency and choice to learners. This means giving students clear information about how AIED systems work and allowing them to opt out of certain features, empowering them to actively participate in their learning journey. Furthermore, ensuring fairness and impartiality within AIED systems is crucial. Rigorous data collection and analysis methods, along with careful curation of features and algorithms, are essential to prevent discrimination and ensure equal access to quality education for all learners. Protecting learner privacy is another cornerstone of this ethical framework. Balancing the collection and use of data with the principles of necessity and proportionality is vital. Learners should also have the ability to access and control their own data. Empowering learners with knowledge is equally important. Educating them about the potential benefits and risks of AIED helps them make informed decisions and exercise informed agency. Additionally, stakeholders like educators, educational designers, institutions, policymakers, and developers can benefit from these insights. Educators can use this information to navigate ethical challenges and use AIED in ways that respect learner autonomy and agency. Policymakers can develop responsible AIED policies, and developers can create ethical and beneficial AIED systems.

#### 5. References

- Adams, C., Pente, P., Lemermeyer, G., & Rockwell, G. (2023). Ethical principles for artificial intelligence in K-12 education. *Computers and Education: Artificial Intelligence*, *4*, 100131.
- Akgun, S., & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 1-10.
- Coghlan, S., Miller, T., & Paterson, J. (2021). Good proctor or "big brother"? Ethics of online exam supervision technologies. *Philosophy & Technology*, *34*(4), 1581-1606.
- Holmes, W., Persson, J., Chounta, I. A., Wasson, B., & Dimitrova, V. (2022). *Artificial intelligence and education: A critical view through the lens of human rights, democracy and the rule of law.* Council of Europe.
- Nemorin, S. (2017). Post-panoptic pedagogies: The changing nature of school surveillance in the digital age. *Surveillance and Society*, *15*(2), 239-253.
- Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, *28*(4), 4221-4241.
- Shum, S. J. B., & Luckin, R. (2019). Learning analytics and AI: Politics, pedagogy and practices. *British journal of educational technology*, *50*(6), 2785-2793.
- Tuomi, I. (2023). A framework for socio-developmental ethics in educational AI.

# Ethical Challenges and Best Practices for Transparency in AIED: A Literature Review and Learner Centric Guidelines

#### Ram Das RAI<sup>\*</sup>, Meera PAWAR & Ramkumar RAJENDRAN Indian Institute of Technology Bombay, India

**Abstract:** Artificial intelligence in education (AIED) is the use of artificial intelligence techniques and tools to enhance and support learning and teaching processes. However, AIED also raises ethical challenges related to transparency, which is the ability to understand how AIED systems make decisions that influence educational outcomes. This paper reviews the literature on transparency in AIED, and presents a set of learner centric guidelines for ensuring transparency. The paper uses the Transparency Index Framework by Chaudhry et al. (2022) as a guiding tool, and discusses the best practices for providing information, allowing oversight, and respecting rights and choices in AIED. The paper also illustrates these practices with examples of ethically and transparently designed AIED systems. The paper concludes by highlighting the trade-offs and precautions involved in ensuring transparency in AIED, and suggests future research directions. This paper contributes to the ethical discourse on transparency in AIED by providing a comprehensive overview, identifying the best practices and challenges, and proposing the learner centric guidelines.

**Keywords:** artificial intelligence in education, ethics, transparency, accountability, safety, fairness, explainability

# 1. Introduction

Artificial intelligence in education (AIED) is a rapidly developing field that aims to create personalized learning experiences for each student, tailored to their individual needs and learning styles (Xu, 2020). AIED can also be used as a technological engine of educational innovation, promoting the reform and development of educational modes and teaching methods. Furthermore, AIED can be used as a technical tool to open the "learning black box," through which educational, psychological, and social knowledge can be accurately quantified and made explicit. However, the adoption of AIED has also led to increasing ethical risks and concerns regarding several aspects such as personal data and learner autonomy (Nguyen et al., 2023). Moreover, AIED poses ethical challenges such as the potential for AI to be used to discriminate against students from marginalized groups, violate students' privacy, and create a culture of surveillance in the classroom. Therefore, it is important to develop ethical guidelines for the use of AIED that ensure that this technology is used in a way that is beneficial to all.

Some of the ethical aspects of AIED that have been identified by researchers and practitioners include cultural integration, accountability, fairness, equity, affordability, security, privacy, and human dignity (Holmes et al., 2022). These aspects reflect the values and principles that should guide the design, development, and deployment of AIED systems. Among these aspects, transparency stands out as a key ethical principle that guides the use of artificial intelligence in education (AIED). Transparency means that AIED systems should provide sufficient information and explanation about how they make decisions and what data they use. Transparency is important for ensuring fairness, accountability, and trust in AIED systems. However, transparency is often challenged by the issue of "black boxes" in AI, which are software that do not reveal how they analyze input data to arrive at their outputs. This can create ethical dilemmas and risks in various domains, such as medical education, where AI

systems are intended to assist physicians with patient care (Katznelson & Gerke, 2021). Therefore, ethicists emphasize the moral requirement for transparency in AI and argue that algorithms should be accompanied by explainability (Farrow, 2023). Moreover, transparency is valued by educators and learners who use AIED systems. For example, Holmes et al. (2022) found that educators might prefer an AI system that sacrifices some accuracy in favor of providing greater transparency in decision-making processes. Yu and Yu (2023) also stated that "transparency stands out as the primary ethical principle guiding the advancement of AI implementation in the field of education" (p. 9).

Previous researchers have explored the ethical challenges of AIED (Adams et al., 2023, Holmes et al., 2022, Nguyen et al., 2023). However, most work has been broad and few studies have focused on specific aspects like transparency or learner autonomy. Chaudhry, Cukurova, and Luckin (2022) addressed transparency in AIED using an ML pipeline framework. They accounted for transparency at various stages of the AI development process from data collection to deployment. They shortlisted studies using popular frameworks for documentation, robustness and reproducibility for each ML pipeline. They also improved their framework based on interviews of education stakeholders. Their framework was robust and new, but mainly technical. Farrow (2023) emphasized using a socio-technical lens to look at ethics in AIED. This means looking beyond technical solutions and incorporating educational, legal and policy dimensions. There is also a need to consider trade-offs when ensuring transparency, such as sacrificing accuracy. This is a complex issue that needs discussion. Therefore, more research is needed in this area using a systematic and socio-technical approach to ethics in AIED.

The purpose of this paper is to review the literature on transparency in artificial intelligence in education (AIED), and to present a set of learner centric guidelines for ensuring transparency. Transparency is an ethical challenge in AIED, as it affects the ability to understand how AIED systems make decisions that influence educational outcomes. The paper uses the Transparency Index Framework by Chaudhry et al. (2022) as a guiding tool, and discusses the best practices for providing information, allowing oversight, and respecting rights and choices in AIED. The paper also illustrates these practices with examples of ethically and transparency in AIED by providing a comprehensive overview of the existing literature, identifying the best practices and challenges, and proposing the learner centric guidelines. The paper is guided by the following research questions:

- What are the causes and consequences of the problem of lack of transparency in artificial intelligence in education (AIED)?
- How can effective and ethical strategies be developed and implemented to address the problem of lack of transparency in AIED and its associated challenges and trade-offs?

In this paper, the authors first conceptualized transparency based on the Transparency Index Framework proposed by Chaudhry, Cukurova, and Luckin (2022). This framework offers a comprehensive and multidimensional perspective on transparency in AI in education. They then derived a keyword string from this framework: (Artificial Intelligence OR AI) AND (Education) AND (Ethics) AND (Transparency OR Accountability OR Safety OR Fairness OR Explainability OR Autonomy). They conducted a systematic search using the PRISMA framework and shortlisted 15 papers that met their inclusion criteria (Moher, D., et al., 2015). In addition to these papers, the authors also referred to the document Recommendation on the Ethics of Artificial Intelligence by UNESCO (2021), which provides a global standard for ethical AI in the educational domain. The details of the PRISMA process are illustrated in Fig 1.



Fig. 1. PRISMA analysis details

This paper has three sections. Section 2 explains the objectives and method of the paper. Section 2.1 discusses the lack of transparency in AIED and its implications for decision-making, learning outcomes, and ethical values. Section 2.2 explores some strategies to address the ethical challenges of AIED, such as legal and policy frameworks, learner autonomy and agency, human oversight and accountability, and trade-offs and precautions. Section 3 discusses the trade-offs involved and the precautions that must be taken when implementing transparency in AIED. It also concludes the paper with a summary and some suggestions for future research.

# 2. Findings

This section presents the findings of the systematic review on ethical transparency in AIED. The section is organized into two subsections: the first one discusses the lack of transparency in AIED and its ethical implications, and the second one proposes strategies to address these ethical challenges. These findings are illustrated in Fig 2 which lists down the findings of this paper.



Fig. 2. A list of the findings on lack of transparency in AIED and strategies to address ethical challenges

# 2.1 Lack of transparency in AIED

# 2.1.1 Concept of transparency in AIED

Transparency is a key concept in artificial intelligence in education (AIED) that has implications for decision-making processes. Transparency can be defined in different ways depending on the perspective and the context of AI use in education or information transfer. However, a common theme is that transparency refers to the availability and accessibility of information about how AI systems make decisions and what data they use. This section presents some possible definitions of transparency from the literature. Turilli and Floridi (2009) defined transparency as the availability of information, the degree of availability, or the role of information in the decision-making process (as cited in Yu & Yu, 2023). Jobin et al. (2019) defined transparency as the requirement that when AI is used in education or information transfer, the specific parameters, source, responsibility, investment, and impact of AI should be disclosed to enhance accountability, modification, interpretation, and communication between AI users, educational researchers, and practitioners (as cited in Adams et al., 2023). Siegel et al. (2018) defined transparency as the degree to which the determinations or predictions of AI systems are revealed to relevant parties in ways that those parties prefer and can understand (as cited in Coghlan et al., 2021). These definitions illustrate the complexity and diversity of transparency in AIED. Transparency is also closely related to other concepts such as explainability, explicability, responsibility, and accountability in the context of AI and data (Nguyen et al., 2023). Yu and Yu (2023) emphasized that transparency has become a prominent concern in the use of AI for educational purposes. Transparency is important for ensuring fairness, trust, and communication in AIED systems.

# 2.1.2 Examples of opaque decision-making in AIED

Transparency in artificial intelligence in education (AIED) is important for understanding how AI systems are making decisions and what are the implications of those decisions for the users. However, transparency can be challenged by the existence of AI systems that are not transparent, meaning that they do not reveal how they make decisions or predictions, or what data or models they use. This is also known as the 'black box' problem, or the lack of transparency of the processes and workings that convert input to output, which is a major structural aspect of the 'black box' model of computation (Farrow, 2023). The 'black box' problem can create various ethical issues and risks for the AIED community and stakeholders, such as:

*False trust and confidence*: Users may not understand, question, or challenge the AI decisions or predictions that are accurate but not transparent (Holmes et al., 2022).

*Negative impact on autonomy, agency, and self-regulation*: Al systems may influence the learning process without providing a detailed account of how the algorithm works, or how it affects the learners' behavior, motivation, or outcomes (Farrow, 2023).

*Reduction of engagement, motivation, and satisfaction:* Al systems may deliver feedback to learners without explaining the criteria, evidence, or rationale behind it, or how it can be used to improve learning (Farrow, 2023).

*Unfair or discriminatory outcomes*: Al systems may use biased or incomplete data sets that favor certain groups of learners over others, or that do not account for the diversity and variability of learners' needs, preferences, and contexts (Holmes et al., 2022).

*Violation of rights, interests, and well-being*: Al systems may not disclose the purposes, methods, or outcomes of their data collection, analysis, or reporting, or that do not obtain informed consent or respect the privacy and ownership of the users' data (Sahlgren, 2023).

*Confusion, inconsistency, or ambiguity*: Al systems may not reveal the regulatory roles, responsibilities, or conditions of their design and use, or that do not address the ethical challenges or concerns raised by the AIED community or other stakeholders (Sahlgren, 2023).

#### 2.1.3 Ethical challenges associated with lack of transparency in AIED

The lack of transparency in different aspects of artificial intelligence in education (AIED) systems can pose various ethical challenges for the AIED community and stakeholders. Chaudhry et al. (2022) classified the lack of transparency into three categories: data collection and analysis, decision making, and deployment of the AIED environments.

Data collection and analysis. The lack of transparency in data collection, analysis, and reporting may raise issues of informed consent, data ownership, data privacy, data quality, and data accountability. These issues can affect the rights, interests, and well-being of the individuals whose data are collected, processed, or used by the AIED systems (Holmes et al., 2022).

*Decision making.* The lack of transparency in the teaching decisions and the system's model of the learner may affect the trust, understanding, and agency of the learners and educators, and may limit their ability to challenge or improve the AIED systems. These issues can affect the learning outcomes, experiences, and opportunities of the learners and educators who interact with the AIED systems (Holmes et al., 2022).

Deployment of the AIED environments. The lack of transparency in the design and deployment of AIED systems may result in unintended or harmful consequences for the learners, educators, and wider society, such as bias, discrimination, manipulation, or exploitation. These issues can affect the quality, equity, and accountability of the AIED systems and their impact on various aspects of society (Holmes et al., 2022).

In addition to these three categories, there are also other ethical issues that can arise from the lack of transparency in AIED systems, such as:

*Ethical guidelines and regulations.* The lack of transparency in the ethical guidelines and regulations of AIED systems may create confusion, inconsistency, or ambiguity among the AIED researchers and developers, and may hinder their ethical responsibility and accountability. These issues can affect the validity, reliability, and credibility of the AIED research and development and their compliance with ethical standards and legal obligations (Holmes et al., 2022).

Benefits and costs of AIED innovation. The lack of transparency in the potential benefits and costs of AIED innovation may prevent the AIED community from balancing the ethical risks and opportunities of their work, and may undermine their social impact and value. These issues can affect the innovation potential, sustainability, and scalability of the AIED systems and their contribution to social good (Holmes et al., 2022).

#### 2.1.4 Examples of transparent decision-making in AIED

Some examples of transparency in artificial intelligence in education (AIED) systems are as follows. One example is bots or intelligent tutoring systems (ITS) that provide clear and understandable feedback, guidance, and explanations to learners and educators about their learning goals, progress, and outcomes. These systems can help to enhance the trust, understanding, and agency of the users and support their learning processes and outcomes (Farrow, 2023). Another example is automated assessment tools that reveal the criteria, methods, and evidence used to evaluate learners' performance and provide constructive suggestions for improvement. These tools can help to ensure the validity, reliability, and fairness of the assessment and foster the learning and development of the users (Farrow, 2023). A third example is predictive analytics dashboards that display the data sources, assumptions, and limitations of the models used to generate predictions and recommendations for learners, educators, and institutions. These dashboards can help to inform the users about the rationale, accuracy, and uncertainty of the predictions and recommendations and enable them to make informed decisions and actions (Farrow, 2023). A fourth example is plagiarism checkers that show the original sources of the text that are matched with the learners' submissions and indicate the degree of similarity and originality. These checkers can help to promote the academic integrity, honesty, and creativity of the users and prevent the misuse or abuse of the AI systems (Farrow, 2023). A fifth example is fairness-promoting algorithms that monitor and mitigate the potential bias or discrimination in the data sets, models, or outcomes of AIED systems and ensure that they benefit all learners equally. These algorithms can help to protect the rights, interests, and well-being of the users and foster the quality, equity, and accountability of the AIED systems (Holmes et al., 2022).

# 2.2 Strategies to address ethical challenges

# 2.2.1 Legal and policy frameworks to address bias and discrimination in AIED

To address the issues of bias and discrimination in AIED and ensure transparency, several legal and policy frameworks have been proposed by researchers. One of them is a code of best practice that the AIED community can rely on in designing and deploying AIED technologies in diverse educational contexts, incorporating guidance addressing the issues of fairness, accountability, transparency, bias, autonomy, agency, and inclusion specifically at the intersection of AI and the learning sciences (Holmes et al., 2022). Another framework is an educational framework that embeds the principles of AI ethics in education, and also ethics issues such as responsibility, inclusion, fairness, security and explainability in conducting educational research, to raise awareness and foster critical thinking among educators and learners (Nguven et al., 2023), Furthermore, a website that publishes and updates the ethical principles for AIED, inviting feedback and improvement suggestions from various stakeholders, such as researchers, educators, learners, policymakers, and the public, has been suggested as a way to facilitate communication and collaboration on AIED ethics (Nguyen et al., 2023). Additionally, an automatic method by text analysis that can evaluate the ethical implications of AIED research papers and systems, providing feedback and recommendations to the authors and reviewers, has been developed to enhance the quality and accountability of AIED research (Holmes et al., 2022). Finally, a regulatory framework that ensures that AIED systems are subject to rigorous testing and auditing before and after deployment, to detect and mitigate any potential harms or biases that might affect the users or the society, has been advocated as a necessary measure to ensure the trustworthiness and responsibility of AIED systems (Holmes et al., 2022).

#### 2.2.2 Strategies to promote learner autonomy and agency in AIED

Researchers have suggested several strategies to promote learner autonomy and agency in AIED and ensure transparency. One of them is developing negotiation-based adaptive

learning systems that allow learners to choose the type and frequency of support, scaffolding of not only knowledge but also metacognition and self-regulation skills, and feedback on their learning progress and outcomes (Nguyen et al., 2023). Another strategy is providing learners with control over their data, such as the ability to access, edit, delete, or share their data with others, and informing them about how their data is collected, stored, processed, and used by the AIED system (Holmes et al., 2022). Moreover, adopting fairness-promoting algorithms that can detect and mitigate any potential biases or harms that might affect different subpopulations and demographics of learners, and ensuring that the AIED system does not favor or disadvantage any groups of learners based on their characteristics or backgrounds, is a crucial strategy to ensure equity and justice in AIED (Holmes et al., 2022). Furthermore, explaining the role and function of the AIED system to the learners in a simple and understandable way, such as using visualizations, examples, or analogies, and providing them with opportunities to ask questions, give feedback, or challenge the system's decisions, is a key strategy to enhance trust and engagement in AIED (Farrow, 2023). Finally, encouraging learners to develop AI literacy skills, such as understanding the basic concepts and principles of AI, its capabilities and limitations, its ethical and social implications, and how to collaborate effectively with AI systems, is a vital strategy to empower learners and prepare them for the future of education (Nguyen et al., 2023).

# 2.2.3 The importance of human oversight and accountability in AIED

Transparency is a fundamental principle for ethical AIED, and it requires human oversight and accountability from the various stakeholders involved in the design and use of AIED systems. One of the reasons why human oversight and accountability in AIED are important is that they ensure that AIED systems are developed and used with good intentions, integrity, and careful consideration of potentially harmful factors, such as bias, discrimination, or manipulation (Nguyen et al., 2023). Another reason is that they empower the human stakeholders, such as educators and learners, to have control over the design and use of AIED systems, and to make informed decisions about when and how to rely on AIED systems for their educational purposes (UNESCO, 2021). Moreover, human oversight and accountability in AIED allow the attribution of ethical and legal responsibility for any stage of the life cycle of AIED systems, as well as in cases of remedy related to AIED systems, to physical persons or to existing legal entities (UNESCO, 2021). Furthermore, human oversight and accountability in AIED enable the protection of personal information throughout the life cycle of the AIED system, by adopting a privacy by design approach that considers the potential risks and harms of data collection, processing, and use (UNESCO, 2021). Finally, human oversight and accountability in AIED facilitate the auditability, traceability, explainability, and communicability of the AIED systems, by providing clear and open information about the data, the system, and the business model of AIED, as well as the purposes, conditions, roles, responsibilities, and trade-offs involved in AIED (Nguyen et al., 2023).

# 3. Discussion and Conclusion

This paper has reviewed the literature on transparency in artificial intelligence in education (AIED), and presented a set of best practices for ensuring transparency in AIED systems. Transparency is an ethical challenge in AIED, as it affects the ability to understand how AIED systems make decisions that influence educational outcomes. The paper has discussed the best practices for providing information, allowing oversight, and respecting rights and choices in AIED, and illustrated them with examples of ethically and transparently designed AIED systems. The paper has also examined the role of transparency in fostering communication and collaboration among the different stakeholders involved in AIED, such as educators, edtech experts, and AI practitioners. The paper has contributed to the ethical discourse on transparency in AIED by providing a comprehensive overview of the existing literature, identifying the best practices and challenges, and proposing the learner centric guidelines. The paper has also acknowledged the trade-offs and precautions involved in ensuring

transparency in AIED, such as privacy and security issues, information and cognitive overload, data or algorithm misuse or abuse, ethical dilemmas or conflicts, and unintended or adverse outcomes. In the ensuing discussion, the paper offers the authors' perspective on the initially posed questions and provides key takeaways for the readers.

# 3.1 Causes and Consequences of Lack of Transparency in AIED

The origins of the deficiency in transparency in AIED are intricately rooted in a combination of factors. These factors include biased algorithms, concerns surrounding data privacy, and a general lack of transparency in AI-driven decision-making processes, all of which contribute to this problem. Consequently, the implications are extensive, impacting not only the impartiality and fairness of educational outcomes but also eroding confidence in AI-driven educational technologies. These challenges, while not insurmountable, demand purposeful and ethically grounded approaches.

# 3.2 Strategies to Address Ethical Challenges

The review outlined a range of strategies to tackle the ethical challenges associated with transparency in AIED. Legal and policy frameworks play a pivotal role in ensuring accountability, fairness, and transparency in the development and deployment of AIED systems. Approaches such as the formulation of a code of best practice, the integration of AI ethics principles into educational frameworks, and the establishment of websites dedicated to ethical guidelines hold significant promise.

Promoting learner autonomy and agency in AIED is imperative to empower students and ensure transparency. Granting learners control over their data, embracing fairnesspromoting algorithms, and facilitating comprehension through explanations are measures that can foster trust and engagement in AI-enhanced educational environments. Additionally, equipping learners with AI literacy skills prepares them to confront the ethical challenges presented by an AI-driven world.

# 3.3 The Importance of Human Oversight and Accountability

Transparency in AIED cannot be achieved without the presence of human oversight and accountability. Active participation in the design and utilization of AIED systems by educators, learners, policymakers, and other stakeholders is paramount. Such involvement ensures that systems are developed and employed with ethical considerations, integrity, and due regard for potential adverse consequences. Moreover, it facilitates the allocation of ethical and legal responsibility and safeguards personal information via a privacy-by-design approach. Crucially, human oversight and accountability enhance the auditability, traceability, and explainability of AIED systems, thereby fostering transparency and trust.

# 3.4 Trade-offs involved when implementing ethical use of AIED to ensure transparency

Ethical use of AIED, particularly to ensure transparency, involves some trade-offs that need to be carefully considered and balanced by the researchers, developers, and users of AIED systems. One of the trade-offs is between accuracy and explainability of AI algorithms, as some AI models, such as deep neural networks, may achieve high performance but are difficult to interpret or justify their decisions (Nguyen et al., 2023). Another trade-off is between scalability and contextualization of AIED systems, as some AIED systems may be designed to reach a large and diverse audience but may not account for the specific needs, preferences, and cultures of different groups of learners or educators (Holmes et al., 2022). Moreover, a trade-off exists between innovation and regulation of AIED research and development, as some AIED research and development may aim to explore new possibilities and opportunities but may face ethical challenges or constraints from existing policies or guidelines (Holmes et al., 2011).

al., 2022). Furthermore, a trade-off occurs between privacy and personalization of AIED systems, as some AIED systems may require a large amount of data from the users to provide tailored and adaptive support but may also pose risks to the users' data protection and ownership (Nguyen et al., 2023). Finally, a trade-off arises between autonomy and oversight of AIED systems, as some AIED systems may enable the users to have more control and agency over their learning but may also require human intervention or supervision to ensure quality and accountability (Holmes et al., 2022).

#### 3.5 Precautions that must be taken while ensuring transparency in AIED

Ensuring transparency in AIED requires some precautions that must be taken by the researchers, developers, and users of AIED systems. One of the precautions is to respect the privacy and confidentiality of the users' information and to obtain well-informed consent from the users before using their data for any purpose (Nguyen et al., 2023). Another precaution is to explain and justify the AI algorithms for specific educational purposes, but to avoid revealing sensitive or personal information about the users, such as their identity, performance. behavior, or preferences (Nguyen et al., 2023). Moreover, a precaution is to be transparent, traceable, explainable, and communicable in the process of establishing, conducting, monitoring, and controlling regulations of AIED, but to balance the interests and rights of different stakeholders, such as developers, educators, policymakers, and users (Nguyen et al., 2023). Furthermore, a precaution is to be transparent in the audibility of the regulation of AIED and to address acknowledgment and responsibility for each stakeholder's actions involved in the design and use of AIED, but to consider the potential trade-offs, risks, and harms that may arise from the use of AIED systems (Nguyen et al., 2023). Finally, a precaution is to accompany the transparency of AIED systems with the provision of guidance, feedback, and support to the users, especially the learners, to help them understand, trust, and challenge the AIED systems, and exercise their agency and rights over their own data and learning outcomes (Nguyen et al., 2023). Additionally, a precaution is to ensure transparency with respect to the needs of the user. An excessive amount of transparency can potentially overwhelm users, for example, when intricate details concerning the AI Model are disclosed to individuals lacking technical expertise (S. Amer-Yahia, personal communication, July 14, 2023). Striking the right balance between transparency and privacy is equally imperative. For instance, in the context of an AIED (Artificial Intelligence in Education) system, disclosure of the reasons behind a student's failure in a specific test should remain undisclosed unless expressly requested by the student, as this information must not be accessible to others (S. Amer-Yahia, personal communication, July 14, 2023).

#### 3.6 Limitations and future research directions

This paper has some limitations that should be acknowledged and addressed in future research. First, it may not capture the full range and diversity of perspectives and approaches on transparency in AIED, as it may have excluded some relevant papers that used different terms or frameworks. Second, it does not consider other ethical principles that are equally important for AIED, such as privacy, beneficence, non-maleficence, justice, and human dignity. These principles are interrelated and may have implications for transparency in AIED. Third, it proposes some strategies to address the ethical challenges of AIED, such as legal and policy frameworks, learner autonomy and agency, human oversight and accountability, and trade-offs and precautions. However, these strategies are not exhaustive or definitive, and they may have their own limitations and challenges. To advance the research on transparency in AIED, we suggest the following future directions based on the limitations of this paper: (1) conduct a more comprehensive and systematic review of the literature on transparency in AIED, using a broader range of search terms and databases, and applying a more rigorous and transparent selection and analysis process; (2) explore the interplay and trade-offs between transparency and other ethical principles in AIED, such as privacy, beneficence, non-maleficence, justice, and human dignity; (3) evaluate the effectiveness and feasibility of the proposed strategies to address the ethical challenges of AIED, such as legal

and policy frameworks, learner autonomy and agency, human oversight and accountability, and trade-offs and precautions.

The paper hopes to stimulate more discussion and debate on the ethical challenge of transparency in AIED, as well as to inspire more innovation and collaboration among the AIED community. Transparency is not only a technical or legal issue, but also a moral and social one. Transparency is not only a means to an end, but also an end in itself. Transparency is not only a challenge for AIED, but also an opportunity for AIED. By embracing transparency as an ethical value and a design principle, AIED can become more trustworthy, responsible, and beneficial for education.

#### Acknowledgements

We would like to thank Dr. Shitanshu Mishra and the rest of the members from the UNESCO MGIEP International Working Group on Ethics of AIED for their help with this paper.

#### References

- Adams, C., Pente, P., Lemermeyer, G., & Rockwell, G. (2023). Ethical principles for artificial intelligence in K-12 education. Computers and Education: Artificial Intelligence, 4, 100131. https://doi.org/10.1016/j.caeai.2023.100131
- Chaudhry, M. A., Cukurova, M., & Luckin, R. (2022). A Transparency Index Framework for AI in Education. In M. M. Rodrigo, N. Matsuda, A. I. Cristea, & V. Dimitrova (Eds.), Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium (pp. 195–198). Springer International Publishing. https://doi.org/10.1007/978-3-031-11647-6\_33
- Coghlan, S., Miller, T., & Paterson, J. (2021). Good Proctor or "Big Brother"? Ethics of Online Exam Supervision Technologies. Philosophy & Technology, 34(4), 1581–1606. https://doi.org/10.1007/s13347-021-00476-1
- Farrow, R. (2023). The possibilities and limits of XAI in education: A socio-technical perspective. Learning, Media and Technology, 48(2), 266–279. https://doi.org/10.1080/17439884.2023.2185630
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in Education: Towards a Community-Wide Framework. International Journal of Artificial Intelligence in Education, 32(3), 504–526. https://doi.org/10.1007/s40593-021-00239-1
- Katznelson, G., & Gerke, S. (2021). The need for health AI ethics in medical school education. Advances in Health Sciences Education, 26(4), 1447–1458. https://doi.org/10.1007/s10459-021-10040-3
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic reviews, 4(1), 1-9.
- Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B.-P. T. (2023). Ethical principles for artificial intelligence in education. Education and Information Technologies, 28(4), 4221–4241. https://doi.org/10.1007/s10639-022-11316-w
- Sahlgren, O. (2023). The politics and reciprocal (re)configuration of accountability and fairness in datadriven education. Learning, Media and Technology, 48(1), 95–108. https://doi.org/10.1080/17439884.2021.1986065
- UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence—UNESCO Digital Library. https://unesdoc.unesco.org/ark:/48223/pf0000381137
- Xu, L. (2021). The Dilemma and Countermeasures of AI in Educational Application. Proceedings of the 2020 4th International Conference on Computer Science and Artificial Intelligence, 289–294. https://doi.org/10.1145/3445815.3445863
- Yu, L., & Yu, Z. (2023). Qualitative and quantitative analyses of artificial intelligence ethics in education using VOSviewer and CitNetExplorer. Frontiers in Psychology, 14. https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1061778.

# Ethical Implications of Utilizing Artificial Intelligence in Education for Assessment

#### Mihir PATKI<sup>a\*</sup>, Saira SANADI<sup>b\*</sup>, Shraddha JADHAV<sup>c\*</sup>, Ashish MUSALE<sup>d\*</sup>, & Kapil KADAM<sup>e\*</sup>

<sup>a,b,c,d,e</sup> Kolhapur Institute of Technology's (KIT) College of Engineering, Kolhapur, India <sup>a</sup>mihirpatki80@gmail.com, <sup>b</sup>sairasanadi29@gmail.com, <sup>c</sup>jadhavshraddha3011@gmail.com, <sup>d</sup>musaleashish2911@gmail.com, <sup>e</sup>kkapilkk@gmail.com \*All authors have contributed equally.

Abstract: The growing integration of Artificial Intelligence (AI) in education has brought forth ethical quandaries, particularly concerning its role in assessment methodologies. This research paper delves into the intricate ethical implications linked to the use of Al-driven technologies for educational evaluation. Through an extensive literature review, this study explores the potential merits and drawbacks of Al application in educational assessments, addressing concerns encompassing fairness, transparency, privacy and the inherent biases embedded in algorithmic decision-making. This article critically assesses the ethical aspects of AI incorporation in formative and summative evaluations, examining its repercussions on student learning encounters and academic outcomes. By scrutinizing the viewpoints of educators, students, policymakers and technologists, this paper provides insights into the complex ethical predicaments arising when AI assumes the role of evaluating human aptitudes and understanding. The findings underscore the essentiality of a well-balanced approach that upholds ethical principles, human discernment and the responsible advancement of AI technologies to ensure equitable and impartial educational assessment practices during the AI-driven epoch.

Keywords: Ethical implication, Al-driven technologies, Educational assessments.

#### 1. Introduction

The rapid integration of Artificial Intelligence (AI) into diverse domains has opened up a world of opportunities and complexities, and the education sector is no exception. In particular, the application of AI in educational assessment has ignited significant interest and discussions. Assessment, a cornerstone of education, traditionally reliant on human evaluation, now stands at the crossroads of technological innovation with the potential for improved efficiency and objectivity. However, this transition brings forth a host of ethical considerations that warrant careful examination.

This paper sets out to explore the intricate ethical implications surrounding the utilization of AI in educational assessment. As AI-driven assessment tools garner momentum, questions arise concerning fairness, transparency, privacy, bias and the broader impact on the quality of learning experiences. Addressing these ethical dimensions becomes pivotal in harnessing the transformative potential of AI while ensuring the preservation of ethical values. By conducting a comprehensive review of the existing literature, this study aims to shed light on the multifaceted ethical dilemmas entwined with the integration of AI in educators to technologists, a nuanced understanding of the complex challenges associated with entrusting algorithms with the evaluation of human knowledge and capabilities will emerge. Moreover, this research critically evaluates the ethical

considerations within both formative and summative assessment contexts, providing a holistic view of AI's possible implications on educational methodologies.

Ultimately, the study underscores the significance of a balanced approach that upholds ethical principles while harnessing the potential of AI in education. By engaging with these ethical concerns, a responsible and effective integration of AI in educational assessment can be fostered, ensuring that technological advancements align with the principles of equity, transparency, and the enhancement of educational practices.

# 2. Literature Review

As part of our research, we have performed two types of literature reviews: meta-synthesis and scoping review. As we delve deeper into the ethical implications of utilizing Artificial Intelligence (AI) in education, it becomes essential to categorize and understand the various domains within which AI is making an impact. From our literature review, we identify a systematic hierarchy for organizing and analyzing AI in education as a) AI in Teaching and Learning, b) AI in Assessment, and c) AI in Literacy Curriculum.

*AI in Teaching and Learning:* Various studies collectively emphasize the growing significance of AI integration in education. In their respective works, Roll and Wylie (2016) delve into the utilization of AI within the context of intelligent tutoring systems, while Sharples (2022) thoroughly examines the promising prospects of AI in enhancing the teaching of creative writing.

*Al in Assessment:* Recent research in education and assessment has emphasized the role of Al in assessment. Jia et al. (2022) advocate for distinguishing assessment from feedback in higher education, Baker et al. (2023) further this exploration in essay scoring, while Schneider (2023) applies Al to autograde short-answer questions, and Swauger and Kalir (2023) investigate Al's potential in detecting online exam cheating. These studies collectively highlight Al's transformative influence on education and assessment.

Al Literacy Curriculum: In recent academic discussions, Holmes (2023) underscores the significance of cultivating Al literacy among teachers to support their students' Al literacy development. Wang and Lester (2023) stress the importance of K-12 Al literacy curricula in preparing students for a changing job market, acknowledging the challenges and the need for evidence-based research in curriculum development. Rets et al. (2023) emphasize the value of interdisciplinary research teams in Al knowledge production, bridging the gap between machine learning and educational theories. Finally, Porayska-Pomsta (2023) calls for a broader understanding within the AIED community of Al's implications on human functioning, learning, and socio-cultural impacts, urging vigilance against potential risks and misuse of Al research and outcomes.

Now, we will discuss ethical concerns regarding a few studies.

# 2.1 *Existing AI-Based Methods for Educational Assessment:*

# 2.1.1 Bi-Directional Recurrent Knowledge Tracing Neural Network:

Delianidi and Diamantaras (2023) present an approach to assessment through the Bi-Directional Recurrent Knowledge Tracing Neural Network (KT-Bi-GRU). This model uses the recurrent neural networks (RNN) to predict student performance and trace their evolving knowledge states over time. The KT-Bi-GRU model incorporates a dynamic sub-network featuring a recurrent Bi-GRU layer for accurate knowledge state estimation and a non-dynamic, feed-forward sub-network for predicting answer correctness. Leveraging student interaction history collected through an Intelligent Tutoring System (ITS), the model

offers personalized educational recommendations. Notably, the model's architecture integrates embedding and convolutional layers for preprocessing input data, which holds immense potential for reshaping assessment practices (Delianidi and Diamantaras, 2023).

Ethical concerns:

- 1. *Transparency and Explainability:* Deep neural networks can be challenging to interpret, making it difficult to explain how decisions are reached. It is important to develop methods for explaining the predictions made by the model, especially when those predictions impact students' educational paths.
- 2. Feedback and Learning Impact: The use of AI in educational assessment could impact the way students receive feedback and learn. Depending on how the technology is used, it could potentially discourage risk-taking or discourage exploring beyond the scope of predicted outcomes.

#### 2.1.2 Data-Driven Methods for Feedback Generation:

Lu and Cutumisu (2021) explore data-driven methods for automated feedback generation, capitalizing on deep-learning algorithms to map student work to expert feedback. This approach includes models such as CNN, CNN+LSTM, and CNN+Bi-LSTM, which form an autograder system aimed at providing personalized feedback on students' essays. The autograder system utilizes AI to analyze essay content and structure, generating constructive feedback based on learned patterns from expert feedback. This AI-driven feedback mechanism streamlines the feedback process, offering timely and consistent support to students in enhancing their writing skills (Lu and Cutumisu, 2021).

Advantages :

- 1. Efficiency: Automated essay scoring can significantly speed up the assessment process, providing rapid feedback to students and teachers.
- 2. Consistent and Transparent Automated Essay Scoring Criteria

Ethical concerns:

- 1. *Bias in the automated scoring algorithm:* There is a potential for the algorithm to be biased, which could result in unfair treatment of certain groups of students. This could be due to factors such as the algorithm being trained on biased data or the algorithm not being able to accurately capture the nuances of certain types of writing.
- 2. Generic feedback: There is a risk that the automated feedback generated by the system could be too generic and not tailored to the individual needs of each student. This could limit their learning and growth, as they may not receive feedback that is specific to their strengths and weaknesses.
- 3. *Reduction in the number of human teachers and evaluators:* The use of AI technology in education could lead to a reduction in the number of human teachers and evaluators, which could have negative implications for employment and the quality of education. This could also lead to a lack of human interaction and support for students, which could impact their motivation and engagement.

#### 2.1.3 Automated Peer Review Evaluation:

Liu et al. (2023) delve into the realm of automated peer review evaluation, proposing a machine-based approach to assess peer assessments through natural language processing and machine learning techniques. The study addresses the challenge of limited labeled data by introducing a pseudo-labeling approach, which enhances the evaluation of peer reviews by identifying problem statements and suggestions in comments. By leveraging a semi-supervised learning method, the proposed approach offers a comprehensive

evaluation of peer assessments, thereby contributing to the enhancement of peer-based learning and assessment practices.

Ethical concerns:

- 1. One concern is the potential for biases and inaccuracies in the data used to train and evaluate AI models. As mentioned in the document (Liu et al., 2023), obtaining labeled data for predicting helpfulness can be challenging, especially when it involves subjective criteria. This raises concerns about the reliability and validity of the data used to train and evaluate models.
- Another concern is the potential for privacy and confidentiality violations when using student reviews as a source of labeled data. Students may not feel comfortable sharing their feedback with others, especially if they fear retaliation or negative consequences. This can lead to incomplete or biased data that may not accurately reflect the helpfulness of peer reviews.

#### 2.1.4 Automated Grading Systems:

Automated grading systems have gained prominence for their potential to expedite assessment processes. However, their limitations in understanding nuances and fact-checking content raise ethical and practical concerns. While these systems offer efficiency, they lack the depth of human understanding and critical analysis required for comprehensive assessment. Human instructors remain essential for providing meaningful feedback and evaluating complex student submissions accurately.

Ethical concerns:

- 1. Regarding security concerns, the paper states that appropriate security precautions must be taken to avoid unauthorized access to the system, particularly with regards to protecting the test program (section 4).
- 2. The paper acknowledges that the use of an automatic grading system may lead to a decrease in the quality of the programming style of students, as even inefficient solutions are accepted (section 3).
- 3. The paper also notes that the use of AI in education may make it easier for students to copy assignments, which goes against the aim of the grading system to promote self-responsibility (section 3).
- 4. The paper acknowledges that the impartiality of the system may lead to a lack of human interaction and feedback, which is an important aspect of education (section 1).
- 2.1.5 Automated Feedback Generation for Student Project Reports A Data Driven Approach:

The work of Jia et al. (2022) underscores the significance of instant feedback in promoting academic achievement and student success. Their focus lies in providing automated feedback for student project reports through the Insta-Reviewer system. This data-driven approach involves a two-step process: an unsupervised method summarizes the original reports to an appropriate length, followed by a supervised text-to-text generation model that produces plausible feedback. The paper emphasizes the need for greater attention to ethical considerations in feedback generation, particularly regarding potential concerns with text-generation methods. Issues such as generating improper or offensive language, as well as unintentionally revealing private information, are highlighted. To mitigate these concerns, the authors fine-tune models with domain-specific data and manually inspect generated feedback to ensure ethical compliance.

#### Ethical concerns:

1. The potential for generating improper or offensive language, which may appear in text generation.

- 2. Another concern arises from the use of pre-trained language models, which can inadvertently produce personally identifiable information due to the lack of careful filtering during the corpus collection process from the internet.
- 3. This unfiltered data can lead to models replicating harmful language or sensitive personal details acquired during pre-training.
- 4. Additionally, there is a lack of systematic methods to evaluate a system's effectiveness in avoiding the generation of inappropriate content.

These concerns are substantiated by references to studies by Celikyilmaz et al. (2020), Gehman et al. (2020), Bender et al. (2021), Li et al. (2022), and Malmi et al. (2022). The paper suggests that researchers in the field should remain vigilant and actively investigate potential ethical issues in text generation.

# 2.2 Various Assessment Methods in Education Utilizing Artificial Intelligence:

The study by Al Braiki et al. (2020) explores diverse assessment methods in education that leverage artificial intelligence. These methods encompass:

- *Multiple-Choice Questions:* Using the approach of computer-based grading and administration enables efficient and automated assessment.
- Automated Essay Scoring: AI tools facilitate subjective evaluations of essays. MIT's Enhanced AI Scoring Engine is an example of an automated system developed for this purpose.
- Calibrated Peer Review: This method involves peers evaluating and providing feedback on each other's work, guided by calibrated criteria.

# 3. Data Collection Methodology Used in Literature Reviewed:

This research paper adopts secondary data collection methodologies which involve using existing data collected by someone else for a purpose different from the original intent. We analyzed and interpreted this data to extract relevant information.

- a. *Published Sources:* In our research, we've extensively referenced published sources, such as books, academic journals, magazines, and newspapers to extract pertinent data and support our study's objectives.
- b. Online Databases: We have harnessed the capabilities of online databases to access a diverse range of secondary data, including research articles, statistical information, economic data, and social surveys, enriching our study with a comprehensive set of information.
- c. *Publicly Available Data:* Leveraging publicly available data shared on various platforms, websites, and social media channels, we have gained unique insights that have added depth to our research by considering diverse perspectives.
- d. *Past Research Studies:* The foundation of our research rested upon thorough examination and analysis of past research studies. By building upon the methodologies and insights of these studies, we have advanced the field by contributing novel perspectives and interpretations.

# 4. Cultivating Personalized Learning and Data-Driven Insights:

The integration of artificial intelligence (AI) and learning analytics with game-based learning holds immense potential for education. Students benefit from personalized learning experiences as AI analyzes their game interactions, adapting content and offering real-time

feedback. This fosters critical skills alongside knowledge absorption. Teachers gain insights from Al-driven analytics, identifying student trends and tailoring instruction. Administrative tasks can be automated, freeing up time for individualized mentoring. Al tools like sentiment analysis provide emotional insights, aiding timely support. Adaptive testing adjusts question difficulty, reducing anxiety and enhancing motivation. Overall, this synergy empowers students, equips teachers, and reshapes education into a dynamic and personalized journey.

# 4.1.1 The Nomads:

The integration of a mathematics educational game, "The Nomads," with a continuous conjunctive model (CCM) for stealth assessment showcases the potential of game-based learning analytics in transforming mathematics education. The game's adaptive expertise training and dynamic problem-solving tasks align with modern pedagogical needs, addressing limitations of traditional classroom teaching. The integration highlights the transformative role of digital game-based learning (DGBL), enhancing student engagement, self-confidence, and mathematical skills. Al-driven analytics, including Bayesian networks, enable real-time assessment, personalized feedback, and skill acquisition tracking. While this study offers valuable insights into student performance and learning progress through gameplay, it also underscores the need for larger and more diverse sample sizes, as well as refinement of assessment models. The fusion of game-based learning, learning analytics, and Al tools holds great promise for reshaping education, empowering both educators and learners to unlock new dimensions of effective and engaging mathematics instruction.

# 4.2 Serious Game Interaction With Medical Emergency Scenario:

Focused on a medical emergency simulation game, the research demonstrates how player interactions can be harnessed to improve educational game design and assessment. By analyzing player behaviors and engagement, the study unveils valuable insights into learning progression and the efficacy of the game's content. The evidence-based assessment approach employed in this study serves as a powerful tool for educators, enabling them to gauge student learning outcomes based on game interactions. This approach holds immense potential for personalized and evidence-driven teaching strategies, facilitating more effective educational experiences.

# 4.3 Maximize Students Sense Of Choice In Elective Subjects:

This study highlights the significant impact of maximizing students' sense of choice in elective subjects on school enrollment and academic performance. By employing the optimized "simulated annealing" algorithm, schools and colleges can efficiently allocate students to their preferred elective subjects, enhancing student engagement and satisfaction. The research underscores the evolving nature of student preferences and the need for educational institutions to adapt accordingly. Through the innovative algorithmic approach, the administrative burden of subject allocation is dramatically reduced, streamlining the process to a mere fraction of the previous time. Ultimately, the implementation of this tool holds the potential to foster a more student-centric learning environment, empowering both educators and students by aligning elective subject choices more closely with individual preferences and aspirations.

# 4.4 Teaching Learning Analytics:

The technique introduces the Analytics Model for Teacher Inquiry (AMTI), emphasizing the integration of Teaching Learning Analytics (TLA) to enhance teaching practices. By bridging classroom interactions and data analysis, TLA offers teachers valuable insights into student engagement, interaction, and outcomes. The study underscores the need for teachers to ask pertinent questions, analyze data, and link insights to classroom practices. TLA holds the

potential to empower educators with data-informed strategies, fostering dynamic learning environments that improve teaching techniques and enhance student-centered learning experiences. Despite obstacles like workload and familiarity with new tools, the study emphasizes the ethical considerations of data collection and the benefits of incorporating TLA into regular teaching practices.

# 4.5 Track Students' Distress During Educational Gameplay:

This assessment technique employs multimodal data fusion to track students' distress during educational gameplay, using techniques like facial expression analysis and gameplay performance. The integration of various data streams enhances the prediction of cognitive-affective states, aiding adaptive support and learning analytics. While focused on the game "Zoombinis," the approach has broader implications for understanding student emotions and engagement. This technique offers educators insights to tailor interventions and curriculum, ultimately creating more effective and supportive learning environments. Further research, including speech data analysis, could extend the method's applicability, benefiting both teachers and students.

# 5. Ethical Dimensions in AIED Assessment

The infusion of AIED into assessment introduces diverse ethical dimensions that demand meticulous examination:

#### 5.1. Privacy:

The accumulation and analysis of personal data by AIED systems evoke apprehensions about student privacy (Selwyn, 2020). Upholding transparent data usage practices and securing informed consent from students are ethical imperatives within this context. Students express concerns about the use of their learning data, emphasizing the need for safeguarding their personal information. Stakeholders, including administrators and diversity leaders, identify concerns with data collection, access, and regulations, underscoring the importance of responsible data handling. The potential for biased decisions based on learner data analysis is acknowledged, raising privacy issues. Moreover, there is a lack of consistent definitions and conceptualizations of privacy and related concepts like ethics and transparency within the context of learning analytics. To promote privacy and autonomy, efforts are made to inform students about data flows and provide them with informed consent regarding data usage.

#### 5.2. Fairness:

AIED systems possess the capacity to perpetuate biases present within the data they learn from (Diakopoulos & Friedler, 2019). Striving for fairness and equity in assessment outcomes mandates continuous monitoring and adjustment of algorithms to mitigate bias. Addressing fairness in learning analytics is a central theme across the literature. Stakeholders emphasize the need to counter biased decisions stemming from learner data analysis. Efforts are made to ensure equity in learning analytics, with recommendations to incorporate diverse perspectives and inclusive practices. Stakeholders also express concerns about the limitations of data and its potential to reinforce inequalities. Fairness is underscored through stakeholder accountability and collaboration in data collection and analysis, aiming to create a more equitable learning environment for all.

# 5.3. Transparency:

The opacity of AI algorithms can hinder learners' comprehension of assessment outcomes (Hill et al., 2019). Ethical utilization of AIED necessitates clear explanations of the decision-making process, enabling learners to foster trust and engagement with the technology. Transparency is a key consideration in responsible learning analytics practices. Ethical reflections within the learning analytics community emphasize the importance of trust and transparency. Stakeholders, including administrators and diversity leaders, play crucial roles in ensuring transparent practices. However, privacy-related concepts like ethics and transparency are often not consistently defined or conceptualized. The literature stresses the significance of informing students about data flows and promoting transparency in the use of their information to build a more transparent and accountable learning analytics ecosystem.

#### 5.4. Autonomy:

Balancing personalized learning facilitated by AIED with learners' autonomy poses a nuanced challenge (VanLehn, 2018). Educators must contemplate the ways in which AIED influences learners' capacity to make informed decisions about their learning pathways. Autonomy emerges as a significant principle in the context of learning analytics. Informed consent is a crucial aspect of autonomy, ensuring that students are well-informed about how their data will be used. Privacy is promoted as a means to support individuals' self-determination, allowing them to control their personal information. Contextual differences in autonomy among various stakeholders are recognized, and efforts are made to engage students in discussions about data privacy and literacy. Ultimately, promoting autonomy in learning analytics involves respecting individuals' control over their data and fostering a sense of ownership and empowerment.

# Conclusion:

In the realm of modern education, the integration of Artificial Intelligence (AI) in assessment has ignited both promise and ethical complexities. This research has delved into these intricacies, revealing a landscape where fairness battles bias, transparency tussles with opacity, and privacy intertwines with data sharing.

From educators to policymakers, the chorus of perspectives underscores the ethical tightrope that AI treads in education. Amidst the potential efficiencies, there is an ethical imperative to ensure AI-driven assessments uphold human values, provide clear explanations, and respect data privacy.

As we steer education toward Al-infused frontiers, this research attempts to serve as a guiding light, urging an ethical framework that harmonizes technological advancement with human-centric educational principles. The compass points to a future where Al enhances learning without eclipsing the core values of fairness, transparency, and equity.

In conclusion, the use of Artificial Intelligence (AI) in education brings both exciting possibilities and important ethical concerns, especially in assessment. We have explored how AI can help evaluate students in new ways, from predicting knowledge to giving personalized feedback. While these advances can make assessments more efficient and accurate, we must be careful about fairness, privacy, and bias. Our study looked at various Al methods for assessment, like grading multiple-choice questions and scoring essays automatically. These methods can save time and offer insights, but they also need to be transparent and unbiased. Through surveys, interviews, and analyzing existing research, we learned that many people worry about how AI might affect education. Teachers, students, and policymakers all have different views. Also the students worry about their data being used in ways they don't understand, so being open and transparent is really important. Some see AI as a useful tool, while others are concerned it might replace human judgment. In the end, finding the right balance between AI and human involvement is crucial. While AI can help us in many ways, it's important to remember that education is not just about numbers and algorithms. Human understanding and ethical values play a big role in assessing students' learning. We need to use AI responsibly, making sure it respects fairness, privacy,

and the well-being of students. By doing so, we can create a future where AI enhances education while keeping ethical values intact.

#### **References:**

- Al Braiki, B., Harous, S., Zaki, N., & Alnajjar, F. (2020). Artificial intelligence in education and assessment methods. Bulletin of Electrical Engineering and Informatics, 9(5), 1998-2007. https://www.beei.org/index.php/EEI/article/view/1984
- Delianidi, Marina, & Diamantaras, Konstantinos. (2023). KT-Bi-GRU: Student Performance Prediction with a Bi-Directional Recurrent Knowledge Tracing Neural Network. Journal of Educational Data Mining, 15(2), 1–21. https://doi.org/10.5281/zenodo.7808087
- Jia, Qinjin, Young, Mitchell, Xiao, Yunkai, Cui, Jialin, Liu, Chengyuan, Rashid, Parvez, & Gehringer, Edward. (2022). Automated Feedback Generation for Student Project Reports: A Data-Driven Approach. Journal of Educational Data Mining, 14(3), 132–161. https://doi.org/10.5281/zenodo.7304954
- Liu, Chengyuan, Doshi, Divyang, Shang, Ruixuan, Cui, Jialin, Jia, Qinjin, & Gehringer, Edward. (2023). Less But Enough: Evaluation of peer reviews through pseudo-labeling with less annotated data. Journal of Educational Data Mining, 15(2), 123–140. https://doi.org/10.5281/zenodo.7304981
- Liu, C., Doshi, D., Bhargava, M., Shang, R., Cui, J., Xu, D., & Gehringer, E. (2023, July). Labels are not necessary: Assessing peer-review helpfulness using domain adaptation based on self-training. In Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023) (pp. 173-183).
- Lu, C., & Cutumisu, M. (2021). Integrating Deep Learning into an Automated Feedback Generation System for Automated Essay Scoring. International Educational Data Mining Society.
- Porayska-Pomsta, K. (2023). A Manifesto for a Pro-Actively Responsible AI in Education. International Journal of Artificial Intelligence in Education, 1-11. https://doi.org/10.1007/s40593-023-00346-1
- Rets, I., Herodotou, C., & Gillespie, A. (2023). Six Practical Recommendations Enabling Ethical Use of Predictive Learning Analytics in Distance Education. Journal of Learning Analytics, 10(1), 149-167. https://doi.org/10.18608/jla.2023.7743
- Ryan S. Baker, Lief Esbenshade, Jonathan Vitale, & Shamya Karumbaiah. (2023). Using Demographic Data as Predictor Variables: A Questionable Choice. Journal of Educational Data Mining, 15(2), 22–52. https://doi.org/10.5281/zenodo.7702628
- Roll, I., Wylie, R. Evolution and Revolution in Artificial Intelligence in Education. International Journal of Artificial Intelligence in Education, 26, 582–599 (2016). https://doi.org/10.1007/s40593-016-0110-3
- Schneider, J., Richner, R. & Riser, M. Towards Trustworthy AutoGrading of Short, Multi-lingual, Multi-type Answers. International Journal of Artificial Intelligence in Education, 33, 88–118 (2023). https://doi.org/10.1007/s40593-022-00289-z
- Sharples, M. (2022). Automated essay writing: An AIED opinion. International journal of artificial intelligence in education, 32(4), 1119-1126. https://doi.org/10.1007/s40593-022-00300-7
- Swauger, S., & Kalir, R. (2023). Learning Analytics and the Abolitionist Imagination. Journal of Learning Analytics, 10(1), 101-112. https://doi.org/10.18608/jla.2023.7813
- Wang, N., & Lester, J. (2023). K-12 Education in the Age of Al: A Call to Action for K-12 Al Literacy. International journal of artificial intelligence in education, 1-5. https://doi.org/10.1007/s40593-023-00358-x

# Fostering Ethics in AI: Perceptions from the Indian AI Curriculum

#### Ashutosh RAINA<sup>a\*</sup>, Kushal MUNDRA<sup>b</sup>, Prajish PRASAD<sup>b</sup> & Shitanshu MISHRA C<sup>c</sup>

<sup>a</sup> Member, Ed Tech Society, India
<sup>b</sup> Flame University, India
<sup>c</sup> UNESCO MGIEP. India

#### \*raina.ashu@gmail.com

The views and content presented in this paper are the sole responsibility of the individual authors and do not represent the views or opinions of the respective organization.

Abstract: In the age of rapid Artificial Intelligence (AI) advancement, universities worldwide have responded by offering specialized AI and Machine Learning (ML) courses to meet industry demands. However, amidst this surge in AI education, AI deployment's ethical and societal implications often need more attention. Unlike traditional algorithmic programming, AI involves intricate decision-making processes that are challenging to predict or explain, demanding a comprehensive understanding of its ethical dimensions. Various global initiatives have highlighted the ethical considerations surrounding AI, resulting in an increased emphasis on integrating ethics education into AI curricula. Despite this, there remains a gap in addressing these vital aspects across the broader AI education landscape, with ethics often relegated to the periphery of computer science courses. In this context, this paper explores the imperative for ethics courses in undergraduate AI education in India. We examine educators ' awareness and perceptions regarding ethics education in AI curricula through interactions with faculty members participating in professional development programs and workshops focused on AI pedagogy. Our analysis reveals that there is a pressing need to extend discussions on ethics beyond mere privacy concerns and traditional performance metrics, integrating real-life scenarios into the curriculum. This paper serves as a preliminary step towards a need for a general framework in structuring ethics education in AI in India, aiming to initiate a more comprehensive and standardized approach to AI ethics education, fostering responsible AI development in the future.

Keywords: AI Education, Ethics, Teacher perception, Active Learning

#### 1. Introduction

The rapid evolution of Artificial Intelligence (AI) has led to significant advancements in various industries, transforming how businesses operate and enabling previously unimaginable innovations. As AI technologies continue to penetrate different sectors, there is a growing demand for a skilled workforce capable of harnessing the power of AI to drive innovation. In response, universities worldwide have begun offering specialized courses and programs in AI and Machine Learning (ML) to equip students with the necessary skills to meet industry demands.

However, amidst this surge in AI education, a critical aspect often receives inadequate attention – the ethical and societal implications of AI deployment (Jobin et al., 2019). Unlike traditional algorithmic programming solutions, AI operates in a distinct manner, often involving complex decision-making processes that are challenging to predict or explain. This unique nature of AI necessitates a comprehensive understanding of its ethical dimensions to ensure responsible and beneficial deployment (Buiten et al., 2019). Recognizing this imperative, various global efforts have raised concerns about the ethical considerations surrounding AI (Hagendorff, 2020). Consequently, the integration of ethics education into the AI curriculum has gained prominence. While some universities have introduced topics and even entire courses dedicated to AI ethics, more attention must be paid to these essential aspects in the broader AI education landscape. Ethics units are frequently relegated to the concluding segments of computer science courses, indicating their lower priority than technical topics. Even within pertinent disciplines like machine learning, where ethical considerations are profoundly relevant, the standard integration practice remains limited (Garrett et al., 2020).

Several noteworthy examples, however, emphasize the significance of addressing AI ethics within educational frameworks (Khan et al., 2022). Initiatives attempting to map AI ethics teaching practices often employ human-driven approaches, where topics of interest are identified and categorized based on instructor descriptions or open coding (Garrett et al., 2020). Despite such endeavors, a comprehensive and standardized approach to integrating AI ethics across curricula still needs to be discovered (Javed et al., 2022).

This paper examines the necessity for incorporating ethics courses into undergraduate Al education, particularly in the context of India. We explore this need by engaging with participants in faculty development programs and workshops dedicated to Al instruction. Through these interactions, we aim to gain insights into the awareness and perspectives of educators within the Indian context, considering the region's socio-cultural diversity and unique ethical beliefs. Our goal is to understand how ethics education can be effectively integrated into Al curricula. By analyzing these perspectives, we contribute to the ongoing dialogue on promoting responsible Al development, advocating for a more comprehensive and standardized approach to Al ethics education. Although this paper does not delve into pedagogy, it serves as an initial step in guiding instructors. We identify the ethical topics currently addressed in Al courses, providing insights into the existing state of Al ethics education and raising pertinent questions. Our intention is to catalyze further research that is tailored specifically to the Indian educational system, ultimately establishing standardized practices for ethics in Al education.

# 2. Ethics in Al

# 2.1 Global Discussion

The current state of literature concerning the inclusion of ethics in AI-based courses reflects the heightened awareness of the ethical dimensions of artificial intelligence (Khan et al., 2022). Numerous institutions and organizations have emphasized the importance of ethical principles in AI, focusing on values such as Transparency, Privacy, Accountability, Fairness, Autonomy, and Explainability (Florida et al., 2018). Several committees and organizations have also introduced their frameworks and guidelines for the ethical implementation of AI technologies. (Floridi et al., 2018 & Jobin et. al., 2019)

However, despite the widespread acknowledgement of these principles and the proliferation of ethical frameworks, a significant challenge still needs to be solved - effectively integrating these ethical considerations into training materials, especially within the courses, to prepare the future AI workforce. The translation of abstract ethical principles into tangible, actionable content for AI education is an ongoing challenge.

Furthermore, there is a recurring concern within the literature about the perceived ambiguity and generality of ethical principles outlined by many organizations (Eitel, 2021). This ambiguity can lead to varied interpretations and applications of these principles, making it challenging to establish a standardized approach to ethical education in AI. Additionally, it has come to light that technology firms involved in AI and autonomous systems development often follow ethical guidelines based on their interpretations and priorities (Saheb, 2023). This further underscores the need for a comprehensive and standardized approach to ethics education in AI courses to ensure that future AI professionals are well-equipped to navigate the complex ethical landscape of the field. As we delve deeper into the evolving discourse on ethics in AI, addressing these challenges and fostering a more consistent and robust ethical foundation for AI education remains a paramount concern.

# 2.2 Ethics in Global Curriculum

Al has made its way into school and college curricula, but the absence of a standardized framework has resulted in diverse implementations (Miao et al., 2022). Instructors often interpret and incorporate Al differently, leading to variations in curriculum content and teaching methods (Javed et al., 2022).

The strategies to achieve ethics education objectives remain subjects of ongoing deliberation (Eitel, 2021). Ethical components are mandated in Computer Science (CS) programs in the United States to secure accreditation. However, universities and educators are responsible for deciding how to execute this mandate (Javed et al., 2022). The advantages of ethics education are evident; discipline-specific ethics instruction has been shown to foster ethical growth and imbue students with the conviction that ethics is intrinsic to their profession, transcending mere public relations considerations (Knight, 2014 & Loescher, 2004). Furthermore, the field of CS has long advocated for a more comprehensive integration of ethics throughout the curriculum, including technical courses (Martin et al., 1996). Despite some progress in this direction, it has yet to become standard practice (Fiesler, 2020 & Saltz, 2019).

Additional queries arise regarding the methods of imparting ethics education, stemming from the notion that an ethics curriculum is optimal in breaking away from isolation and giving it a structured form (Eaton, 2018). Furthermore, proposals suggest incorporating project-based learning and utilizing science fiction as a tool for speculative exploration (Burton, 2015). Researchers have additionally pinpointed that global curricula demonstrate deficiencies in disciplinary breadth. They frequently exhibit limited geographical coverage, a bias toward Western cultures, and a tendency to offer courses primarily at just one academic level (Saltz et al., 2019).

This paper endeavors to initiate a discourse on integrating ethics into India's undergraduate (UG) curriculum for Artificial Intelligence (AI). This objective is achieved through interactions with UG computer science instructors and examining AI-based curricula to assess the inclusion of ethical components.

# 3. Ethics in Al Curriculum in India

Within educational research, scrutinizing syllabi and soliciting instructors' viewpoints on the subject matter is a customary approach to evaluating curricular requisites (Chong, 2016). This methodology has been frequently employed to offer insights into pedagogical strategies within the field of computer science (Becker, 2019). Assessing a syllabus is an effective means of ascertaining the components of knowledge encompassed within a course (Becker, 2019). Prior research has effectively applied this methodology to delineate the content and objectives of general technology ethics courses (Fiesler, 2020).

In this study, our focus centers on the domain of AI ethics within the context of AI courses and dedicated ethics-oriented courses. To steer our analysis, our principal research query was formulated as follows: What themes pertaining to Ethics in AI (Principles) are encompassed within AI-based courses offered at a selection of prominent universities in India? What is the comprehension and viewpoint of undergraduate (UG) instructors from various Indian universities regarding the instruction of ethics in AI?

Our dataset was bifurcated into two segments for analysis. We gathered instructors' perspectives on Ethics in AI during AI education workshops. Their responses were systematically collected through the utilization of Mentimeter activities. Simultaneously, the

scrutiny of syllabi was conducted by mining online resources for syllabi related to Al education within the top 100 undergraduate institutions in India specializing in engineering and technology.

# 3.1 Instructors' Perspective on Ethics in AI

Sixty responses were collected from various instructors across three AI Education workshops. A specific segment of these workshops was dedicated to engaging instructors in discussions regarding their perspectives on the necessity of ethics in AI education. These interactions were conducted using open-ended questions through the utilization of Mentimeter. Mentimeter is an interactive presentation and polling tool that enables real-time audience engagement and feedback during presentations and events (Mentimeter, 2023). The responses were recorded and later presented to the participants to guide subsequent discussions.

The questions posed during these interactions were divided into three sections. Initially, instructors were asked about their teaching experiences in AI. Following this, their viewpoints on *'What ethical challenges they encountered or perceived in the design and implementation of AI-based solutions*' were sought. Subsequently, a scenario-based discussion was conducted, requiring instructors to design a lunch recommendation system for school children while considering the perspectives of various stakeholders, including management, students, parents, and doctors (Payne, 2019). Following an exercise involving an ethics matrix (Payne, 2019), the instructors were once again queried regarding their perspectives on the integration of ethics within the AI curriculum. Figure 1 illustrates a selection of responses recorded through Mentimeter.

For analysis, the responses were categorized into challenges related to data quality or responses and ethical and moral perspectives. This categorization followed coding based on associations derived from the existing literature (Khan et al., 2022).

	Join at menticom use code 2235 6775	ad Mentin		Join at menti.com use code 3216 9067	Monti
Have you taugł	nt AI / ML /Soft Comp	outing?	What ethical challenges do you face or perceive in the design and use of Al based solutions ? 42 Responses		
			Am i talking to a human or machine	Privacy	Privacy
3% No P		Is it the correct gender/authenticity	Concern about data usage and data sharing	Originality	
	ex Partially	Privacy	Privacy	Privacy	
		Bias,	Cost	Validity and Reliability of information shared via AI system	
		Fake news(deepfakes of politicians),	Rahietaee	Privacy of informative use of	

Figure 1: Mentimeter response recorded during one of the sessions.

Based on the data, it was observed that 90% (54) of the instructors had taught a course in artificial intelligence. The perceived ethical challenges were evenly distributed among ethical and moral aspects, with data quality challenges encompassing privacy as the most frequently reported concern in the former category and issues of reliability, authenticity, and validity in the latter category, as indicated in Figure 2. Following the discussion session and the activity involving the design of a learning exercise for teaching ethics in Al using an ethics matrix, participants were once again invited to share their perspectives. It was observed that some of the earlier responses had been consolidated, reflecting a more balanced distribution of concerns related to privacy, explainability, fairness, and reliability.



Figure 2: Participant perspectives on ethics in AI before and after the ethics session.

During focused group interactions with the participants, a recurring theme emerged regarding the need for a structured framework to guide the development of teaching materials and activities related to ethics. As relevant topics arose, these issues were raised sporadically without a dedicated session or structured approach. Furthermore, participants noted that using an ethics matrix with practical examples, such as designing recommender systems, helped them streamline their thought processes. Such a framework would facilitate the creation of activities that enable learners to engage with ethical considerations while designing Al-based solutions.

As discussed in previous sections, the absence of a universal framework has been identified as a challenge. Most of the participants in this study hailed from institutions where course curricula are structured at the university level. Thus, we embarked on the second phase of this research to investigate whether any structured methods or frameworks, either as individual courses or integrated into broader curricula, were being employed in teaching ethics in AI within Indian academia. In this phase, we explored universities and autonomous institutions where instructors had the autonomy to shape course curricula.

# 3.2 Ethics in AI in UG curriculum.

In this study, we compiled a list of the top 100 institutions in India specializing in engineering and technology by amalgamating data from the QS World Ranking and India's national equivalent ranking system. Subsequently, we systematically crawled the webpages of each of these institutions to amass comprehensive information regarding courses related to Artificial Intelligence (AI), Machine Learning (ML), Deep Learning, Soft Computing, and Ethics in AI. We meticulously recorded both the names and contents of these courses. Additionally, we scrutinized the course details to identify the utilization of terms commonly associated with ethics frameworks, as outlined by (Khan et al., 2022). It is crucial to acknowledge that due to disparities in website structures among institutions, a substantial amount of manual data extraction was requisite, and we continue to engage in ongoing efforts to maintain this data collection process. The forthcoming analysis encompasses findings from a 25% subset of institutions, selected randomly from public and private institutions for web crawling. Notably, our dataset exclusively comprises autonomous universities, and colleges affiliated with universities were excluded if the parent university was already represented in our dataset.



Figure 3: Inclusion of Ethics in AI for Institutions that offer AI-based courses.

Based on our initial data analysis, as shown in Figure 3, it is evident that 44% of the institutions offering AI-related curricula do not explicitly mention ethics in AI within their course descriptions. On the other hand, 40% of the institutions refer to ethics and allocate one or two sessions for discussions on ethics, while 6% incorporate ethics into the course title. Upon examining the course curricula, it becomes apparent that most courses within the 40% and 6% categories predominantly emphasize ethics within the context of engineering design without explicitly addressing the ethical principles relevant to AI. These courses touch upon terms like privacy, reliability, accuracy, and robustness, but often in assessing solution efficiency or comparing different algorithms. In some instances, there appears to be a tentative shift toward discussing the real-world implications of accuracy and robustness within specific scenarios. Notably, one elective course, "Ethics in AI", explicitly sets an objective to "Apply ethical principles and commit to professional ethics and responsibilities and norms of engineering practice." Only two courses delve into the ramifications of ethical practices in the use of AI, particularly about policy, with one of these being specific to AI in education.

Based on this preliminary analysis, it is apparent that while most courses underscore the importance of robustness, accuracy, and reliability as performance metrics, a significant portion confines discussions of ethical considerations in AI to professional ethics in engineering. This underscores the pressing need for a standardized framework, even at the national level, to guide the incorporation of ethics in AI education. Only two-course descriptions allude to a structured framework encompassing privacy, explainability, robustness, and accuracy.

# 4. Discussion and Conclusion

Based on the preceding analysis, there is an urgent necessity to expand the discourse surrounding privacy and other efficiency parameters beyond the conventional performance metrics of AI. It is imperative to explore their ramifications within the context of the contemporary world, particularly within real-life scenarios. Furthermore, providing learners with opportunities to grapple with ethical dilemmas by extending these discussions across multiple perspectives, as exemplified in the school recommender system workshop, is essential. Privacy garners extensive attention, a trend reflected in the responses obtained from participants in our pre-workshop perception survey. However, a perceptible shift in participants' perspectives occurred following discussions and activities utilizing the ethics matrix within real-life scenarios. Their concerns extended to encompass explainability, robustness, accuracy, and more. Our observations suggest that implementing a

standardized framework, in our case, the ethics matrix, enables participants to systematize their thoughts concerning ethical concerns.

Additionally, it facilitates their ability to empathize with the predicament of selecting the appropriate perspective for fine-tuning algorithms. In such cases, it prompts contemplation on the steps required to enhance the transparency of the algorithmic processes and tuning parameters, allowing users to make informed decisions regarding reliance on the output and its extent. Intriguingly, we noted substantial discussions on ethics in AI within Indian school curricula, predominantly centered on the four pillars of ethics: privacy, explainability, robustness, and accuracy (Miao et. al. 2022). However, this emphasis must be noticed in undergraduate (UG) curricula.

Our data collection process is ongoing as we engage with instructors to gain insights into their perspectives on ethics in AI across various institutions. Further in-depth analysis, coupled with exploring questions such as the impact of university-based curricula and the variability among curricula in autonomous institutions, will significantly contribute to our understanding of and structuring ethics education in AI. Such analysis can facilitate the development of a unified framework for teaching and learning ethics in AI, grounded in real-life contexts and experiential learning.

#### References

- Becker, B. A., & Fitzpatrick, T. (2019, February). What do cs1 syllabi reveal about our expectations of introductory programming students? In *Proceedings of the 50th ACM technical symposium on computer science education* (pp. 1011-1017).
- Buiten, M. C. (2019). Towards intelligent regulation of artificial intelligence. *European Journal of Risk Regulation*, *10*(1), 41-59.
- Burton, E., Goldsmith, J., & Mattei, N. (2015, January). Teaching AI Ethics Using Science Fiction. In *Aaai workshop: Ai and ethics*.
- Chong, F. (2016). The pedagogy of usability: An analysis of technical communication textbooks, anthologies, and course syllabi and descriptions. *Technical Communication Quarterly*, *25*(1), 12-28.
- Eaton, E., Koenig, S., Schulz, C., Maurelli, F., Lee, J., Eckroth, J., ... & Williams, T. (2018). Blue sky ideas in artificial intelligence education from the EAAI 2017 new and future AI educator program. *AI Matters*, *3*(4), 23-31.
- Eitel-Porter, R. (2021). Beyond the promise: implementing ethical AI. AI and Ethics, 1, 73-80.
- Fiesler, C., Garrett, N., & Beard, N. (2020, February). What do we teach when we teach tech ethics? A syllabi analysis. In *Proceedings of the 51st ACM technical symposium on computer science education* (pp. 289-295).
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2021). An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Ethics, governance, and policies in artificial intelligence*, 19-39.
- Garrett, N., Beard, N., & Fiesler, C. (2020, February). More than" If Time Allows", the role of ethics in Al education. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 272-278).
- Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and machines*, *30*(1), 99-120.
- Mentimeter (2023), Interactive presentation software. Mentimeter. https://www.mentimeter.com
- Javed, R. T., Nasir, O., Borit, M., Vanhée, L., Zea, E., Gupta, S., ... & Qadir, J. (2022). Get out of the BAG! Silos in AI ethics education: an unsupervised topic modelling analysis of global AI curricula. *Journal of Artificial Intelligence Research*, *73*, 933-965.
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature machine intelligence*, *1*(9), 389-399.
- Khan, A. A., Badshah, S., Liang, P., Waseem, M., Khan, B., Ahmad, A., ... & Akbar, M. A. (2022, June). Ethics of Al: A systematic literature review of principles and challenges. In *Proceedings of the 26th International Conference on Evaluation and Assessment in Software Engineering* (pp. 383-392).
- Knight, H. (2014). *How humans respond to robots: Building public policy through good design.* Brookings.
- Loescher, K. J. (2004). The impact of an "Ethics across the curriculum" initiative on the cognitive moral development of business school undergraduates. Nova Southeastern University.

Martin, C. D., Huff, C., Gotterbarn, D., & Miller, K. (1996). Implementing a tenth strand in the CS curriculum. *Communications of the ACM*, *39*(12), 75-84.

Miao, F., & Shiohira, K. (2022). K-12 AI curricula. A mapping of government-endorsed AI curricula.

- Payne, B. H. (2019). Ethics of artificial intelligence curriculum for middle school students. *MIT Media* Lab Personal Robots Group. Retrieved Oct 10, 2019.
- Saheb, T., & Saheb, T. (2023). Topical review of artificial intelligence national policies: A mixed method analysis. *Technology in Society*, 102316.
- Saltz, J., Skirpan, M., Fiesler, C., Gorelick, M., Yeh, T., Heckman, R., ... & Beard, N. (2019). Integrating ethics within machine learning courses. *ACM Transactions on Computing Education* (*TOCE*), 19(4), 1-26.

# ChatGPT in Education: Risks to Fairness of Access

Sumitra SADHUKHAN<sup>a\*</sup>, Shitanshu MISHRA<sup>b#</sup>, Sridhar IYER<sup>a</sup> <sup>a</sup> Indian Institute of Technology, Bombay, India <sup>b</sup>UNESCO MGIEP, India <sup>\*</sup>sumitra.sadhukhan@iitb.ac.in

#The views and content presented in this paper are the sole responsibility of the individual author and do not represent the views or opinions of the respective organization.

**Abstract:** In the rapidly changing ChatGPT world, human skills training is the need of the hour. With a large amount of data, ChatGPT is training itself day by day to understand human prompts and to provide appropriate results. A lot of prompt engineering courses educate learners in structuring the prompts based on how the ChatGPT engine processes any prompt and response. However, these courses do not target specific cognitive skills required by users (learners) to unfold, understand, and express what they really need and want to know from ChatGPT. In this paper, we look at ChatGPT as a tool to support knowledge acquisition and we discuss the questioning skill as an essential cognitive skill required to interact with and make optimal use of ChatGPT capabilities. We argue that risk to the fairness of access does not just stem from the monetary availability of such technologies, but it can also arise from the unpreparedness of the target users. The AI and ethics of AI literacy communities should also focus on training individuals on the cognitive skills needed to become optimal users of such technologies.

Keywords: ChatGPT, Cognitive skill, Question Posing, Prompt Engineering

#### 1. Introduction

The advancement of AI-powered chatbot, ChatGPT has opened a new door to education. It is trained to perform complex tasks and give answers like humans. In education, ChatGPT is a more efficient tool compared to traditional search engines as it searches possible multiple sources and offers a written answer in a human-understandable form instead of only providing a list of multiple sources (Cascella et al., 2023). It can help students to access fine-grained information in a simplified way using prompts. ChatGPT can be used by learners to solve tasks such as writing code and explaining code and can be used to scaffold instructors in creating personalized programming exercises in computing education (Prather et al., 2023). ChatGPT has the potential to provide personalized support and feedback to students at different levels of complexity (Farrokhnia et al., 2023). It also has shown the capability to stimulate critical thinking among students by providing a set of questions designed based on every student's knowledge level and preferences to challenge them (Cotton et al., 2023).

However, a large array of articles has also discussed the limitations and drawbacks of ChatGPT. Gao et al. (2023), emphasize on lack of thorough comprehension of the meaning of the words that ChatGPT processes. Although it can spot patterns and arrive at reasonable solutions, it does not fully understand the meaning of the words (Bogost, 2022). This could lead to responses that occasionally lack depth and insight (Borji, 2023). Particularly for performing tasks that require a nuanced understanding of specific domain knowledge (Dimitrov, 2023). As pointed out by Zhong et al. (2023), while exploring the understanding ability of ChatGPT, it was observed that ChatGPT may generate some contradictory or unreasonable responses. They proposed that it can be overcome by advanced prompting strategies, i.e., the manual few-shot or chain of thought (CoT), which provides manual intermediate reasoning steps. Zheng (2023), tried to understand the failures of ChatGPT in complex open-domain question-answering

sessions. They found that the model gives general answers, not specific ones, and found the problem of ChatGPT in knowledge memorization, recall, and reasoning. So, they recommended providing background information, and external knowledge as specific as possible and decomposing complex problems into subproblems.

Loconte et al. (2023), showed that a large language model like ChatGPT lacks cognitive ability that needs to integrate with human prefrontal lobes, known as "prefrontal functions". They investigated ChatGPT's intelligence power using the same test used to evaluate prefrontal functioning in humans. They found poor planning abilities and difficulty in understanding others' intentions and mental states. Hence ChatGPT lacks such demonstrable cognitive abilities.

#### 2. User Competencies needed for ChatGPT

#### 2.1 Prompt Engineering Skill

Prompt engineering allows learners to communicate effectively with ChatGPT by providing specific refined prompts. There exist different courses that aim at improving learner's prompt engineering skills. These courses center on instructing prompt patterns and structuring sentence components to address specific challenges. The courses primarily aim to enhance learners' understanding of Natural Language Generation (more specifically GPT) systems, such that the learners can determine the refinements of prompt patterns before engaging with ChatGPT. The process of refining prompt patterns involves selecting appropriate words while defining the prompt's focus, scope, and boundary. White et al. (2023), introduced a comprehensive catalog of prompt engineering techniques, classified into six pattern categories: input semantics, output customization, error identification, prompt improvement, interaction, and context control. The input semantics category focuses on understanding human prompts to generate relevant output, serving particularly well when faced with ill-structured prompts. Within the prompt improvement category, the use of question refinement prompt patterns allows ChatGPT to produce improved versions of user prompts.

#### 2.2 Question Posing Skill

We argue that improved prompt generation necessitates not only a nuanced comprehension of generative AI technology, as taught through prompt engineering courses, but also demands cognitive skills essential for formulating relevant inquiries, commonly known as "question posing" (chin et al., 2010). The optimal utilization of the tool hinges on the capacity to formulate a well-constructed prompt that in turn depends on both the learner's prompt engineering skills and their skill to precisely articulate the needed question. Crafting adequately useful prompts involves an individual's competency at various levels including:

- 1. **Analysis and Identification of Knowledge Gap:** Identifying the gaps, inconsistencies, and conflicts in the prior understanding of the context. These are commonly expressed in the form of questions.
- 2. **Prompt Generation:** Generation of prompts based on the identified question and prompt engineering techniques.

Users should be able to express what they truly need from the generative AI system. Individuals must be able to reflect and express what they don't know and what they want to know. They must be able to decompose their knowledge needs based on their priorities. ChatGPT may solve complex open-question answering problems if prompts are carefully decomposed based on user priorities. Users who have been trained with question-posing (and prompt engineering) skills have a completely different ChatGPT experience than users who have not been trained.



Figure 1. Antecedent cognitive processes to prompt generation

As illustrated in Figure 1, before writing any prompt, learners need to know what they know, what they don't know, or what they want to know, in order to receive a more specific response from ChatGPT. It is essential that users are aware that "what I know" represents prior knowledge checking, and "what I don't know" represents the context gaps. By identifying gaps in knowledge, users will be able to frame more focused questions from the beginning.

Prompt engineering has its own challenges like achieving the desired results on the first try, controlling the level of creativity of the result, and understanding and evaluating the reasoning behind the generated responses. Question posing (QP) can help to overcome some of the challenges of prompt engineering. QP is a cognitive tool that helps learners to think at a deep level and articulate their conflicts and beliefs to take meta-linguistic moves and formulate their concepts (Sasson et al., 2018). Question posing contains three parts. (1) Learners need to formulate what they know in the context based on their prior experience in the learning situation (Hwang et al., 2020). (2) Learners need to know what they don't know in the current context. (3) Frame questions targeting the gap (Mishra et al., 2015). According to Hwang et al. (2020), learners have difficulty in posing questions. Question posing is a skill that may help to pose the correct question on the first try.

#### 3. Generative AI: Threats to Fair Access

In the context of education, today or in the future, if the quality of education is closely linked to artificial intelligence in education (AIED) systems, it becomes imperative to address the potential threats to equitable access to these systems. Ensuring equitable access to such technologies will directly influence equitable access to quality education. Among the common risks associated with these AIED systems is the risk of "access to the system", specifically, access to its capabilities.

Most of the time, we discuss external factors, such as the digital divide, the unavailability of devices for learners from underprivileged socio-economic backgrounds, and barriers to technology integration in education (Tsai et al., 2012; Hsu, 2016; Kopcha, 2012). However, as discussed before, learners' internal characteristics such as questioning skills can also significantly block one's access to the capabilities of intelligent AIED systems, including ChatGPT, as discussed before. In Figure 2, we have classified threats to fairness of access into external and internal factors. For example, the digital divide falls within the external category, while internal threats encompass an individual's lack of knowledge, cognitive skills, or competencies that directly hinder one's ability to make use of the affordances and capabilities that technology has to offer.



Figure 2. Categories of Threats to Fairness of Access

#### 4. Discussion

UNESCO reports that during COVID-19, 850 million pupils, or half of the world's student population, were not able to attend school or college due to economic and social factors. This was stark evidence of inequitable access to e-learning primarily due to the digital divide. According to Singh et al. (2022), e-learning plays an important role in facilitating access to quality education. In the last eight years, the Government of India launched several e-learning initiatives, such as the Internet Saathi program, DIKSHA platform, Unnati project, E-pathshala, etc., as part of the Digital India campaign. These and similar projects throughout the world demonstrate the already significant reliance on information technology in education. It will not be surprising when quality education will similarly begin to rely on artificial intelligence-based educational technology in the very near future. We can easily anticipate such a future given the example of the technological storm brought in by the popular emergence of generative AI technologies in recent times. It should be noted that ChatGPT was launched on November 30, 2022, and the number of ChatGPT users exceeded 100 million by February 23. At this point, however, it is imperative to ask: How many of these 100 million users already had or have acquired the right competencies to make optimal use of ChatGPT-like technologies? The forthcoming reliance on AI in the educational domain is unavoidable, and at this point, it is imperative that we identify necessary preparations to take these technologies to individuals, including the needed training to develop the necessary competencies in individuals to make optimal use of such technologies.

We did not intend, in this paper, to present an exhaustive categorization of internal and external threats to fair access. Instead, our objective is to emphasize the need to identify and dig deeper into the essential internal competencies needed by users (learners, teachers, etc.) in order to effectively utilize existing and evolving artificial intelligence-based educational technologies, and develop training programs to develop such competencies among individuals. More specifically, we intended to emphasize the need to scaffold the development of question-posing skills with a lens that students' questions are critical to their exploration of new knowledge and deeper conceptual understanding (Mishra et al., 2015), especially in the context of Al supported self-learning modalities.

# References

Bogost, I. (2022). ChatGPT is dumber than you think. *The Atlantic*, 7.

Borji, A. (2023). A categorical archive of chatgpt failures. arXiv preprint arXiv:2302.03494.

- Cascella, M., Montomoli, J., Bellini, V., & Bignami, E. (2023). Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios. *Journal of Medical Systems*, *47*(1), 33.
- Chin, C., & Osborne, J. (2010). Students' questions and discursive interaction: Their impact on argumentation during collaborative group discussions in science. *Journal of research in Science Teaching*, 47(7), 883-908.

Dimitrov, M. (2023). What business leaders should know about using LLMS like ChatGPT.

Gao, J., Zhao, H., Yu, C., & Xu, R. (2023). Exploring the feasibility of chatgpt for event extraction. *arXiv* preprint arXiv:2303.03836.
https://www.unesco.org/en/articles/half-worlds-student-population-not-attending-school-unescolaunches-global-coalition-accelerate

- Hsu, P. S. (2016). Examining current beliefs, practices and barriers about technology integration: A case study. *TechTrends*, *60*, 30-40.
- Hwang, G. J., Zou, D., & Lin, J. (2020). Effects of a multi-level concept mapping-based question-posing approach on students' ubiquitous learning performance and perceptions. *Computers & Education*, *149*, 103815.
- Kopcha, T. J. (2012). Teachers' perceptions of the barriers to technology integration and practices with technology under situated professional development. *Computers & Education*, 59(4), 1109-1121.
- Loconte, R., Orrù, G., Tribastone, M., Pietrini, P., & Sartori, G. (2023). Challenging ChatGPT' Intelligence 'with Human Tools: A Neuropsychological Investigation on Prefrontal Functioning of a Large Language Model. *Intelligence*.
- Mohammadreza Farrokhnia, Seyyed Kazem Banihashem, Omid Noroozi & Arjen Wals (2023): A SWOT analysis of ChatGPT: Implications for educational practice and research, Innovations in Education and Teaching International, DOI: 10.1080/14703297.2023.2195846
- Mishra, S., & Iyer, S. (2015, June). Question-Posing strategies used by students for exploring Data Structures. In Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education (pp. 171-176)
- Prather, J., Denny, P., Leinonen, J., Becker, B. A., Albluwi, I., Caspersen, M. E., ... & Savelka, J. (2023, June). Transformed by transformers: Navigating the ai coding revolution for computing education: An iticse working group conducted by humans. In *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 2* (pp. 561-562).
- Sasson, I., Yehuda, I., & Malkinson, N. (2018). Fostering the skills of critical thinking and question-posing in a project-based learning environment. *Thinking Skills and Creativity*, *29*, 203-212.
- Singh, S., Singh, U. S., & Nermend, M. (2022). Decision analysis of e-learning in bridging digital divide for education dissemination. *Procedia Computer Science*, 207, 1970-1980.
- Sison, A. J. G., Daza, M. T., Gozalo-Brizuela, R., & Garrido-Merchán, E. C. (2023). ChatGPT: More than a weapon of mass deception, ethical challenges and responses from the human-Centered artificial intelligence (HCAI) perspective. *arXiv preprint arXiv:2304.11215*.
- Tsai, C. C., & Chai, C. S. (2012). The" third"-order barrier for technology-integration instruction: Implications for teacher education. *Australasian Journal of Educational Technology*, 28(6).
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., ... & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. *arXiv preprint arXiv:2302.11382*.
- Zheng, S., Huang, J., & Chang, K. C. C. (2023). Why does chatgpt fall short in providing truthful answers. *ArXiv preprint, abs/2304.10513*.
- Zhong, Q., Ding, L., Liu, J., Du, B., & Tao, D. (2023). Can chatbot understand too? A comparative study on chatgpt and fine-tuned bert. *arXiv preprint arXiv:2302.10198*.

## AIED in K-12 Classrooms: Challenges and Opportunities from an Ethics Lens

#### Anabil MUNSHI

Efficacy Research & Learning Engineering Team, Edmentum Inc., USA \*munshi.anabil@gmail.com

**Abstract:** The recent growth of the AIED field has made available several new tools to facilitate personalized learning and teaching. However, this has also raised concerns over the ethical principles and inclusivity guidelines that should monitor the use of AI tools in the classroom, especially in K-12 settings. This paper looks at some primary use cases of AI in education and reviews the concerns and challenges raised by AIED researchers over applying AI in K-12 educational spaces. We then discuss the ethical guidelines suggested by global agencies, government organizations, and AIED scholars, to direct educational policies that can help integrate AI tools in K-12 classrooms in a more ethical and inclusive manner.

Keywords: Al in K-12 education, Personalized learning, Ethics, Inclusivity, AIED Policy

#### 1. Introduction

The research field of Artificial Intelligence in Education (AIED) has grown considerably over the last decade, leading to the availability of a range of new AI-driven tools to facilitate learning. The primary end-user of most of these learning tools, viz., intelligent tutoring systems, is the learner; while other tools aimed more directly to support classroom educators, such as teacher dashboards and automated assessment systems, have also been developed. More recently, the growth and popularity of generative AI and large language models like ChatGPT (OpenAI, 2023) has led to further interest in leveraging the power of generative AI for applications in education.

While personalized learning tools driven by learning sciences principles and Al technologies have facilitated a greater understanding of learner difficulties and therefore enabled the provision of adaptive support to address student needs during learning, the application of AI models in classroom settings comes with its own ethical challenges. This becomes even more important when AIED tools are applied to K-12 (elementary through high school) spaces, where researchers have noted an ethics policy gap in governing the use and impact of AI for children (Adams et al., 2023).

In this context, this paper presents an overview of some of the major use cases of Al in education, outlines the key ethical concerns and risks identified by scholars and practitioners on the use of Al in K-12 educational spaces, and discusses guidelines and opportunities that can be used to develop educational policies to address these concerns for Al-integrated K-12 classrooms.

#### 2. Applications of Artificial Intelligence in Education

The rapid development of artificial intelligence (AI) technologies in the last few decades has presented several opportunities for computer scientists and learning sciences researchers to develop AI-powered learning tools and systems to support and enhance student learning. Personalized computer-based learning environments such as intelligent tutoring systems (Anderson et al., 1985; Azevedo et al., 2022), and open-ended learning environments (Land, 2000) track student activities during the learning process to inform "learner models" that attempt to capture an understanding of the dynamic cognitive, emotional, and behavioral

states of the learner as they learn. This information is used to detect the key moments during learning when the learner is facing difficulties and may benefit from receiving external support. This support, or *adaptive scaffolding*, is then often provided through system prompts or via pedagogical agents who engage the learner in conversational dialogues to gain a more contextual understanding of their current learning needs and guide them accordingly towards more efficient learning strategies and improved learning outcomes (Segedy et al., 2013; Munshi et al., 2022). The use of these personalized learning systems in the classroom also has the potential to reduce teachers' workload, especially in large classrooms where it may be difficult for the teacher to keep track of the obstacles faced by each individual learner as they work on a problem or task. Since teachers are at the epicenter of the classroom learning process, researchers have developed "teacher dashboards" to communicate the learner assessment information gained by these AI systems to the teacher (Aleven et al., 2022).

Beyond applications to directly foster student learning via intelligent learning environments, AI also has other use cases in K-12 education. For example, AI-driven automated essay scoring or exam grading tools can be used to assist the educator with administration tasks. AI models can also act as an instructional aid to the teacher by analyzing syllabus and course materials to propose personalized content (Chen et al., 2020).

Over the last year, with the availability and increased attention on generative AI tools like ChatGPT, there has also been discussions in the research community on using generative AI to address challenges in education. Kasneci et al. (2023) suggest how large language models like ChatGPT can help elementary school students develop reading comprehension skills by providing explanations and summaries of complex texts, and also help semi-automate students' task grading by identifying both the strengths and weaknesses of the learning task at hand (Kasneci et al., 2023; Soc & Heng, 2023). However, researchers also caution against potential unethical and discriminatory use of AI technologies like ChatGPT in learning contexts (Mhlanga, 2023).

In summary, the field of AIED has grown in conjunction with recent advances in educational psychology and the learning sciences to come up with AI-powered learning systems that can both provide learners with a personalized learning experience and support educators in their instructional process in the classroom. However, ethical concerns raised around the use of AI systems in K-12 education need to be identified and addressed before these systems are integrated into classrooms. We discuss some of the major concerns below.

#### 3. Ethical Concerns Around the Use of Al in K-12 Education

Akgun et al. (2022) outline the key ethical concerns associated with AIED applications in K-12 education along four categories: (1) privacy; (2) surveillance; (3) autonomy; and (4) bias and discrimination.

Al tools often require the user to consent to sharing sensitive personal information (gender, demographics, location, etc.) as part of the initial user agreement. This presents privacy concerns, especially in cases where the use of these tools is required by the school and there is less agency on part of the end-users (teacher and students) to decide on consenting to the collection of personal information by the tool. AIED systems that use machine learning and predictive modeling algorithms to better understand learner behaviors and adapt the learning process to their needs may also need to track and monitor learner activities and/or conversations in the learning environment, thereby leading to surveillance concerns. In this context, the knowledge of being monitored may also make the learner feel unsafe and inhibit their level of participation in the learning task (Akgun et al., 2023). AIED platforms that are less open-ended and compel the learner to move along a specific learning trajectory driven purely by the judgments of a predictive algorithm may also take away learners' decision-making *autonomy* on their own learning process. Additionally, research has shown that AI models may have their own inherent biases (Roselli et al., 2019; Ntoutsi et al., 2020). If a K-12 learner is given feedback based on incorrect interpretation of their mental states or other behavioral characteristics by a machine learning model, it may disengage the learner or have even more harmful effects on their mental, socio-emotional and cognitive processes. Therefore, a learning environment that is purely Al-driven - without a human

educator in the loop – can inadvertently promote discriminatory assessment practices that affect learner grades and are detrimental for their future career prospects.

To address these concerns, it is imperative to develop guidelines and educational policies that ensure that AI systems are used in K-12 classrooms to foster learning in a fair, responsible, and inclusive manner.

#### 4. Educational Policies to Govern the Use of Al in Classrooms

Initial conversations around ethical guidelines on the use of AI in education were started more than twenty years ago by Aiken and Epstein (2000), but they did not see widespread adoption in the development of AIED systems. However, the more recent development of AIED systems and the associated recognition of the power of AI in transforming the classroom learning space has led to a renewed focus and attention on the ethics of using AI technologies in educational contexts. We discuss recommendations from some AIED ethics documents below. We adopted the following inclusion criteria for selecting the documents reviewed in this section: (a) the document was published by a global agency or government organization; (b) it provides policy recommendations on the ethical use of AI in schools; and (c) the document was released between the years 2018 and 2023, thereby more accurately reflecting the changing education landscape during the COVID period and the increasing shift towards digital learning in a post-COVID world that makes it even more important to develop ethical policies for the use of AI in education.

Adams et al. (2023) performed a content analysis of four globally relevant AIED ethics guidelines released between the years 2019 and 2021 by the World Economic Forum (2019), IEAIED (2021), UNESCO (2021), and the United Nations Children's Fund (2021) – and identified four key ethical principles to guide AI policy development specifically pertaining to K-12 applications:

(1) ensuring *pedagogical appropriateness* of the AI system by weighing its positive effects (adaptive learner support) against potential negative social or development outcomes;

(2) protecting *children's rights*, by taking steps such as actively involving children in AI policy development or by ensuring the AI is explainable in age-appropriate language that allows the child to provide informed assent to its use;

(3) taking initiatives to develop AI literacy for students, teachers, as well as parents; and

(4) ensuring *teachers' well-being* and clearly defining teachers' roles and preparing them to work effectively in AI-integrated learning spaces.

These principles are further echoed in an "Insights and Recommendations" document released by the US Department of Education in May 2023 on the use of AI in schools, which discusses the need to develop a national education-focused AI policy (Cardona et al., 2023). The guidelines in this document emphasize keeping humans (teachers and students) in the instructional loop, aligning AI models to a shared vision for education, designing these models using modern learning principles, prioritizing trust, informing and involving educators, and focusing research and development efforts on addressing context and enhancing trust and safety in AIED.

The European Union's recent ethical guidelines on the use of AI in teaching and learning focus on ensuring agency and children's rights in system design and use, maintaining AI-driven learning systems' technical robustness and resilience to attack, and overall building "trustworthy AI" that is transparent, accessible, and inclusive considering use by children of all genders and demographic characteristics, with a special focus on learners with special needs (European Commission, 2022).

Al policy strategies produced by several other countries also focus on taking measures to promote ethical and inclusive AI in education with more contextual focus on the unique education landscapes of these countries. For instance, the National Strategy for Artificial Intelligence released by NITI Aayog in India calls for developing "AI technologies that are capable of imparting quality education to India's linguistically diverse population" and setting up "a consortium of Ethics Councils at each Centre of Research Excellence" to ensure that the development of AI products adhere to standard ethical practice guidelines (NITI Aayog, 2018).

UNICEF's recent policy guidance for the use of AI on children further emphasize that AIED businesses must create system guides and have internal codes of conduct in place that ensure that their software and hardware development teams implement these AI systems keeping the ethical considerations and their effects on children in mind (UNICEF, 2021).

In the changing educational space where AI technologies are heavily applied in the classroom, AIED researchers further caution against developing systems that "attempt to replace the teacher" or "diminish the human role and the human potential for learning and growth" by relying only on inferences provided by data-driven systems. Instead, AIED models should be transparent and explainable, and help teachers better understand learner needs and optimize their teaching process, and additionally provide teachers with "new and creative roles that might not have been possible before the use of technology" (Holmes et al., 2022).

#### 5. Towards Ethical and Inclusive AI-Rich K-12 Classrooms

This paper reviewed some major application areas for artificial intelligence in the field of education and outlined ethical challenges and concerns raised by AIED researchers, policy-makers and other stakeholders. While AI-driven personalized learning models have the potential to positively transform the classroom teaching and learning process by providing educators and learners with a wide range of tools and scaffolding mechanisms to become more efficient at the task at hand, AI system developers must understand the risks and concerns of their end-users and take all possible steps to build ethical and inclusive AI learning environments. Clear government policies on ethical AIED, driven by the guidelines reviewed in this paper, can be a great first step towards ensuring accountability, transparency, privacy, and fairness in AIED system development, so that educators and all learners can feel safe in leveraging the benefits of integrating AI into their classrooms.

#### References

- Adams, C., Pente, P., & Lemermeyer, G. (2023). Ethical principles for artificial intelligence in K-12 education. *Computers and Education: Artificial Intelligence, Volume 4*.
- Aiken, R. M., & Epstein, R. G. (2000). Ethical guidelines for AI in education: Starting a conversation. *International Journal of Artificial Intelligence in Education*, 11, 163–176.
- Akgun, S., Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI Ethics, 2, 431–440.* <u>https://doi.org/10.1007/s43681-021-00096-7</u>
- Aleven, V., Blankestijn, J., Lawrence, L., Nagashima, T., & Taatgen, N. (2022). A dashboard to support teachers during students' self-paced Al-supported problem-solving practice. In *European Conference on Technology Enhanced Learning* (pp. 16-30). Cham: Springer International Publishing.
- Anderson, John R., C. Franklin Boyle, and Brian J. Reiser. (1985). Intelligent tutoring systems. *Science* 228.4698: 456-462.
- Azevedo, Roger, et al. (2022). Lessons learned and future directions of metatutor: Leveraging multichannel data to scaffold self-regulated learning with an intelligent tutoring system. *Frontiers in Psychology* 13: 813632.
- Cardona, M., Rodriguez, R.J., and Ishmael, K. (2023). Artificial Intelligence and the Future of Teaching and Learning: Insights and Recommendations. *Office of Educational Technology, US Department of Education.*
- Chen, L., Chen, P. and Lin, Z. (2020) Artificial Intelligence in Education: A Review. *In IEEE Access, vol. 8, pp. 75264-75278*, doi: 10.1109/ACCESS.2020.2988510.
- European Commission, Directorate-General for Education, Youth, Sport and Culture (2022). Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators, *Publications Office of the European Union*, <u>https://data.europa.eu/doi/10.2766/153756</u>
- Holmes, W., Porayska-Pomsta, K., Holstein, K. et al. Ethics of Al in Education: Towards a Community-Wide Framework. Int J Artif Intell Educ 32, 504–526 (2022). <u>https://doi.org/10.1007/s40593-021-00239-1</u>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education.

Land, S. (2000). Cognitive requirements for learning with open-ended learning environments. *Educational Technology Research and Development*, 48:61–78

- Mhlanga, D. (2023). Open AI in Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning. *SSRN 2023, 4354422.*
- Munshi, A., Biswas, G., Baker, R., Ocumpaugh, J., Hutt, S., & Paquette, L. (2022). Analyzing adaptive scaffolds that help students develop self-regulated learning behaviors. *Journal of Computer Assisted Learning*, Vol 39 (2), 351-368.
- NITI Aayog, India. (2018). India: National Strategy for Artificial Intelligence #AIFORALL. Delhi, India: NITI Aayog.
- Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M. E., ... & Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(3), e1356.
- OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model]. https://chat.openai.com/chat
- Roselli, D., Matthews, J., & Talagala, N. (2019). Managing bias in Al. *In Companion Proceedings of The 2019 World Wide Web Conference* (pp. 539-544).
- Segedy, J., Kinnebrew, J., and Biswas, G. (2013). The effect of contextualized conversational feedback in a complex open-ended learning environment. *Educational Technology Research and Development*, 61(1):71–89.
- Sok, S. & Heng, K. (2023). ChatGPT for education and research: A review of benefits and risks. *Cambodian Journal of Educational Research*.
- The Institute for Ethical AI in Education (2021). The ethical framework for AI in education. *University of Buckingham.*

United Nations Children's Fund (2021). Policy guidance on AI for children 2.0.

- United Nations Educational, Scientific and Cultural Organization (2021). Al and education: Guidance for policy makers.
- World Economic Forum (2019). Generation AI: Establishing global Standards for children and AI.

# W10

## The 12th International Workshop on ICT Trends in Emerging Economies (WICTTEE 2023)

## Factors Affecting Sustainable Use of Minecraft-based Lessons

#### Dominique Marie Antoinette B. MANAHAN<sup>a\*</sup> & Maria Mercedes T. RODRIGO<sup>b</sup>

<sup>a</sup>Ateneo Laboratory for the Learning Sciences, Ateneo de Manila University, Philippines <sup>b</sup>Ateneo Laboratory for the Learning Sciences, Ateneo de Manila University, Philippines \*dmanahan@ateneo.edu

**Abstract:** This paper examines the factors that affect the sustainable use of Minecraft-based lesson plans. We focus specifically on teachers who received training on using What-If Hypothetical Implementations in Minecraft (WHIMC), a set of Minecraft worlds designed to teach astronomy, geosciences, and ecology. We commissioned a group of teachers to develop lesson plans using at least one of the WHIMC worlds and pilot their implementation in their classes. Another group of teachers received training on developing learning modules with WHIMC but had the freedom to choose when and how to implement their lesson plans during the school year. The teachers cited logistical impediments, scheduling impediments, lack of technical resources, lack of curricular alignment, changes to teaching modalities, limited time for practice, and anxieties about classroom management as barriers to sustainable use of WHIMC. They suggested the training of additional teachers to create a community of practice whose members can collaborate and support each other's use of WHIMC for STEM education.

Keywords: WHIMC, Philippines, Game-based Learning

#### 1. Context

In 2020, the Ateneo de Manila University (ADMU) and the University of Illinois Urbana-Champaign (UIUC) entered into a partnership in which UIUC allowed ADMU to use its What-If Hypothetical Implementations in Minecraft (WHIMC; https://whimcproject.web.illinois.edu/) in schools in the Philippines. WHIMC is a set of computer simulations built in Minecraft Java Edition to engage students in "what if" questions such as "What if Earth had no moon?" or "What if Earth had a slightly colder sun?" WHIMC immerses students in these alternate versions of Earth, giving them opportunities to see how the planet's geography and life forms would differ under these circumstances. It enables students to use science tools to collect data about these alternate Earths and record their observations.

The UIUC team successfully used WHIMC in summer school programs (Lane, et al., 2022). It also recently partnered with several planetariums to include WHIMC installations among the planetariums' exhibits. Furthermore, UIUC has used the data collected from their WHIMC deployments to study learner-related phenomena such as the structure and skill level of learner scientific observations (Hum, et al., 2022) and STEM interest behaviors (Gadbury & Lane, 2022), among others.

Leveraging on UIUC's success, the goal of the ADMU team was to determine whether WHIMC could cultivate Filipino learners' interest in STEM. ADMU researchers were motivated by the need to improve Science, Technology, Engineering and Mathematics (STEM) achievement in the Philippines. The PISA National Report on the Philippines (Philippines Department of Education, 2019) said that, compared to the OECD average of 489 in math and 489 in Science, Filipino students scored a low 353 and 357 respectively. Only 1 out of 5 attained the minimum proficiency level in math. These results are corroborated by students' performance in the National Achievement Test, where only 25% demonstrated mastery levels in math and only 5% of test takers demonstrated mastery levels in science.

Through a 2021 grant from the Philippines' Department of Science and Technology, the ADMU purchased 212 Minecraft Java Edition licenses and rented a server that allowed it to mirror the UIUC WHIMC site. The ADMU team then worked with teachers to help them gain familiarity with WHIMC and develop lesson plans that integrated WHIMC in science, math, and even English classes. The ADMU team also provided the teachers with technical support during actual conduct of these lessons.

Since then, the ADMU team has trained 23 teachers for WHIMC Java Edition and 24 teachers for WHIMC Education Edition. In the ADMU WHIMC Java Edition server, over 304 students were able to explore it. Furthermore, it developed a Minecraft Education Edition version of some of the worlds (see teachers' guide through <u>bit.ly/WHIMCEETeachersGuide</u>). This version has been used by 78 students and 5 teachers in 2 Philippine schools. The ADMU team also used the data from its deployments to study the effects of WHIMC on student STEM interest (Tablatin, Casano, & Rodrigo, 2023), the relationship between STEM interest and student affective states such as frustration and boredom (Esclamado, Rodrigo, & Casano, 2022), and to compare the quality of Philippine and US student observations (Casano & Rodrigo, 2022).

Despite these successes, continued use of WHIMC is not guaranteed. Indeed, now that the DOST grant has ended, the overarching question we pose in this paper is one of sustainability. Will teachers continue using WHIMC as the ADMU project team starts removing the scaffolds? In this paper, we attempt to examine what factors contribute or hinder the sustainability of WHIMC use in Philippine classrooms.

#### 2. Sustainability of Educational Technology Initiatives

In the context of educational technology initiatives such as WHIMC, sustainability is defined as persistent and ongoing change of the educational culture (Neiderhauser, et al., 2018). Sustainability is a process that unfold when support systems are established and the community in which the initiative takes place maintains the initiative over time. When an initiative is sustained, the innovation or intervention becomes richer and more sophisticated as it evolves. Initiatives factor in changes in context. Increased knowledge of and comfort with the innovation prompts stakeholders to maximize the innovation's affordances.

These characteristics of sustainability are evident in many Minecraft applications. The work of Lincenberg and Eynon (2021) discusses how educators use Minecraft to transform classrooms into social spaces where teachers and students could visit or challenge each other, spaces where students could respond to the curriculum in their own ways, e.g. by solving problems that their teachers gave or that they defined themselves. A review by Baek, Min, and Yun (2020) shows how Minecraft has been successfully integrated in science, math, social science, and language subjects, and has been shown to increase creativity, improve technology skills, increase collaborative skills, and encourage communication.

There are, however, many factors that hinder the sustainability of educational technology initiatives. Ertmer (1999) classifies these barriers into two: first-order barriers are extrinsic to teachers. These include lack of equipment, time, training, or support. Second-order barriers are intrinsic to teachers. These include teachers' underlying perceptions about how teaching and learning should take place. First-order barriers are often cited particularly in developing world contexts. In Tanzania, for example, barriers included a lack of basic infrastructure and a lack of motivation and support from school management are impediments (Kafyulilo, Fisser, & Voogt, 2016). In the Philippines, the cost of Internet access, the lack of curricular alignment, and insufficient teacher professional development make it difficult for schools, teachers, and students to incorporate technology in their classes (Rodrigo, 2021).

Second-order barriers have been discussed in the context of Minecraft-based lessons. Thinking of ways to incorporate Minecraft in lessons necessitates familiarity with technology in general and Minecraft in particular. This creates work that is over and above

regular teaching tasks. Innovations such as Minecraft also necessitate a constant upgrading of skills and constant adaptation, which challenges teachers' already limited time (Thorsteinsson & Niculescu, 2016).

This paper reports feedback from teachers who were trained to use WHIMC about the factors that helped or hindered them from using WHIMC in their classes, and the likelihood that they will use WHIMC (again) in the future. By asking teachers for their feedback, we assess WHIMC's sustainability. Is it likely to continue being a useful tool for teachers and students in Philippine classrooms or will first- or second-order barriers become so insurmountable that the use of WHIMC is more likely to fizzle out as so many education technology initiatives do?

#### 3. Data Collection

#### 3.1 Participant Selection

All of the partner teachers were endorsed by the researchers' institutional partners and teacher participants who attended the module development training using WHIMC under the project were invited to participate voluntarily in an online focus group discussion (FGD) session. The module development training provided varied, as there were school schedules that allowed a five-day online training while others had only one full working day available but held on-site in the school's own laboratory. In general, the format of the trainings held provided first a short orientation about the educational framework of WHIMC, a discussion on the features of Minecraft, and a short demonstration on how to navigate through the WHIMC worlds. Then, the teachers were provided an opportunity to initially explore the WHIMC worlds assigned using lent Minecraft accounts and a training guide. The training guide contained three to five guide guestions for each WHIMC world that would allow the teacher participants to look into the content and environment of the WHIMC world. After being able to explore the WHIMC worlds, the teachers were given the space to design and draft their learning modules for their own classes. This was done individually for some schools, while other schools designed a module as a group according to the grade level they taught. The presentation included a feasibility assessment by the teacher on whether they could fully implement and integrate Minecraft in their classes. Afterwards, everyone is given a chance to present their module drafts to the rest of the participants and trainers, and hear feedback about how they could improve the module developed. Finally, the trainers gave an orientation on how to prepare their classes for a WHIMC lesson.

Upon invitation to attend the FGD, the teachers were provided session schedule options with a maximum of 1.5 hours each session. The confirmed FGD participants were sent an electronic calendar invite with an online conference link where the FGD would be held at least two weeks before the FGD session schedule they had chosen. On the FGD session schedule, the teacher participants gathered at the online meeting room hosted by the project manager and the project leader.

#### 3.2 Focus Group Discussions

The project leader and project manager prepared a presentation and an online word document for the FGD. The online word document served as the main data gathering tool to collect the FGD participants' basic information and notes on their responses to the FGD questions.

The session began with the project manager and project leader welcoming the participants. The project manager proceeded to state the session objectives, which was "to identify the factors that motivate or hinder the implementation of a WHIMC module in the classroom. During the FGD, participants were given five (5) minutes to post their responses the question applicable to their situation on the Google Doc file provided. The project manager and project leader facilitated the discussion in the online meeting room, where the participants were given the space to share about their responses to the group.

There were two (2) questions raised during the online session. The first question was applicable to the teachers who had already developed and implemented a lesson plan in at least one class, and read by the project leader or manager, "If you implemented a WHIMC lesson plan before, would you implement it again? Why or why not?" The other question was applicable to the teachers who had not implemented a WHIMC lesson plan yet, and asked "If you did not implement one before, why? Would you implement one now? Why or why not?"

#### 4. Results and Discussion

The project team was able to interview a total of nine (9) participants, five (5) male and four (4) female basic education teachers, working in Philippine private and public basic education institutions. The participants had an average age of thirty-five (35) years old, with an average of eleven (11) years of basic education teaching experience, and an average of ten (10) years of teaching a STEM-related subject.

Four (4) teachers who were trained to use WHIMC did not pursue the implementation of a WHIMC lesson plan in their actual classes. The remaining five (5) did so. All of them expressed that, if given an opportunity, they would implement a WHIMC lesson in their classes again. One female teacher from a private school shared that she had observed the depth of learning from her students as they used WHIMC in their class. The students' discussion went beyond the intended topic, and had seen how theoretical concepts actually happen through their exploration inside the WHIMC worlds. She also observed that it developed their affective and social skills, since the students were able to play, communicate and collaborate on their movements within the WHIMC Java Edition server during their online WHIMC class. Two (2) other teachers expressed that they would choose to implement a WHIMC lesson as a supplement or enrichment activity in their classes. For example, it was shared by one of them that there are topics, such as the solar system, where it is difficult to think of experiments for. Thus, the WHIMC worlds exploration could be a good tool to help the students learn deeper about the topic. Then, one teacher from a public school shared that she saw how eager the students were during class which encouraged her to plan for another WHIMC lesson.

The first-order barriers cited by both groups of teachers were similar:

- 1. Logistical impediments. Given that the schools only have a common computer laboratory, which was usually used for the computer education subject, venue availability for a WHIMC class was seen as a concern. It was raised that the school would hopefully increase the computer laboratories and/or update their policies to allow the students to bring their own devices or laptops.
- 2. Lesson scheduling. The WHIMC lessons tended to be quite lengthy, exceeding the usual class period length for some schools. One teacher noted that they had limited contact time with the students within a week (e.g. three times a week to teach one module). The challenge to teachers was to integrate WHIMC into these limited time slots. Having more time would make implementing a WHIMC module more feasible.
- 3. **Technical resources.** For schools that implemented a WHIMC lesson through an online modality, it was said to be more convenient if the school had available computers to lend to avoid any delays in the student's accomplishment of the WHIMC class' activities.
- 4. **Curriculum alignment.** Of the four (4) teachers who were not able to implement a WHIMC lesson, three (3) were assigned to teach computer education. Their curriculum goals and competencies were not aligned to the content of WHIMC, and focused on other STEM areas such as robotics and programming. Meanwhile, for one of the teachers under this group that taught Science (specifically, Physics), it was raised that the developed module during the teacher training was not fit for the lesson topics his class had in the pipeline.
- 5. **Changes in teaching and learning modalities.** Throughout the school year, policies on the teaching and learning modalities have constantly shifted due to the changing COVD-19 restrictions, to which the schools adjusted to every now and then. The change from flexible learning, partially on-site and fully on-site classes affected the lesson planning of the teachers.

The second-order barriers cited by both groups of teachers included the following:

- 1. Limited time to practice the lesson design or for dry runs. Teachers said that they had limited time to explore WHIMC and practice their lesson designs. Additional time to practice would give them more chances to think of and implement WHIMC-based lesson in their classes. They could also use the time to improve upon the lessons that they had already created.
- 2. **Concerns about classroom management.** Related to the second first-order barrier, teachers expressed anxiety about how they would manage a WHIMC-based class during a 50-minute class period. They were not sure they would be able to orchestrate a proper exploration of the worlds, given the limited time.

The teachers did have some suggestions about how these barriers could be overcome. They suggested that WHIMC training be made available to more STEM teachers. This creates a community of practice whose members can work together on common lesson plans, thereby allowing the school to implement WHIMC-based lessons across more sections of a grade level, as the STEM subjects are usually taught by multiple teachers.

#### 5. Conclusion

We found that for both groups of teachers, those who were and were not able to implement a WHIMC lesson, encountered similar first- and second-order barriers. They noted the feasibility of running a WHIMC lesson in their classes was hindered due to logistical, technical and curriculum concerns. To be able to address said barriers, the school administration's support is needed to be able to realize the material and content support needs of the teachers. The difference seemed to be that those who were able to implement were also able to overcome the limited time for practice and any anxieties they might have had about classroom management. The teachers did not mention barriers similar to the ones cited in prior literature (see Thorsteinsson & Niculescu, 2016) but it is consistent with some of the findings of general impediments to the use of technology in developing world classrooms (see Rodrigo, 2021).

In summary, to successfully sustain the implementation of Minecraft-based lessons, it requires the support of school leadership for the needed facilities and technological infrastructures. Having a community of teachers that are equipped to develop and manage such kinds of lessons is also important to push for the institutionalization of lesson plan development and implementation. Additionally, having a more stable mode of teaching and learning as well as a community of practice would allow teachers to successfully plan and consider integrating Minecraft in their class topics.

#### Acknowledgements

We thank H Chad Lane and Jeff Ginger for their enthusiastic collaboration. We thank our partner teachers and schools. We thank Ms. Rosario Madjos for her administrative support. We thank our colleagues at the Ateneo Laboratory for the Learning Sciences. Finally, we thank the Ateneo de Manila University for making this work possible through the grant entitled, "The Impact of Technology Use on Learning Outcomes and the Learning Experience."

#### References

Baek, Y., Min, E., & Yun, S. (2020). Mining educational implications of Minecraft. *Computers in the Schools*, *37*(1), 1-16.

Casano, J. D., & Rodrigo, M. M. T. (2022, July). A Comparative Assessment of US and PH Learner Traversals and In-Game Observations within Minecraft. In Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium: 23rd International Conference, AIED 2022, Durham, UK, July 27–31, 2022, Proceedings, Part II (pp. 267-270). Cham: Springer International Publishing.

- Ertmer, P. A. (1999). Addressing first-and second-order barriers to change: Strategies for technology integration. Educational technology research and development, 47(4), 47-61.
- Esclamado, M. A., Rodrigo, M. M. T., & Casano, J. A. (2022). Achievement, Behaviors, and STEM Interest of Frustrated and Bored Learners Using Minecraft. In 30th International Conference on Computers in Education (ICCE 2022).
- Gadbury, M., & Lane, H. C. (2022, July). Mining for STEM Interest Behaviors in Minecraft. In Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium: 23rd International Conference, AIED 2022, Durham, UK, July 27–31, 2022, Proceedings, Part II (pp. 236-239). Cham: Springer International Publishing.
- Hum, S., Stinar, F., Lee, H., Ginger, J., & Lane, H. C. (2022, July). Classification of Natural Language Descriptions for Bayesian Knowledge Tracing in Minecraft. In Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium: 23rd International Conference, AIED 2022, Durham, UK, July 27–31, 2022, Proceedings, Part II (pp. 250-253). Cham: Springer International Publishing.
- Kafyulilo, A., Fisser, P., & Voogt, J. (2016). Factors affecting teachers' continuation of technology use in teaching. *Education and Information Technologies*, *21*, 1535-1554.
- Lane, H. C., Gadbury, M., Ginger, J., Yi, S., Comins, N., Henhapl, J., & Rivera-Rogers, A. (2022). Triggering STEM Interest With Minecraft in a Hybrid Summer Camp. Technology, Mind, and Behavior, 3(4). DOI: 10.1037/tmb0000077
- Lincenberg, J., & Eynon, R. (2021). Crafting worldly spaces: The role of the educator in shaping student agency in Minecraft. Review of Education, Pedagogy, and Cultural Studies, 1-23.
- Niederhauser, D. S., Howard, S. K., Voogt, J., Agyei, D. D., Laferriere, T., Tondeur, J., & Cox, M. J. (2018). Sustainability and scalability in educational technology initiatives: Research-informed practice. Technology, Knowledge and Learning, 23, 507-523.
- Philippines Department of Education. (2019, December). PISA 2018 National Report of the Philippines. Accessed from <u>https://www.deped.gov.ph/wp-content/uploads/2019/12/PISA-2018-Philippine-National-Report</u>
- .pdf Rodrigo, M. M. T. (2021, July). Impediments to AIS Adoption in the Philippines. In Adaptive Instructional Systems. Design and Evaluation: Third International Conference, AIS 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part I (pp. 412-421). Cham: Springer International Publishing.
- Tablatin, C. L. S., Časano, J. D., & Rodrigo, M. M. T. (2023, March). Using Minecraft to Cultivate Student Interest in STEM. In Frontiers in Education (Vol. 8, p. 143). Frontiers.
- Thorsteinsson, G., & Niculescu, A. (2016). Pedagogical insights into the use of Minecraft within educational settings. *Studies in Informatics and Control, 25*(4), 507-516.
- WHIMC Research Office. (no date). What-If Hypothetical Implementations in Minecraft (WHIMC). Accessed on 7 April 2023 from the WHIMC website: <u>https://whimcproject.web.illinois.edu/</u>

## A Mathematical App for the Conceptual Understanding of Area and Perimeter

#### Jumela F. SARMIENTO<sup>a,\*</sup>, Debbie Marie B. VERZOSA<sup>b</sup>, Maria Alva Q. ABERIN<sup>a</sup>, Ma. Louise Antonette N. DE LAS PEÑAS<sup>a</sup>, Agnes D. GARCIANO<sup>a</sup>, Juan Carlo F. MALLARI<sup>a</sup>, & Mark Anthony C. TOLENTINO<sup>a</sup>

<sup>a</sup>Department of Mathematics, Ateneo de Manila University, Philippines <sup>b</sup>Department of Mathematics and Statistics, University of Southern Mindanao, Philippines \*jsarmiento@ateneo.edu

**Abstract:** This paper discusses an app that was developed to build a strong understanding of the concepts of area and perimeter in students. An important feature of the app is the three-component feature which highlights progressive learning: *Explore,* designed for the learning of the conceptual understanding of area and perimeter; *Apply,* where area and perimeter concepts are applied; and *Create* intended for constructing representations to develop higher order thinking skills. The pedagogical basis for the creation of the app, the game design elements employed in the app as well as the integration of the app in the classroom will be presented.

**Keywords:** Mathematical app, area and perimeter, area, perimeter

#### 1. Introduction

There is a longstanding and general consensus worldwide (Hiebert & Carpenter, 1992; Skemp, 2006) and in the Philippines (Department of Education, 2016) that mathematics must be learned relationally; that is, as an integrated network of connections between and among mathematical concepts. Teaching area and perimeter meaningfully entails going beyond the teaching of formulas or procedures so that students could justify these formulas and reconstruct them in case they forget (Lithner, 2017). Toward this end, designing instructional materials to facilitate students' learning of area and perimeter is essential.

A problem arises in emerging economies whereby many schools routinely cater to illprepared students, which forces teachers to resort to rote-based strategies in teaching area and perimeter. In (Bansilal, 2011), classroom discussion facilitated by a teacher in South Africa focuses largely on helping her students remember the formula rather than on allowing her students to reason about the authentic performance-based assessment mandated by their national board of education. The reason for this teachers' actions was due to her learners' non-readiness for the task, preventing the students from engaging in more meaningful mathematical activity. In Taiwan, which has historically performed well in international assessments, there are teachers who present formulas directly instead of engaging their learners to explore the geometric basis for these formulas (Huang, 2017).

The potential of technological tools for teaching area and perimeter has been explored (Hwang et al., 2020). This paper discusses the *Area and Perimeter* app aimed to provide students the opportunity to understand the conceptual underpinnings of area and perimeter formulas. This app was designed in consideration of scholarly literature and was one of the outputs of a government-funded project in the Philippines (De Las Peñas et al., 2023). It was designed in a game-like environment to address particular learning competencies in mathematics. It is compatible with and runs efficiently in a wide range of Android devices (e.g., smartphones, tablets) and Windows PC (laptops or desktops) without the internet.

#### 2. Pedagogical Basis

The Area and Perimeter app was designed based on the concrete-pictorial-abstract approach developed by Jerome Bruner (1996). The idea is to carefully sequence representations that would enable learners to transition from concrete to abstract (Ding & Li, 2014). As such, area formulas (which are abstract concepts) are strongly tied to geometric shapes (which are pictorial representations). The app also consists of three levels: *Explore*, *Apply*, and *Create*. By *exploring* geometric objects, students can think about area and perimeter using particular cases. By playing the app repeatedly, they can generalize, understand the conceptual basis for the area and perimeter formulas, and *apply* them to compute the area and perimeter of a given geometric figure. Finally, students are provided an opportunity to develop higher-order thinking by *creating* geometric objects given the area or perimeter. This sequence of *Explore*, *Apply*, and *Create* allows the student to develop relational understanding instead of relying solely on procedural formulas of area and perimeter.

#### 3. The Area and Perimeter App

#### 3.1 App Description

In the Area and Perimeter app, students have the option to choose the following topics: Area, *Perimeter*, or Area and Perimeter. The Area component includes the following subtopics: *Introduction to Area, Rectangle/Square, Triangle, Parallelogram* or *Trapezoid*. The app allows the student to investigate area properties of rectangles, triangle, parallelogram and trapezoid. The *Perimeter* component includes *Introduction to Perimeter, Rectangle* or *Rectilinear Figures*. Herein, a student can learn perimeter properties of rectangle or rectilinear figures. In the Area and Perimeter component, students find the area and perimeter of rectilinear shapes and construct rectilinear shapes of given area and perimeter.

There are three levels in the app, to reflect the pedagogical intention to enable students to develop deep understanding of area and perimeter, beyond knowing formulas by rote.

*Explore.* In this level, students learn the conceptual underpinnings of both area and perimeter. Under the topic *Area* and subtopics *Introduction to Area* and *Rectangle/Square*, students are shown a rectilinear shape then they are supposed to click on each unit square within the shape (Figure 1(a)) and are asked how many squares are enclosed in the figure. This way, they connect the idea of area to the measure of the surface inside the given enclosed region.



Figure 1. Screenshots of the app; Topic: Area and Level: Explore.

For the sub-topics of triangles, parallelograms, and trapezoids the aim of the *Explore* component is to allow students to learn the reasoning behind the area formulas. For example, for *Triangle*, a rectangle is first shown and then a diagonal is drawn (Figure 1(b)) to show that half of the rectangle is a triangle. This reinforces the idea that the area of a triangle is half the

area of a corresponding rectangle. For *Parallelogram*, an auxiliary line segment is shown (Figure 1(c)) to illustrate that translating the triangle at one end of the parallelogram to the other end would result in a rectangle. This reinforces the idea that the area of a parallelogram is the same as that of a rectangle with the same height and width. Finally, for *Trapezoid* (Figure 1(d)), a trapezoid (in green border) is shown and is duplicated (whole figure) and placed adjacent to the former so that the combined figure is a parallelogram. This reinforces the idea that the area of a trapezoid is half the area of a parallelogram with the same height and base. A score on the lower right corner of the screen allows the student to monitor his or her own progress and decide if he or she can choose another feature of the app.

*Apply.* Unlike the *Explore* level, in the *Apply* level, there are no extra steps, students are directly asked, "How many square units are in the figure?" At the *Apply* level, students use the ideas in the *Explore* component to determine the area of given geometric figures.

*Create.* In this level, students construct representations themselves, enabling them to develop higher-order thinking skills. Specifically, students are asked to draw a shape by clicking and connecting vertices to produce the required shape with a given area (Figure 2(a)). On the right side of the screen, the target area and current area are given. Thus, if a student constructs a shape with an incorrect area (Figure 2(b)), there is an instant feedback that provides an opportunity for the students to construct a new shape until they get one with the correct area.



Figure 2. Screenshots of the app, Topic: Area and Level: Create

Under the *Perimeter* option, the same levels (*Explore, Apply* or *Create*) are also available.

*Explore.* In this level, the students are asked to click on the edges of the shape and count the number of red edges such as in an example under *Rectilinear Figure* that is shown in Figure 3(a). When the correct answer is given, the app affirms that the shape has the given number as its perimeter. Through this, students connect the idea of perimeter with the number of edges around the shape.

*Apply.* In this level, students must directly input the perimeter of the given shape (Figure 3(b)) using the ideas from the *Explore* component to be able to find the perimeter.



Figure 3. Screenshots of the app, Topic: Perimeter and Levels: Explore (a), Apply (b)

*Create.* Here, students are tasked to construct a rectilinear shape with a given perimeter (Figure 4(a)). On the right side of the screen, both the target perimeter and the current perimeter are given. As the students add edges to form a shape, they will see that the current perimeter is updated, and this will guide them in constructing the correct shape having the desired target perimeter. For example, as shown in Figure 4(b), the target perimeter is 10 while the current perimeter is 7 and it is not possible to continue adding edges to construct a rectilinear shape having the target perimeter. Thus, students who meet such situations will construct new shapes until they successfully meet the target perimeter.



Figure 4. Screenshots of the app, Topic: Perimeter and Level: Create.

When the topic *Area and Perimeter* is chosen from the Main Menu, students may perform tasks that involve both concepts, to help them attain the reasoning skills to distinguish between these two concepts (Huang, 2017).

*Explore.* Under the *Explore* level, students first click on each unit square within the figure to show yellow squares and then find the number of unit squares (Figure 5(a)). When the answer is correct, they now click on the edges of the shape to show red edges and then count them (Figure 5(b)). When the answer is correct, the app confirms that the area of the shape is the number of unit squares given and the perimeter is the number of red edges.



Figure 5. Screenshots of the app, Topic: Area and Perimeter and Level: Explore

Apply. In this level, students directly input the area and perimeter of the given shape.

*Create.* In this level, students draw a rectilinear shape with the given area and perimeter (Figure 6(a)). There may be instances when students construct a rectilinear shape with the correct perimeter but incorrect area or correct area but incorrect perimeter (Figure 6(b)). The current area and current perimeter given on the right side of the screen provide the immediate feedback that will enable the students to construct new shapes until they successfully meet the target area and perimeter.



Figure 6. Screenshots of the App, Option: Area and Perimeter and Difficulty Level: Create.

#### 3.2 Game Design Factors

Creating an innovative instructional app in a game-like environment such as *Area and Perimeter* requires a sound framework on which its development can be based. Focusing on how the game is aimed at enabling players to make the leap from the concrete to the abstract, we used Shi and Shih's game-based learning (GBL) design factors (Shi and Shih, 2015) as a guide, apart from "sociality," which *Area and Perimeter* does not inherently feature as it is a single-player game.

The game goal is to develop an understanding of area and perimeter formulas. This is achieved via a combination of the abovementioned *Explore*, *Apply*, and *Create* difficulty levels, which constitute the game mechanism. The player interacts with the game by tapping squares to highlight areas, dragging between points to construct edges, and tapping a simple number

pad to provide answers to area questions. As mentioned above, players are free to select what concept (area or perimeter) and what shape to focus on; the *Explore*, *Apply*, and *Create* difficulties are also available for all these concepts from the beginning. Moreover, in the *Create* difficulty, players are tasked to construct any shape that has particular area and perimeter values, giving them autonomy and testing their creativity. Each concept, shape, and difficulty combine to form an inherently challenging level that requires the player to accomplish a task (e.g., highlight areas, tap edges).

While the *Area and Perimeter* app does not have a narrative by the strictest definition, it does feature a sense of progression through the various topics and difficulties. The game fantasy centered around an interactable grid, which is also the main medium with which the game provides sensation. The game's mystery arises from the variety of the topics and levels, as well as the randomness in the questions generated. These factors taken together give rise to the game value of *Area and Perimeter* as a learning tool for mastering abstract concepts such as area and perimeter formulas using concrete examples.

#### 4. Integration and Use of the Area and Perimeter App

This section explores the use of the *Area and Perimeter* app as a pedagogical tool based on the RAT (Replacement, Amplification, Transformation) framework (Hughes et al., 2006). The app can replace the traditional approach of teaching area and perimeter without modifying the content or learning outcomes. For instance, teachers frequently ask students to calculate the areas and perimeters of hand-drawn figures of triangles, squares, rectangles and other quadrilaterals. These figures are now presented in the lattice configuration of the app. Finding the area or perimeter still must be done, but instead of doing it in a static way, geometric figures are now shown in a colorful and dynamic way. Teachers can use the app during drill and practice time.

Secondly, learning may be enhanced by the integration of the app since it can explain perimeter and area ideas with ease and efficiency. The lattice design accurately represents unit squares, which can help learners comprehend the conceptual underpinnings of area and perimeter calculations. This understanding is also possible since students engage in a variety of activities that the app can generate.

Finally, the app may be used to change instructional methods and the learning process. Teachers may conduct classroom activities that are more student-centered by using the app as a tool for group inquiry and projects that promote cooperation among students and provide a supportive learning environment. The *Create* option has the potential to improve student engagement and alter the way that students learn. Thus, teachers are encouraged to utilize this feature during class discussions and activity periods to expose students to questions about area and perimeter that aren't commonly asked.

Teachers can screencast to a smart TV or projector if students don't have their own devices. The proposed exercises can be completed by the entire class at the same time.

The app's three learning levels—Explore, Apply, and Create—can accommodate various learning demands of the students. While the levels on Explore and Apply can be utilized for remedial work, the level on Create is suitable for student enrichment activities.

#### 5. Conclusion and Future Direction

This paper describes the features of the *Area and Perimeter* app designed to help students gain a better understanding of the notions of area and perimeter and help them make sense of the usual formulas associated with these concepts. The three levels of difficulty of the app leads students from basic exploration of the area or perimeter of polygons to application of these ideas in calculating these parameters, and finally to the creation of shapes with a given area or perimeter.

Efforts to improve numeracy and literacy among Filipino learners are even more relevant at this time considering the recognition by the Department of Education (DepEd) of

the Philippines that there have been setbacks in learning targets caused by school closures during the Covid-19 pandemic. The DepEd issued the *Adoption of the National Learning Recovery Program* (DepEd, 2023) to help address these setbacks. The *Area and Perimeter* app may be considered as a tool that can contribute to this recovery plan.

The Area and Perimeter app may be expanded to include the exploration of areas and perimeters of non-polygonal figures such as circles and other irregular shapes. This will further help students understand the concepts and not resort to sheer memorization of formulas. An avenue for future work would be the translation of the app into Filipino to make it more accessible to students in the Philippines. Finally, the authors plan to conduct further research to evaluate the effectiveness of the app in the learning of students.

#### Acknowledgements

This paper is one of the outputs of the Ateneo Mathplus Resources Laboratory housed at the Department of Mathematics, School of Science and Engineering, Ateneo de Manila University. The authors thank the Department of Science and Technology-Philippine Council for Industry, Energy, and Emerging Technology Research and Development (DOST-PCIEERD) and the University Research Council (URC), Ateneo de Manila University for the support of the development of the mathematical resources for Grades 1 to 11 Mathematics. Many thanks go to our app developers Victor Antonio M. Ortega, Mr. Nigel Benedict Cargo, Mr. Jose Teodoro Lacson, and our artist Mr. Amiel Damian F. Justiniani.

#### References

- Bansilal, S. (2011). Assessment reform in South Africa: opening up or closing spaces for teachers? *Educational Studies in Mathematics, 78,* 91-107.
- Bruner, J. (1966). *Toward a theory of instruction*. Cambridge: Belknap Press.
- De Las Peñas, M. L. A. N., Aberin, M. A. Q., Garciano, A. D., Sarmiento, J. F., Tolentino, M. A. C., Verzosa, D. M. B., Mallari, J. C. (2023). *Terminal report (Department of Science and Technology-Philippine Council for Industry, Energy and Emerging Technology Research and Development): Mathematical resources for distance learning utilizing community LTE networks and television frequencies.*
- Department of Education (2016). *K to 12 curriculum guide: mathematics*. www.deped.gov.ph/wp-content/uploads/2019/01/Math-CG\_with-tagged-math-equipment.pdf
- Department of Education. (2023). Adoption of the National Learning Recovery Program in the Department of Education. www.deped.gov.ph/wp-content/uploads/DO\_s2023\_013.pdf
- Ding, M., & Li, X. (2014). Transition from concrete to abstract representations: the distributive property in a Chinese textbook series. *Educational Studies in Mathematics, 87,* 103-121.
- Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 65-98). New York: Macmillan.
- Huang, H. E. (2017). Curriculum interventions for area measurement instruction to enhance children's conceptual understanding. *International Journal of Science and Mathematics Education, 15,* 1323–1341.
- Hughes, J. E., Thomas, R., & Scharber, C. (2006). Assessing technology integration: The RAT Replacement, Amplification, and Transformation – Framework. In C. M. Crawford, R. Carlsen, K. McFerrin, J. Price, R. Weber, & D. A. Willis (Eds.), *Proceedings of SITE 2006--Society for Information Technology & Teacher Education international conference* (pp. 1616-1620). Orlando, Florida, USA: Association for the Advancement of Computing in Education (AACE).
- Hwang, W., Hoang, A., & Tu, Y. (2020). Exploring authentic contexts with ubiquitous geometry to facilitate elementary school students' geometry learning. *Asia-Pacific Education Researcher, 29,* 269–283.
- Lithner, J. (2017). Principles for designing mathematical tasks that enhance imitative and creative reasoning. *ZDM*, *49*, 937-949.
- Shi, Y.-R., & Shih, J.-L. (2015). Game factors and game-based learning design model. *International Journal of Computer Games Technology*, 2015, 1-11.
- Skemp, R. R. (2006). Relational understanding and instrumental understanding. *Mathematics Teaching in the Middle School*, *12*(2), 88–95. Originally published in *Mathematics Teaching* (1976).

## Design of a Mobile App to Promote Understanding and Fluency in Finding the Equation of a Line

Agnes D. GARCIANO<sup>a,\*</sup>, Maria Alva Q. ABERIN<sup>a</sup>, Ma. Louise Antonette N. DE LAS PEÑAS<sup>a</sup>, Juan Carlo F. MALLARI<sup>a</sup>, Jumela F. SARMIENTO<sup>a</sup>, Mark Anthony C. TOLENTINO<sup>a</sup>, & Debbie Marie B. VERZOSA<sup>b</sup> <sup>a</sup>Department of Mathematics, Ateneo de Manila University, Philippines <sup>b</sup>Department of Mathematics and Statistics, University of Southern Mindanao, Philippines \*agarciano@ateneo.edu

**Abstract:** This paper focuses on the design of a mobile app called *Pick or Fish* that fosters comprehension and mastery of the concepts of slopes, *y*-intercepts and equations of lines. The app's pedagogical value lies in its potential to help students understand and become proficient in these concepts. The app is suitable for use on low-cost mobile devices. It functions within an engaging game-like setting featuring visual elements that enable students to see the effect of parameter changes on the direction of a line. The beginner and advanced levels of the app have scaffolding features that gradually introduce the students to the key aspects of linear functions. The mechanics of the app, its pedagogical basis and how integration in the classroom may be achieved as teachers plan the lesson, facilitate open-ended discussion and encourage independent use of the app are also discussed.

**Keywords:** Lines, slopes, *y*-intercept, equations of lines, mobile app

#### 1. Introduction

There is a view that from both a mathematical and pedagogical perspective, the core and fundamental concept underlying algebra, trigonometry, probability, statistics and calculus is that of *function* (Schwartz & Yerushalmy,1992). In the Philippine high school mathematics curriculum, linear function is one of the first types of functions that students learn. In Grade 8, the topics of slope, *y*-intercept, and equation of a line are discussed. An in-depth understanding of these concepts is necessary in problem solving and mathematical modeling involving linear functions. Furthermore, this is also necessary for doing more advanced work in linear regression, interpolation, and rates of change in calculus. The skill of writing equations of lines enables students to make correct mathematical models of real-world data exhibiting linear trends and make forecasts. Thus, the acquisition of conceptual knowledge and procedural fluency on calculations involving linear functions are important for high school students to have.

In this paper, we present the design of a mobile app called *Pick or Fish* that is intended to help students gain a better understanding of slope and y-intercept and acquire fluency in forming the equation of a line. The app is one of the technological tools developed and designed by the authors made available for use in Philippine schools through the internet, community LTE networks (De Las Peñas et al., 2022) or local networks powered by datacasting technology (De Las Peñas et al., 2023). These tools support the twin goals of mathematics education in the Philippines, which is to develop critical thinking and problem solving (DepEd, 2016). In emerging economies such as the Philippines, studies have shown that mathematics is only superficially learned (Verzosa & Vistro-Yu, 2019), and that the focus of mathematics classrooms is on rules and procedures (Nag et al., 2014; Verzosa, 2020).

*Pick or Fish* was designed as a mathematical mobile application (app) that runs in an interactive, game-like environment. It is intended to be played by students to gain proficiency

in finding slopes and forming equations of lines. Interactive digital tools such as mobile apps designed in a game -like environment have the potential to enhance the interest of students (Chao et al, 2018). Interest to carry out the tasks is important for students to gain mastery and fluency in the topic. While there are other existing apps and software that allow the exploration of slopes and equations of lines, one of the important features of *Pick or Fish* is that it is built in a game like environment, and there is the inclusion of scaffolding mechanisms that allow a student to work at a particular level of difficulty based on their learning progression on the topic. The latter feature makes the app apt for student centered learning. As an emerging economy, the Philippines has also comparatively limited educational resources. Further, many Filipino students come from low-income families who could not afford graphics calculators, laptops/computers or medium- to high-range mobile phones. Thus, the *Pick or Fish* was designed to be freely available and able to run on earlier Android versions of mobile phones.

#### 2. Description of the Pick or Fish App

The app is designed to help students develop their understanding of the concept of slope and *y*-intercept of a line and consequently help build their skill in forming the equation of a line. There are two levels: *Beginner* and *Advanced* and there are several topics per level. In each level, the task is to pick a fruit or fish by either varying the value of the slope, or the *y*-intercept, or both.

#### 2.1 Beginner Level

In all topics in the *Beginner* level, the screen shows a hand with a rod and a target object (either a mango fruit or a fish). The student needs to press the "up" and "down" buttons found at the bottom of the screen to move the rod so that it would touch the fruit or the fish. Doing this, the student can see how adjusting the slope corresponds to changing the steepness of the line as represented by the rod. By playing the game repeatedly, the student can connect positive and negative slopes to the orientation of the line (Figures 1(a) and 1(b), respectively).



Figure 1. Screenshots of the Beginner level of Pick or Fish

The app scaffolds instruction by first focusing on slopes having integer values, then fractional values. The student can see how the slope also corresponds to the ratio of the vertical and horizontal changes in the line. For example, in Figure 1(c), the slope is -5/3 which corresponds to the ratio in a triangle with vertical dimension 5 and horizontal dimension 3.

In the succeeding topic, *Slopes—Equation of a Line,* an equation of a line in the form y = mx + b is given, but only the value of *m* needs to be changed (Figure 1(d)). This is the first introduction of the student to the slope-intercept form of the equation of a line. Familiarity with this form is important as it shows that a line is determined by two parameters *m* and *b*. The next topic is the *y*-intercept. where the student presses the up and down buttons to get the correct value of the *y*-intercept (Figure 1(e)). Here, the student sees how changing the *y*-

intercept will not change the steepness but would rather shift the line up or down. The last topic under the Beginner Level is the *Slope-Intercept Form of a Line*. Here, the student provides the values of the slope m and y-intercept b in the equation y = mx + b.

#### 2.2 Advanced Level

In the *Advanced* level, the student can utilize their developing notions of slope and *y*-intercept to make the line reach the targeted mango or fish. The topics in this level are like those in the *Beginner* level, except that the rod is not visible at the start–only the hand and target are shown (Figure 2(a)). Thus, the student must apply what they learned in the *Beginner* level and press the necessary buttons (without visual scaffolding) to produce the correct equation that would make the rod reach the target (Figure 2(b)).



*Figure 2.* Screenshots of the *Advanced* level, Topic: *Slope-intercept Form of a Line* rod is not shown at the start; (b) rod is shown after pressing "Confirm" button.

A student who makes a mistake has the chance to correct himself. For instance, when working on a sample question related to Slope-Intercept Form of a Line, both slope and *y*-intercept values are needed. The student may correctly input the value of the *y*-intercept, which is visible on the screen. However, determining the value of the slope requires more thought. Without visual help he must calculate the ratio of the vertical and horizontal distances between the hand and the object. If the student enters an incorrect value, he receives feedback that the rod misses passing through the object. This feedback serves as a clue, helping the student identify whether the slope must be increased or decreased and gives him an opportunity to input the correct answer.

#### 2.3 Topic/Level Progression

The scaffolded questions built into the app help students to work at the appropriate level of difficulty. Students can go back to any of the levels and topics at any time if they need more practice, or they can do so as guided by their teacher. In each level, once the student presses "Confirm," the app provides visual feedback and a bell or buzzer sound for the student to know if the answer entered is right. In addition, a score found at the top part of the screen enables students to monitor their progress and explore other topics in the app. These features help learners gauge their level of mastery and guide them as they progress through the topics.

#### 3. Pedagogical Basis

The *Pick or Fish* app was designed to elucidate how variables express a relationship in a linear equation. The concept of *variable* is not easily learned because students may not have limited conceptions of how variables are used in mathematics. For Küchemann (1981, p. 104), a

variable represents "a range of unspecified values, and a systematic relationship is seen to exist between two such sets of values." Ely and Adams (2012) identified two critical properties from Küchemann's definition: First, a variable is indeterminate, and can stand for a range of values such as in "y = x + 7," and second, it is part of a systematic relationship such that when one variable changes, another quantity may vary with it. Drawing on the history of mathematics, they further argue that the development of the notion of a variable represented a major breakthrough from using letters merely to represent an unknown or determinate quantity, such as in the equation "2 + x = 5." This development took centuries, making it understandable why students struggle with the transition from unknown to variable—they think that a letter must represent a specific value (Ely & Adams, 2012). This is possibly why students find it difficult to solve tasks involving representational transitions from a graph to an equation or vice versa (Ceuppens et al., 2018).

The visual scaffolds in the *Pick or Fish* facilitate the cognitive shift required to help students see relationships between the equation of a line to its graph and gain a better understanding of the systematic relationship between the variables in the equation. This is done through a game environment, particularly in the *Beginner* level, where students can see how adjusting coefficients in a linear equation leads to changes in the corresponding graph. These scaffolding elements which are introduced gradually through the progression of topics, along with visual aids and feedback mechanisms, help students cultivate a strong conceptual understanding of slope and *y*-intercept concepts, as well as to improve their fluency and proficiency in forming the equations of lines.

#### 4. Game Design Features

Shi and Shih's (2015) Game-Based Learning (GBL) Design Model was used as the framework for the development of *Pick or Fish*. This model highlights 11 game design factors (*game goal, game mechanism, interaction, freedom, challenge, game fantasy, narrative, sociality, sensation, mystery, and game value*) for the design, analysis, and evaluation of game-based learning applications. Ten of these factors were incorporated into *Pick or Fish* and are discussed in detail below.

The game goal for Pick or Fish is to deepen one's understanding of slopes and the formulation of the equations of lines. This is achieved via its drill-like game mechanism where the player must provide the correct slope, *y*-intercept or equation of a line that passes through two given points (the hand and the fruit or fish). This drill-like mechanism complements the game goal as it allows players to repeatedly practice with different questions or tasks provided in the app to develop their fluency in the topic. Contributing to the game mechanism are the factors *interaction, freedom, and challenge.* Players *interact* with the game by tapping buttons to increase or decrease the slope and/or *y*-intercept of a line. They are given the *freedom* to select from multiple topics and levels that, as previously mentioned, allow for scaffolded instruction. These levels *challenge* the player to correctly answer five questions in succession, after which they will receive a prompt to move on to the next topic. These elements are important for developing learners' fluency in the topic or level until they have developed sufficient mastery) or to address specific learning gaps (i.e., selecting the topic or level that corresponds to a competency that they need to improve).

As *Pick or Fish* can be considered a single-player arcade game, it places no emphasis on *sociality* and features a simple *game fantasy* and *narrative* in which players are tasked to pick fruits off trees or catch fish by correctly identifying the equations of lines. The *game fantasy* is then further enhanced by the game's *sensation* and *mystery* aspects. For *sensation*, the game uses a simple but vibrant background with an overlain Cartesian plane, as well as sound prompts to let the player know when they input a correct or incorrect answer. It must be noted that on this Cartesian plane, for the slope topics under the game's *Beginner* level, adjusting the slope not only changes how the line is tilted but also changes the line's "rise" and "run" as represented by the appropriate triangle. This topic or level progression adds *mystery* to the game, with each new topic or level providing a new experience to the player. Finally, the *game value* arises from the combination of all these factors, as *Pick or Fish* is aimed at providing players an engaging environment in which they can develop their understanding of equations of lines.

#### 5. Integration and Use of the Pick or Fish App

This section describes how *Pick or Fish* can be utilized as a pedagogical tool. The steps below were adapted from the work of Woods et al (2018).

Step 1. Determine learning goals. The app is appropriate and useful in attaining the learning competencies involving slope, *y*-intercept and equation of a line. The design makes it easier for students to make the shift needed to see connections between a line's equation and graph and to comprehend the systematic link between the variables in an equation.

Step 2. Plan ahead. The teacher should plan for the logistical requirements of using the free to download app. Students should be taught how to properly download the software on a desktop or an Android device by their teachers. An Android emulator is required to run the program if it was downloaded to a PC.

Step 3. Present and practice. This step entails the presentation of the lesson and a demonstration on how to play the app. The teacher can ask, "Should the rod be moved up or down, left or right?" or "What happens to the direction of the rod when the up button is clicked?" These instruction questions will prevent students from using guesswork as a strategy to find the slope, *y*-intercept and equation of the line. Students have the chance to practice and become comfortable with the app during this phase.

Step 4. Process and verbalize. Finally, open-ended questions will allow students to explain their thinking process and reasoning. This can be done by asking reflection questions like "What helps you determine which button to click?", "How do you know that slope should be positive? or negative?", "What is the meaning of *y*-intercept?", or "How can one form the equation of a line?"

For practice and mastery, the teacher may advise the pupils to utilize the app asynchronously or outside of class. Students can share screenshots of their outputs over time in order to monitor their development. A follow-up discussion can be done to address difficulties and misconceptions of the students.

#### 6. Conclusion and Future Direction

A deep conceptual understanding and fluency in creating equations of lines is a basic learning competency expected of high school students. However, many students still find the ideas of slope and *y*-intercept abstract and lack the proficiency in using these elements to formulate the equation of a line. This problem is carried on to more advanced work in mathematics where the concept of linear functions and equations are necessary.

In this paper, the *Pick* or *Fish* app is introduced as an interactive tool that can help students gain more insight and fluency in finding the slope and *y*-intercept of a line and write its equation. The app is aligned with the Philippine Department of Education's Most Essential Learning Competencies in Grade 8 Mathematics (DepEd, 2020). The app's game-like features help motivate learning and entice students to play to improve their understanding of abstract concepts. The popularity of mobile phones and tablets in the Philippines makes the app easily accessible as it can be readily downloaded and require low storage.

The next step in this research is to conduct studies on determining the effectiveness of the app in student learning. The app may still be improved by the inclusion of undefined slopes of vertical lines. An aspect that may be pursued in the future is the incorporation of more scaffolding mechanisms that can offer help to students who may struggle with the transition from beginner to advanced levels. Gathering more feedback from students and teachers on the use of the app will also help direct future research in this area.

#### Acknowledgements

This paper is a product of the research work of the Ateneo Mathplus Resources Laboratory of the Department of Mathematics, School of Science and Engineering, Ateneo de Manila University (ADMU). The authors express their gratitude to the Department of Science and Technology-Philippine Council for Industry, Energy, and Emerging Technology Research and Development and the University Research Council, ADMU for their valuable support in the development of the mathematical resources made for Grades 1 to 11 Mathematics. We also thank our app developers Victor Antonio M. Ortega, Mr. Nigel Benedict Cargo, Mr. Jose Teodoro Lacson, and Mr. Amiel Damian F. Justiniani for the apps and video art assets.

#### References

- Ceuppens, S., Deprez, J., Dehaene, W., & Cock, M. D. (2018). Design and validation of a test for representational fluency of 9th grade students in physics and mathematics: The case of linear functions. *Physical Review Physics Education Research*, *14*, 02015.
- Chao, W.-H., Yang, C.-Y., Hsien, S.-M., & Chang, R.-C. (2018). Using mobile apps to support effective game-based learning in the mathematics classroom. *International Journal of Information and Education Technology*, *8*(5), 354-357.
- De Las Peñas, M. L. A. N., Aberin, M. A. Q., Garciano, A. D., Mallari, J. C. F., Sarmiento, J. F., Tolentino, M. A. C., & Verzosa, D. M. B. (2022). Deployment of mathematical resources to a Philippine high school through a community LTE network. In W. C. Yang, M. Majewski, D. Meade, & W. K. Ho (Eds.), *Proceedings of the 27th Asian technology conference in mathematics* (pp. 126-140). Mathematics and Technology, LLC.
- De Las Peñas, M. L. A. N., Kwong, J. C. T., Banzon, P. A. B., Martinez, P. A., Lapada, W. I. G., Eballa, J. N., Sebastian, J. E. C., Asido, M. D., San Juan, J. D. M., Versoza, D. M. B., Sarmiento, J. F., Garciano, A. D., Tolentino, M. A. C., Aberin, M. A. Q., & Mallari, J. C. F. (2023). Mathematical mobile apps via Rural Casting. In P. Kommers, I. A. Sánchez, & P. Isaías (Eds.), *Proceedings of the international conferences on e-Society 2023 and mobile learning 2023* (pp. 343-350). International Association for Development of the Information Society.
- Department of Education. (2016). *K to 12 Curriculum Guide: Mathematics (Grade 1 to Grade 10)*. https://www.deped.gov.ph/wp-content/uploads/2019/01/Math-CG\_with-tagged-mathequipment.pdf
- Department of Education. (2020). *K to 12 most essential learning competencies with corresponding CG codes.*
- Ely, R., & Adams, A. E. (2012). Unknown, placeholder, or variable: what is x? *Mathematics Education Research Journal*, 24(1), 19–38.
- Küchemann, D. E. (1981). Algebra. In K. M. Hart (Ed.), *Children's understanding of mathematics* 11– 16 (pp. 102-119). London: John Murray.
- Nag S., Chiat S., Torgerson C., & Snowling, M. J. (2014). Literacy, foundation learning and assessment in developing countries: final report. *Education Rigorous Literature Review*. Department for International Development, United Kingdom.
- Schwartz, J., & Yerushalmy, M. (1992). Getting students to function on and with algebra. In E. Dubinsky & G. Harel (Eds.), *The concept of function: aspects of epistemology and pedagogy* (pp. 261–289). Mathematical Association of America.
- Shi, Y.-R., & Shih, J.-L. (2015). Game factors and game-based learning design model. *International Journal of Computer Games Technology, 2015*, 1-11.
- Verzosa, D.M.B. (2020). Teachers' approaches and drawn representations in explaining solutions to fraction tasks. Manuscript in preparation.
- Verzosa, D.M.B., & Vistro-Yu, C.P. (2019). Prospects and challenges in implementing a new mathematics curriculum in the Philippines. In C.P. Vistro-Yu & T.L. Toh (Eds.), *School mathematics curricula* (pp. 207-226). Singapore: Springer.
- Woods, D.M., Geller, L.K., & Basaraba, D. (2018). Number sense on the number line. *Intervention in School and Clinic*, 53(4), 229-236.

## Towards Identifying the Learning Affordances of Social Media as Telemedicine Platforms among Physicians in a Developing Economy

#### Michelle BERNABE<sup>a,b\*</sup>, Rowena CAGUIAT<sup>a</sup> & Ryan EBARDO<sup>a</sup>

<sup>a</sup>De La Salle University Manila, Taft Avenue, Manila, Philippines <sup>b</sup>University of St. La Salle, Bacolod City, Philippines \*michelle bernabe@dlsu.edu.ph

**Abstract:** This preliminary study investigates the integration of social media platforms in telemedicine from the viewpoint of physicians, utilizing the affordance theory. Through qualitative interviews with eight experienced physicians, the research explores the learning affordances of social media in healthcare in the Philippines. The findings reveal that social media facilitates digital information exchanges through remote communication, real-time symptom verification, and virtual fetal monitoring, leading to enhanced patient-provider interactions. However, certain restrictions, such as prior checkups and patient referrals, are identified to ensure responsible healthcare practices. The study provides essential insights for policymakers and healthcare professionals to optimize social media's role in telemedicine, revolutionizing healthcare delivery and improving patient outcomes. By understanding the specific learning affordances and challenges, this research contributes to enhancing the integration of social media platforms in healthcare practices and fostering patient-centered solutions through improved exchange of information between physicians and patients.

Keywords: Telemedicine, Affordance Theory, Social Media, Facebook, health technology

#### 1. Introduction

Incorporating social media into telemedicine has been the subject of numerous studies in the recent past (O'Connor & Aardema, 2019; Wu, Chen, & Lin, 2020; Lee & Yoo, 2021). O'Connor and Aardema (2019) note that these platforms present possibilities to improve the efficiency and effectiveness of healthcare delivery via telemedicine. Identifying the potential benefits and risks of this developing trend requires understanding the role of affordance theory in the context of social media as telemedicine platforms (O'Connor & Aardema, 2019; Wu, Chen, & Lin, 2020; Lee & Yoo, 2022).

The use of social networking sites as part of telemedicine has the potential to revolutionize patient care. Social media allows doctors to remotely consult with patients, check their health, and disseminate informational resources (Wu, Chen, & Lin, 2020). Patient involvement and satisfaction can be increased, healthcare access expanded, and geographical obstacles removed with this integration (Wu, Chen, & Lin, 2020).

Despite the potential benefits, there are knowledge gaps regarding the specific affordances of social media platforms in the context of telemedicine from the physician's perspective (Smith, Johnson, & Davis, 2022). Most research has ignored the unique viewpoints and experiences of doctors who utilize social media as telemedicine platforms in favor of focusing on patient-centered outcomes and user perspectives (Smith, Johnson, & Davis, 2021). Telemedicine's benefits and drawbacks can be better understood if we gain insight into clinicians' perspectives and experiences using social media platforms for telehealth (Smith, Johnson, & Davis, 2022).

The main objective of this research is to examine, from a medical practitioner's vantage point, how social media can be used to take advantage of their educational affordances in the context of telemedicine. To determine the advantages and disadvantages of incorporating social media into telemedicine (Brown, Williams, & Wilson, 2022; Chen, Chen, & Lin, 2023) and the specific learning affordances that influence physicians' perceptions and practices. A qualitative method was used to learn more about how doctors use social media platforms for telemedicine (Brown, Williams, and Wilson, 2022; Chen, Chen, and Lin, 2023) and to record their rich experiences.

This study will contribute to healthcare policymaking by identifying the learning opportunities offered by social media in telemedicine. Additionally, it will expand our understanding of social media's potential benefits in resource-constrained healthcare settings, as it is conducted in a developing economy. Lastly, using a qualitative approach, the study delves deeper into the learning aspect of affordance theory within the context of social media as a telemedicine platform.

Section 2 summarizes existing research on the use of social media as a platform for telemedicine. The third section investigates the applicability of affordance theory to social media. Discussion follows in Section 4. In the final section, we discuss the preliminary findings of this ongoing research and its potential limitations.

#### 2. Related Review of Literature

O'Connor and Aardema (2019), Wu, Chen, and Lin (2020), and Lee and Yoo (2021) emphasize the potential revolutionary impact of integrating social media into telemedicine. Notably, O'Connor and Aardema (2019) and Smith et al. (2022) argue that social media platforms offer specific affordances that can enhance the effectiveness and efficiency of telemedicine, particularly for healthcare professionals like physicians. To comprehensively understand the potential benefits and challenges of this innovative approach, it is essential to consider the perspectives of physicians regarding these affordances and their influence on healthcare practices.

Social media's affordances within telemedicine are diverse and extensive. As highlighted by Wu, Chen, and Lin (2020), social media platforms enable remote interactions, allowing physicians to establish virtual connections with patients, offer medical guidance, and monitor their health status. These platforms also facilitate the exchange of medical knowledge, educational resources, and health-related updates, enabling the dissemination of critical information (Lee & Yoo, 2021). Additionally, social media platforms serve as valuable tools for creating networks among physicians, fostering collaboration, and facilitating discussions on specific cases or the solicitation of expert opinions from peers (O'Connor & Aardema, 2019). Furthermore, the interactive features inherent in these platforms promote patient engagement and empowerment (O'Connor & Aardema, 2019).

While the integration of social media into telemedicine offers substantial advantages, it also presents challenges. Wu, Chen, and Lin (2020) underscore the potential benefits, including expanded healthcare service accessibility, increased patient participation and satisfaction, and the ability to overcome geographical limitations. Nevertheless, integrating social media platforms into telemedicine requires addressing significant concerns, including patient privacy and confidentiality, data security, and the potential for misinterpretation or misdiagnosis during virtual consultations (Chen, Han, & Cheng, 2022). In this context, healthcare professionals must effectively manage regulatory and ethical considerations associated with the use of social media platforms (Chen, Han, & Cheng, 2022).

To optimize the utilization of social media platforms in telemedicine, healthcare professionals and policymakers should establish explicit guidelines and protocols for patient validation and referrals, ensuring the continuity of care and minimizing the potential risks of medical malpractice (Bergmo, 2015; Desai et al., 2020; Hollander & Sage, 2020; Ohannessian et al., 2016). Moreover, providing training and support to healthcare professionals in the appropriate use of social media platforms can enhance their competence (Bergmo, 2015; Desai et al., 2020; Hollander & Sage, 2020; Ohannessian et al., 2016). Integrating telemedicine education into the medical curriculum can better prepare future healthcare professionals to efficiently leverage the advantages offered by this

technology (Bergmo, 2015; Desai et al., 2020; Hollander & Sage, 2020; Ohannessian et al., 2016).

#### 3. Theoretical Foundations

The Affordance Theory which delves into the perceived opportunities and capabilities that are presented by social media platforms to healthcare professionals within the area of telemedicine. This study seeks to get valuable insights into physicians' motives and experiences regarding using social media as a telemedicine tool through the theoretical lens of affordances to advance healthcare delivery in this study's context.

Gruzd, Staves, and Wilk (2017) argue that the utilization of digital technology in social media platforms holds the capacity to revolutionize the field of health and social care. These platforms offer many chances for remote contact, sharing of information, establishment of support networks, utilization of collaborative features, access to educational resources, and engagement with other interactive activities. The aforementioned affordances are of paramount importance in augmenting the provision of healthcare, surmounting geographical impediments, enhancing accessibility to healthcare services, and bolstering patient involvement and contentment.

This study utilized the Affordance Theory to enhance comprehension of the perceived opportunities and capabilities provided by social media platforms to physicians. These platforms enable remote communication, information sharing, support networks, and other interactive features that contribute to the improvement of healthcare delivery. The theoretical framework's emphasis on the capabilities offered by technology is congruent with the research's aim to investigate the utilization of social media as a telemedicine tool from the viewpoint of medical professionals. This approach facilitates a thorough analysis of the potential advantages and obstacles associated with the platform.

#### 4. Methodology

This study will employ a qualitative descriptive approach with thematic analysis, drawing from methodologies articulated by Sandelowski (2000), Vaismoradi et al. (2013), and Braun & Clarke (2006). The purpose is to deeply understand participants' experiences in their natural settings. Thematic analysis, as detailed by Braun and Clarke (2006), will assist in identifying patterns and themes in the data. Purposive sampling will select eight physicians with diverse expertise based on criteria such as consistent use of social media for telemedicine and a minimum active period of two years in this domain (Palinkas et al., 2015) which is presented in Table 1.

Data collection will involve semi-structured interviews with the physicians, adhering to the approach delineated by Smith et al. (2018). Interviews will be audio-recorded and transcribed verbatim. An interview guide, derived from study questions, will explore topics like functionalities of social media platforms, influencers on physician practices, and the pros and cons of integrating social media into telemedicine. The analytical procedure, as explained by Braun & Clarke (2006), will span from understanding the data to constructing and categorizing themes, supported by participant statements. After filtering, eight pertinent codes were identified and grouped into four distinct themes, further detailed in the following sections.

Participant ID	Specialization	Gender	Age	Years using Social Media Telemedicine
001	OB-GYN	Female	34	3
002	Internal Medicine	Male	42	3
003	Family Medicine	Female	36	3
004	Pediatrician	Female	58	3
005	Pulmonology	Male	43	3
006	Allercology	Female	46	6

Table 1 Participants

007	Opthamology	Male	37	10
008	Opthamology	Male	42	5

#### 5. Preliminary Results and Limitations

This results section, framed by the affordance theory, examines the impact of social media as a telemedicine platform on healthcare professionals and physicians. It specifically highlights the features and capabilities of social media platforms and how they influence physicians' practices. Exploring these affordances reveals how social media improves healthcare delivery, overcomes geographical limitations, enhances access to services, and boosts patient engagement. To illustrate these findings, the researchers provide representative quotes from the interviews and detailed themes in Table II - Resulting Themes for clarity and insight.

Theme	Codes	Exemplar		
Validation	Real-Time Symptoms Verification	"I find it helpful that I can easily validate and verify a patient's symptoms in real-time using telemedicine through video call" – Case 2		
	Virtual Fetal Monitoring	"I use virtual fetal monitoring during prenatal checkups to ensure the baby's health." - Case 1		
Restrictions	Prior Checkup	"I only accept telemedicine consultations if I have previously checked up on the patient." – Case 6		
	Patients Referral	"I accept telemedicine consultations only when		
Fase of Use	Effortless Access	"easy access and hassle-free way to connect with patients, enabling us to provide medical care and support without the limitations of physical distance." – case 5		
	Comfortable Engagement	"I can engage with my patients in a way that feels comfortable and caring, ultimately enhancing the quality of their healthcare experience." – Case 4		
Backtracking	Medication History	"Telemedicine enables me to efficiently track and review my patient's medication history, ensuring that I can make well- informed decisions and avoid potential drug interactions during our remote consultations." – Case 3		
	Medical Record Retrieval	"I can easily access and review my patients' medical history, allowing me to provide more personalized and informed care during our virtual consultations." – Case 5		

Table 2. Resulting Themes

#### 5.1 Validation

Telemedicine enhances healthcare via remote technologies such as accurate patient information validation, real-time symptom verification, and virtual fetal monitoring during prenatal checkups. These capabilities enable physicians to deliver effective care remotely, overcoming geographical limits and enhancing patient outcomes (Bergmo, 2015; Desai et al., 2020; Kreps et al., 2019).

#### 5.2 Restrictions

Telemedicine offers various affordances, including remote consultations, information sharing, and establishing social connections with patients. However, prior checkups and patient referrals are often required to ensure responsible healthcare practices. These measures help validate patient information, maintain continuity of care, and minimize medical malpractice risk. By balancing the benefits and restrictions of telemedicine, healthcare professionals can optimize its use while upholding quality and safety standards (Bergmo, 2015; Desai et al., 2020; Hollander & Sage, 2020; Ohannessian et al., 2016).

#### 5.3 Ease of Use

Telemedicine platforms offer ease of use, enabling smooth interactions between healthcare professionals and patients (Sundararaman et al., 2017). Their accessibility allows patients to connect from anywhere with internet, enhancing patient experiences and communication (Wade et al., 2017; Donelan et al., 2019). Such benefits can revolutionize healthcare and improve outcomes (Bashshur et al., 2016; Dorsey et al., 2017).

#### 5.4 Backtracking

Telemedicine's affordances, such as backtracking of treatment plans (Bhavnani et al., 2018), medication history reviews (Gagnon et al., 2016), and swift medical record access (Lau et al., 2018), enhance healthcare. These affordances streamline workflows, improve diagnostic accuracy, and benefit both patients and providers (Terry et al., 2017; Wade et al., 2020).

#### 6. Conclusion and Recommendations

Telemedicine, leveraging social media platforms, revolutionizes healthcare by offering affordances like remote consultations and virtual monitoring (O'Connor & Aardema, 2019; Wu, Chen, & Lin, 2020; Lee & Yoo, 2021). These features improve patient access and engagement while overcoming geographical limitations (Gibson, 1979). Nevertheless, there are inherent challenges in this integration, necessitating measures such as prior checkups to maintain ethical healthcare (Bergmo, 2015; Desai et al., 2020; Hollander & Sage, 2020; Ohannessian et al., 2016).

To optimize the use of social media in telemedicine, healthcare professionals and policymakers can establish clear guidelines and protocols for patient validation and referrals, ensuring continuity of care and minimizing the risk of medical malpractice (Bergmo, 2015; Desai et al., 2020; Hollander & Sage, 2020; Ohannessian et al., 2016). Additionally, providing training and support to healthcare professionals in effectively utilizing social media platforms for telemedicine can enhance their proficiency (Bergmo, 2015; Desai et al., 2020; Hollander & Sage, 2020; Ohannessian et al., 2016). Integrating telemedicine education into medical curricula can further equip future healthcare providers with the skills to leverage these affordances effectively (Bergmo, 2015; Desai et al., 2020; Ohannessian et al., 2016).

Future research should delve into the affordances of social media in telemedicine for both patients and professionals (Smith et al., 2022). Longitudinal studies are needed to evaluate its long-term effects on healthcare outcomes and experiences. With technological shifts and policy changes, continuous investigation is vital to guide evidence-based practice in this evolving domain (Smith et al., 2022). Such endeavors can lead to enhanced, patient-focused telemedicine through social media.

#### 7. References

Bashshur, R. L., Shannon, G. W., Bashshur, N., & Yellowlees, P. M. (2016). The Empirical Evidence for Telemedicine Interventions in Mental Disorders. Telemedicine and e-Health, 22(2), Bergmo, T. S. (2015). An economic analysis of telemedicine: A systematic review. Journal of

Telemedicine and Telecare, 21(6), 323-331. doi:10.1177/1357633X15574087 Brown, E., Williams, J., & Wilson, M. (2022). Understanding the perceptions and experiences of physicians in utilizing social media platforms for telemedicine: A qualitative investigation. Health Communication, 37(2), 209-217. doi:10.1080/10410236.2020.1811617

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77-101.
- Chen, P., Han, X., & Cheng, C. (2022). The impact of social media affordances on the effectiveness and efficiency of healthcare delivery in telemedicine. Journal of Health Communication, 27(6),
- Chen, S., Chen, Y., & Lin, C. (2023). Exploring the specific affordances of social media platforms in the context of telemedicine: A qualitative study with physicians. Journal of Telemedicine and Telecare, 29(1), 53-63. doi:10.1177/1357633X23946157
- Desai, T., Shariff, A., & Shariff, A. (2020). Telemedicine and remote monitoring: Paradigm shift of healthcare in COVID-19 era. Journal of Family Medicine and Primary Care, 9(8), 3783-3787.
- Donelan, K., Barreto, E. A., Sossong, S., Michael, C., Estrada, J. J., Cohen, A. B., & Wozniak, J. (2019). Patient and Clinician Experiences With Telehealth for Patient Follow-up Care. American Journal of Managed Care, 25(1), 40–44.
- Dorsey, E. R., Topol, E. J., & Telemedicine Study Group. (2017). State of Telehealth. New England Journal of Medicine, 375(2), 154–161. doi:10.1056/NEJMra1601705
- Gibson, J. J. (1979). The ecological approach to visual perception. Houghton Mifflin Harcourt.
- Gruzd, A., Staves, K., & Wilk, A. (2017). Connected health: Perceptions, experiences, and opportunities. In A. Gruzd, B. Wellman, & Y. Takhteyev (Eds.), The Oxford Handbook of Social Media and Society (pp. 403-424). Oxford University Press.S
- Hollander, J. E., & Sage, J. I. (2020). Telemedicine in the Age of COVID-19. Journal of the American Medical Association, 323(16), 1693–1694. doi:10.1001/jama.2020.4683
- Johnson, R., Ramanathan, S., & Williams, A. (2023). Understanding the role of social media in telemedicine: A comprehensive review. Journal of Telemedicine and Telecare, 29(1), 32-42.
- Kreps, G. L., Neuhauser, L., & Kreps, G. L. (2019). New directions in eHealth communication: Opportunities and challenges. Patient Education and Counseling, 92(2), 273-275.
- Lee, J., & Yoo, J. (2021). The role of social media in telemedicine: A systematic review. International Journal of Medical Informatics, 146, 104331. doi:10.1016/j.ijmedinf.2020.104331
- Lee, S., Kim, J., & Park, H. (2023). Optimizing the use of social media platforms in telemedicine: Practical recommendations for healthcare professionals, policymakers, and technology developers. International Journal of Medical Informatics, 157, 105073.
- Liamputtong, P. (2010). Qualitative research methods (4th ed.). Oxford University Press.
- Lim, S., & Kim, H. (2023). Bridging the gap in knowledge: A comprehensive understanding of the affordance theory in the use of social media as telemedicine platforms. Journal of Health Communication, 28(4), 346-356. doi:10.1080/10810730.2022.2039615
- Moorhead, S. A., Hazlett, D. E., Harrison, L., Carroll, J. K., Irwin, A., & Hoving, C. (2013). A new dimension of health care: Systematic review of the uses, benefits, and limitations of social media for health communication. Journal of Medical Internet Research, 15(4), e85.
- O'Connor, P. J., & Aardema, F. (2019). Affordances: Exploring the relationship between mental and relational affordances in social media platforms. Journal of Computer-Mediated Communication, 24(2), 71-88. doi:10.1093/jcmc/zmz004
- Ohannessian, R., Duong, T. A., & Odone, A. (2016). Global Telemedicine Implementation and Integration Within Health Systems to Fight the COVID-19 Pandemic: A Call to Action. JMIR Public Health and Surveillance, 6(2), e18810. doi:10.2196/18810
- Rodriguez, A. R., Davis, H. A., & Anderson, J. L. (2021). Benefits and challenges of integrating social media into telemedicine: A systematic review. Journal of Medical Internet Research, 23(2
- Smith, J. A., Flowers, P., & Larkin, M. (2009). Interpretative phenomenological analysis: Theory, method, and research. Sage Publications.
- Smith, J., Johnson, A., & Davis, H. (2022). Exploring the affordance theory in the use of social media as telemedicine platforms: A qualitative study. Telemedicine Journal and E-Health, 28(3).
- Sundararaman, L. V., Edwards, P., & Mount-Campbell, A. (2017). Acceptance of Telemedicine in Remote Locations: An Empirical Test. Journal of Health Informatics in Developing Countries
- Thompson, C., & Davis, R. (2023). Informing healthcare professionals, policymakers, and technology developers about the potential of social media platforms in enhancing telemedicine practices: Insights from a qualitative study. Health Informatics Journal, 29(1), 85-95.
- Wu, T., Chen, P., & Lin, Y. (2020). Integrating social media platforms into telemedicine practices: Potential benefits and challenges. Journal of Medical Internet Research, 22(11), e22155.

## Analyzing Sentiments of ChatGPT Users: Philippine Setting

Gisele DIZON<sup>a</sup> & Dr. Ma. Rowena R. CAGUIAT<sup>b</sup>

<sup>a</sup>De La Salle University, Philippines <sup>b</sup>De La Salle University, Philippines <sup>\*</sup>gisele\_dizon@dlsu.edu.ph

**Abstract:** The research is focused on a statistical examination of data from Twitter, a popular social networking platform where lively debates may be exchanged on a wide range of topics, including ChatGPT. The primary goal of this research is to conduct sentiment analysis on the ChatGPT-related tweets in order to get insight into how Twitter users in the Philippines feel about this AI-powered language model. In addition, the study delves into the n-grams and most commonly used words associated with AI, highlighting the preponderance of AI-related phrases and their practical significance. However, 16.57% of tweets expressed negative sentiments, suggesting areas for improvement and privacy concerns. In conclusion, this research provides valuable insights into user sentiments surrounding ChatGPT during a period of significant milestones and growth. The findings offer guidance to developers and stakeholders for addressing potential challenges and further enhancing ChatGPT's success in the ever-evolving AI landscape.

Keywords: ChatGPT, sentiment analysis, artificial intelligence, NLP

#### 1. Introduction

Artificial intelligence (AI) has recently emerged as a game-changing technology that has the potential to significantly impact many fields. AI is a field of computer science that aspires to program computers to do tasks that normally require human intelligence. AI's core goal is to create machines with human-level intelligence (Wang et al., 2023). Language comprehension, pattern identification, problem solving, and deductive reasoning are all necessary for this task. Machine learning, natural language processing (NLP), robotics, and many more technologies are all part of the wide subject of AI. The widespread implementation of technologies that use AI has had a tremendous impact on the areas of industry, healthcare, education, the military, cyber security, and even defense (Firat, 2023).

The use of NLP algorithms to simulate human speech in conversational bots like ChatGPT (Generative Pre-trained Transformer) is a notable use of artificial intelligence in various fields. ChatGPT is considered as a language-based artificial intelligence system that belongs to a wider group of transformers. As stated by (Kooli, 2023), deep neural networks such as transformers, were developed specifically to process and create sequences of data, such as text. OpenAI's ChatGPT, which was opened to the public in November 2022, has quickly gained popularity, with one million people signing up in just five days. Facebook took 300 days, Twitter 720 days, and Instagram 75 days to reach this number (Biswas, 2023). Moreover, Iskender (2023) emphasizes that ChatGPT is a chatbot that can create responses that resemble those of actual people to a variety of inquiries and prompts since it was trained on a sizable dataset of text from the internet. Despite ChatGPT's novelty, understanding its users' sentiments is crucial for a number of reasons. To begin, ChatGPT adopters' opinions and sentiments can contribute to shaping the general public's image of new technology. The success or failure of ChatGPT may be determined in large part by the insights revealed by this data. Second, ChatGPT adopters may have discovered issues or problems with this new

technology; their feedback can aid in finding and fixing these problems before they become common. As a result, the potential of ChatGPT's success on the market might be improved by investigating the perceptions of the tool's users.

Researchers collected data from Twitter, which enables users to read and post short messages (called "tweets"), in order to get insight into the sentiments of ChatGPT users. The ability of the Twitter platform to present daily trending topics on a national and global scale and to allow a large number of users to voice their thoughts and opinions on these subjects at the same time makes it more timely and popular than other social media platforms (Taecharungroj, 2023). Twitter was chosen for this research because of its widespread use and open data policy. Twitter is used for sentiment analysis research because, unlike other platforms like Facebook and Instagram, it makes its data and tweets public. Twitter is ranked number 4 on Alexa's list of the most frequented websites throughout the globe (Munggaran et al., 2023). Twitter's public nature encourages more conversation and the development of a larger user base that can quickly respond to and discuss breaking news and popular issues.

#### 2. Objectives

Despite the advancements in sentiment analysis and AI language models, several gaps remain in the literature concerning sentiment analysis in the Philippine setting. The purpose of this research is to delve into and dissect the feelings and thoughts of ChatGPT users in the Philippines. Because of its cultural richness, multilingual environment, and particular emotional manifestations, the Philippines provides a fascinating and novel context for the study of sentiment analysis. The widespread use of AI technologies in the Philippines makes it all the more important to investigate the emotions conveyed by users in conversations with Generative AI language models like ChatGPT.

#### 3. Research Questions

- a. What are the predominant sentiments expressed by ChatGPT users in the Philippines?
- b. What is the polarity of sentiments by ChatGPT users in the Philippines expressed on Twitter?

#### 4. Research Framework

Leveraging data obtained from Twitter tweets specifically mentioning ChatGPT, this study employs sentiment analysis techniques to discern the prevailing attitudes of users towards the language model.



*Figure 1*. Three main stages of the research framework

The research framework involves three main stages: data collection, data preprocessing, and sentiment analysis. First, Twitter's API is utilized to gather a substantial dataset of tweets containing references to ChatGPT. Next, the collected data undergoes preprocessing to eliminate noise, handle privacy concerns, and convert text into suitable formats for analysis. Finally, sentiment analysis is conducted using lexicon-based approaches to determine whether the sentiment expressed in tweets is positive, negative, or neutral.

#### 5. Literature Review

#### 5.1 Related Studies

Due to the novelty of ChatGPT, the researchers were unable to locate any large studies that dealt specifically with this topic. There is, however, a body of work devoted to the GPT family of text-generating AIs (including, but not limited to, GPT-2 and GPT-3). Furthermore, we found enough information on Twitter data mining to investigate users' sentiments. Therefore, we describe certain papers that are very relevant to our study below.

To study public opinion on hybrid work arrangements using tweets, Trivedi et al. (2022) developed a Robustly Optimized BERT Pre-training Approach (RoBERTa). According to the RoBERTa, the consensus opinion of users is favorable toward the hybrid work approach. Another article by Alhijawi & Awajan (2022) used Twitter temporal data mining to forecast the popularity of a film. In order to foretell user pleasure and the popularity of movies among users, the authors suggested a rating prediction model and a temporal product popularity model. Similarly, Sanjaya et al. (2022) presented a sentiment analysis using Twitter data to discover perceptions of electronic wallets usage. The results of the sentiment analysis reveal that Twitter users in Indonesia are less likely to have a positive opinion of electronic wallets. Moreover, Haque et al. (2022) performed sentiment analysis to qualitatively analyze selected set of tweets. Results showed that the vast majority of early adopters are enthusiastic about the potential benefits of this technology in areas like software development disruption, entertainment, and creative expression. And only a limited percentage of users worried about things like the possible abuse of Chat-GPT.

#### 5.2 Generative AI models

In recent years, generative AI models have received a great deal of attention and interest from the general public due to their capacity to generate material that is very similar to that generated by humans. These models are capable of responding to a wide variety of inputs, from pictures and text to audio and video (Dwivedi et al., 2023) and even human voice. The two most well-known GPT-based AI products published by OpenAI in 2022 are ChatGPT (OpenAI, 2022b) and DALL-E (OpenAI, 2022a). Some other Generative AI models may also make user images called Magic Avatars (Pavlik, 2023); for example, Stable Diffusion from Stability.ai and Lensa. The Language Model for Dialogue Applications (LaMDA) is the foundation of Google's latest Generative AI system, Bard (Pichai, 2023) which was launched recently.

#### 5.3 Sentiment Analysis

Since the development of NLP, texts have been seen as data sources, and studies have started to look at the subjective meanings of phrases in the context of data extraction. Information on how individuals feel about certain events, circumstances, goods, or even political beliefs may now be gleaned from their online discussions of these topics, thanks to techniques known as "opinion mining" and "sentiment analysis." Many new academic subfields have emerged in response to the proliferation of social networking sites, all with the goal of mining useful data from these platforms (Wankhade et al., 2022). Several practical contexts call for in-depth exploration using sentiment analysis. For example, doing a product analysis to learn what features or characteristics a client's value.

Subhashini et al.'s (2021) article offers the findings of a thorough evaluation of current research on opinion mining. It also discusses representing knowledge in views, classifying them, and extracting text characteristics from opinions that include noise or ambiguity. A method for adaptable aspect-based lexicons for sentiment categorization is proposed by

Mowlaei et al. (2020). The authors presented a statistical technique and a genetic algorithmbased strategy for building two dynamic lexicons to help in the categorization of attitudes based on their characteristics. More accurate scoring of ideas dependent on context is made possible by an automatically updated dynamic lexicon (Kumar & Uma, 2021).

To the best of the researchers' knowledge, the study presented is the first to perform a comprehensive examination of the Twitter sentiments on ChatGPT in the Philippines. This research contributes to the existing body of knowledge by giving a glimpse of Twitter sentiments to this newest innovation.

#### 6. Research Methodology

In this investigation, sentiment analysis method was utilized to understand perception of ChatGPT users in the Philippines. The users' perception toward AI chatbots like ChatGPT can be used as a measure of sentiment because of the widespread impact of its outputs. In the dictionary-based sentiment analysis, each word is given a value that is derived from its definition in the dictionary, beyond its positive or negative connotations. The sum of these values identifies the emotional value of the sentence. Natural Language Toolkit (NLTK) was used to perform sentiment analysis of the obtained tweets in this study. NLTK represents a Python library. It offers a solid foundation for constructing Python applications and organizing data. The toolkit is also crucial in transforming textual data into a certain format, from which sentiments may be retrieved (Tunca et al., 2023). The primary purpose of the NLTK is to analyze human language data for use in natural language processing. Additionally, NLTK also offers a variety of pre-processing functions to support all NLP techniques, such as part-of-speech tagging, tokenizing, lemmatizing, stemming, parsing, and performing sentiment analysis for selected datasets.

#### 7.1 Data Collection

The data comprising the tweets used in this study were obtained from Kaggle, a free and online community platform which allows users to find and publish data sets. The data is formatted as CSV (Comma Separated Value) which is a flat text file about tweets. The Twitter dataset consists of 478,379 tweets from public Twitter accounts related to the search term "ChatGPT". Twitter's Stream API was used to accumulate records in the dataset from January 4th to March 29th, 2023. In addition, only general tweets in the English language were processed. Other types of tweets such as mentions, replies retweets, and retweets with comments were discarded.

#### 7.2 Data Preprocessing

When expressing their thoughts and feelings about an issue, individuals often fail to employ correct proper language structure. Instead, the study of these structures becomes more challenging by the use of slang, misspelling, alternative emoticons, and abbreviations, and sometimes puns. As the quality of the output is dependent on the quality of the input, the complexity of the textual data has a major impact on the efficiency with which sentiments are analyzed (Elbagir & Yang, 2019). Therefore, a number of preprocessing procedures have been executed to remove unnecessary data from the collected tweets in order to make the data more suitable for mining and feature extraction, consequently increasing the reliability and precision of the results. Preprocessing the tweets also removed retweets and duplicates, resulting in a dataset containing 507 tweets. To preprocess these data, NLTK was utilized. To begin, a regular expression (Regex) was executed in Python to identify and remove URLs ("http://url"), retweets (RT), user mentions (@), and undesired punctuation from tweets. Hashtags (#) are included as part of the tweet, but the "#" sign has been deleted since they frequently clarify the subject of the tweet and provide important information relevant to the

topic of the tweet. After a dataset has undergone the preprocessing steps, it is ready to be classified based on sentiment using the VADER algorithm. For the purpose of determining the frequency of words and, consequently, their visualization in Word Cloud, a number of NLTK functions are employed as additional preprocessing steps.

#### 7.3 Sentiment Classification

In general, sentiment analysis may be used to investigate the polarity of a given context, whether that context be positive, negative, or a neutral sentiment. In this phase, Valence Aware Dictionary for Sentiment Reasoning (VADER) was employed. VADER is both a language for expressing emotions on social media and a rule-based tool for analyzing user sentiment (Tunca et al., 2023). VADER analyzer was used to examine both the polarity of sentiments and intensity in each tweet.

#### 8. Results

This section provides the results of a Twitter sentiment analysis conducted using VADER and NLTK.

#### 8.1 Descriptive Analysis

During the data collection period from January 4 to March 29, 2023, there were primarily 478,379 tweets associated with the ChatGPT topic, which provides ample opportunities to observe trends.



Figure 2. Number of Tweets per Day

A large number of tweets were observed between November 30, 2022 to April 8, 2023 during which ChatGPT reaches 100 million monthly users at a faster rate than Instagram or TikTok (Chow & Perrigo, 2023), and receives an estimated 1 billion visitors, an average of over thirty-five million visitors per day (Brandl & Ellis, 2023).



*Figure 3.* Number of Tweets per Week

Figure 3 reveals that the week beginning March 12 had a high number of tweets, most likely because ChatGPT Plus was released as a subscription service, providing faster responses and priority access to new features than the free version of ChatGPT (Deutscher, 2023), and possibly because ChatGPT experienced a data breach when a bug exposed users' personal information (Wagenseil, 2023).


Figure 4. Top 10 Most Unigram Frequent Words

As shown in Figure 4, numerous words are associated with artificial intelligence, such as "chatgpt", "ai", "gpt", "just", "use", "new", "like", "asked", "openai", and "industry". In the same way, Figure 5 depicts the ten prevalent bigrams and common words, including (artificial, intelligence), (powerful, dominating), (dominating, industry), (dominating, crypto), (crypto, industry), (large, multimodal) and so on. For instance, the terms "ai" or "artificial intelligence", and "industry" can be seen in both unigram and bigram graphs, indicating that people believe artificial intelligence will be the dominant industry of the near future.





It is worth noting that the terms 'chat', 'gpt', 'chatgpt' were excluded from these n-grams to get better results focusing on other aspects related to ChatGPT.



Figure 6. Word Cloud of Tweets

Figure 6 shows a visual representation of the most frequently occurring words in the processed tweets from the dataset. This Word Cloud is a useful tool for evaluating widespread sentiment since it illustrates at a glance the most discussed topics related to ChatGPT on Twitter. Words like "chatgpt," "ai," "use," "new," and "will" stand out because they are frequently used in discussions about ChatGPT.



Figure 7. Distributions of sentimental polarities

The lexicon-based VADER technique was used to extract features by loading the appropriate NLTK library. The module's applied techniques yielded three sentiment scores (positive, negative, and neutral) as the module's output scores. These scores are measure whether each tweet is a positive tweet, a negative tweet, or a neutral tweet with a value ranging from [-1 to +1]. Figure 7 presents the tweets' sentiments with corresponding values. The values of [-1, 0, 1] were set to refer to positive, neutral, and negative sentiments.

	Table 1.	Count of	Tweets	Accordina	to	Polaritv
--	----------	----------	--------	-----------	----	----------

	Count	Percentage
Positive Tweets	263	51.87%
Neutral Tweets	160	31.56%
Negative Tweets	84	16.57%
Total	507	100.00%

Table 1 depicts that the majority of tweets contained positive and neutral sentiments about ChatGPT. Remarkably, (263) 51.87% of the collected tweets contained positive sentiments, (160) 31.56% contained neutral opinions, and (84) 16.57% expressed negative sentiments. In comparison to other classifications, the number of positive tweets was the highest, indicating that the majority of users expressed a positive attitude toward ChatGPT.



Figure 8. Sentiment Analysis of Tweets

Figure 8 displays number of tweets by sentiment of each polarity in a bar chart. Positive tweets constitute the majority of the dataset, accounting for 263 of all the analyzed tweets in the Philippines. This suggests that a significant portion of the Twitter users' sentiments were positive during the time the data was collected. The high percentage of positive tweets could indicate that people were expressing joy, satisfaction, excitement, or contentment in their experience with ChatGPT is probably because users perceive it as a viable means of enhancing commercial operations (Taecharungroj, 2023). The capability of ChatGPT is exerting a significant influence on different technology industries. For example, the platform gives the opportunity for students to pose inquiries and address any uncertainties, augmenting

the whole educational encounter. Despite the inherent restrictions, ChatGPT is anticipated to have practical applications in real-world scenarios, and businesses have a strong inclination to utilize it for the sake of generating profits (Haleem et al., 2022).

Neutral tweets account for 160 tweets or 31.56% of the tweets in the Philippines. These tweets represent a balance between positive and negative sentiments, as they do not strongly lean towards any particular emotion. Neutral tweets may include informational tweets, facts, or updates on ChatGPT without expressing a specific opinion or emotion. This significant proportion of neutral tweets could indicate that many users were simply sharing information or updates about ChatGPT rather than explicitly expressing their emotions.

On the other hand, there are only 84 negative tweets which make up about 16.57% of the tweets in the Philippines. While this percentage is lower than positive tweets, it still represents a considerable proportion of the analyzed tweets expressing negative emotions, such as anger, sadness, disappointment, or frustration about ChatGPT is likely due to the potential of this technology to generate and disseminate deceptive news and information that appears credible has caused significant concern among many individuals (Taecharungroj, 2023).

In order to mitigate negative sentiments expressed towards ChatGPT, developers can employ several strategies, like educating users about the capabilities and limitations of the system, which can help manage their expectations and reduce potential frustrations. Granting users, a sense of control over the system, for instance, by allowing them to customize its behavior within certain bounds, can enhance user satisfaction and mitigate negative sentiments. Lastly, developers should promptly address specific concerns raised by users, demonstrating responsiveness and a commitment to resolving issues.

Overall, the sentiment analysis provides valuable insights into the prevailing emotions among Twitter users in the Philippines during the data collection period. The higher percentage of positive tweets might suggest a generally positive atmosphere on ChatGPT, while the presence of negative and neutral tweets also reflects the diversity of opinions and discussions happening on Twitter.

#### 9. Conclusion

The findings from the analysis of the Twitter dataset related to ChatGPT provide valuable insights into the sentiments and discussions surrounding the AI-powered language model. The dataset was obtained during the time which ChatGPT witnessed a significant surge in popularity, reaching 100 million monthly users at a faster rate than Instagram or TikTok. During the week beginning March 12, there was a notable increase in the number of tweets, possibly driven by the release of ChatGPT Plus, a subscription service offering faster responses and priority access to new features. In addition, concerns about privacy and security may have contributed to the spike in tweets after a data breach in which user information was exposed due to a security breach.

The examination of n-grams and frequently-used terms found a strong connection with AI, showing the importance of this issue in ChatGPT-related talks. It also reflects a view that AI will play a preponderant role in defining future sectors and a curiosity with AI technology and its possible uses. However, the VADER lexicon-based system for analyzing sentiment found that most tweets were either neutral or favorable towards ChatGPT. Around 51.87% of tweets were positive, indicating that the platform garnered a favorable response from users. Neutral tweets constituted 31.56% of the dataset, reflecting informative and objective discussions without strong emotional sentiments. Negative tweets represented 16.57% of the tweets, expressing concerns, criticisms, or frustrations about certain aspects of ChatGPT.

It appears that users are generally pleased with ChatGPT, as seen by the positive attitude of the majority of tweets. However, the existence of critical tweets indicates that the platform has room for development and may face issues that need to be addressed. Limitations of the approach and context-specific elements that may impact the sentiments expressed on social media platforms should always be taken into account when doing any sort of sentiment analysis. ChatGPT's developers and stakeholders may utilize the insights collected from this study to better understand user impressions and input, which might lead to future enhancements and improvements.

As a result of its impressive performance and widespread popularity, ChatGPT may now be considered a frontrunner among AI-powered language models. Its promise as a gamechanging technology is evidenced by both its subscription model and its quick user growth. If ChatGPT is to sustain its current success in the rapidly changing AI world, its developers and stakeholders will need to stay cautious in responding to user concerns and protecting user data. In addition, a proactive strategy to monitoring user sentiment is essential as the platform continues to explode in popularity. This necessitates persistent monitoring of social media for conversations, user input, and emerging issues. Engaging with users actively through official channels, such specialized support teams or feedback mechanisms, can aid in addressing issues promptly. Further research and development are anticipated for ChatGPT and natural language processing in order to augment their capabilities and address existing constraints. Some of the areas that can be focused on for further development are enhancing comprehension of context, refining training methodologies, mitigating biases, advancing dialogue systems, accommodating many languages, and integrating with other artificial intelligence technology.

#### References

- Brandl, R., & Ellis, C. (2023, July 19). CHATGPT statistics and user numbers 2023 OpenAl Chatbot. Tooltester. https://www.tooltester.com/en/blog/chatgpt-statistics/
- Biswas, S. (2023, February). Role of ChatGPT in education. https://ssrn.com/abstract=4369981
- Choudhury, A., & Shamszare, H. (2023). Investigating the Impact of User Trust on the Adoption and Use of ChatGPT: Survey Analysis. *Journal of Medical Internet Research*, 25, e47184.
- Chow, A. R., & Perrigo, B. (2023, February 17). The Ai Arms Race is on. start worrying. Time. https://time.com/6255952/ai-impact-chatgpt-microsoft-google/
- Davis, F.D. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Q. 1989, 13, 319.
- Deutscher, M. (2023, April 26). OpenAI adds new data controls to CHATGPT and previews business version. SiliconANGLE. https://siliconangle.com/2023/04/25/openai-adds-new-data-controls-chatgpt-previews-business-version/
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023).
  "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International *Journal of Information Management*, 71, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642
- Elbagir, S., & Yang, J. (2020). Sentiment analysis on Twitter with Python's Natural Language Toolkit and vader sentiment analyzer. IAENG Transactions on Engineering Sciences. https://doi.org/10.1142/9789811215094 0005
- Firat, M. (2023). What CHATGPT means for universities: Perceptions of scholars and students. Journal of Applied Learning and Teaching, 6(1). https://doi.org/10.37074/jalt.2023.6.1.22
- Haleem, A., Javaid, M., & Singh, R. P. (2022, October). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, 2(4), 100089. https://doi.org/10.1016/j.tbench.2023.100089

Haque, M. U., Dharmadasa, I., Sworna, Z. T., Rajapakse, R. N., & Ahmad, H. (2022). " I think this is the most disruptive technology": Exploring Sentiments of ChatGPT Early Adopters using Twitter Data. arXiv preprint arXiv:2212.05856.

Iskender, A. (2023). Holy or unholy? interview with open Al's chatgpt. *European Journal of Tourism Research*, 34, 3414. https://doi.org/10.54055/ejtr.v34i.3169

Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. Sustainability, 15(7), 5614. https://doi.org/10.3390/su15075614

Mowlaei, M.E., Abadeh, M.S., & Keshavarz, H. (2020) Aspect-based sentiment analysis using adaptive aspect-based lexicons. Expert Syst Appl 148:113234

OpenAI. (2022a, November 2). DALL·E API now available in public beta. https://openai.com/blog/dalleapi-now-available-in-public-beta

Pichai, S. (2023, February 6). An important next step on our AI journey. Google. https://blog.google/technology/ai/bard-google-ai-search-updates/

Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., ... & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, 6(1).

Sanjaya, A. D., Pudjiantoro, T. H., Ningsih, A. K., & Renaldi, F. (2022). Sentiment Analysis Of E-Wallets on Twitter social media With Naïve Bayes and Lexicon-Based Method

Subhashini, L., Li, Y., Zhang, J., Atukorale, A.S., & Wu, Y. (2021) Mining and classifying customer reviews: a survey. Artif Intell Rev 54:6343–6389

Taecharungroj, V. (2023). "What can chatgpt do?" analyzing early reactions to the innovative Al chatbot on Twitter. *Big Data and Cognitive Computing*, 7(1), 35. https://doi.org/10.3390/bdcc7010035

Wagenseil, P. (2023, May 4). From the headlines: Chatgpt and other AI text-generating risks. SC Media. https://www.scmagazine.com/resource/from-the-headlines-chatgpt-and-other-ai-text-generating-risks

Wang, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023).
 Exploring the Potential Impact of Artificial Intelligence (AI) on International Students in Higher Education: Generative AI, Chatbots, Analytics, and International Student Success.
 https://doi.org/10.20944/preprints202305.0808.v1global research collaboration. *Research and Practice in Technology-Enhanced Learning, 1*(1), 3-29.

Wankhade, M., Rao, A. C., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. Artificial Intelligence Review, 55(7), 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

## Implementing an Inquiry Learning with Mobile-supported Mathematical Board Game to Promote Primary Students' Attitude toward Mathematics Learning

Sakda CHALEEPLIAM<sup>a</sup>, Sasivimol PREMTHAISONG<sup>a</sup> & Pawat CHAIPIDECH<sup>b\*</sup> <sup>a</sup>Khon Kaen University Demonstration School International Division, Khon Kaen University,

Thailand <sup>b</sup>Faculty of Education, Khon Kaen University, Thailand \*Pawach@kku.ac.th

**Abstract:** One of the teaching approaches most used by educators is the inquiry-based learning, and with the development of technology, students have access to more games. Therefore, many researchers have integrated investigative learning methods with mobile-supported board games. The purpose of this research is to investigate students' conceptual understanding and their attitude toward mathematics learning using inquiry-based learning methods and mobile-supported board games. In this study, the participants were primary school students in the northeastern region of Thailand. The finding demonstrated that students have better conceptual understanding after using an inquiry learning with mobile-supported mathematical board game. Moreover, students had a positive attitude toward mathematics learning. Specifically, it allowed students to gain confidence and motivated them to study mathematics. Consequently, this study revealed that an inquiry-based learning combined with mobile-supported board game has the potential to foster positive attitudes towards mathematics learning with technology and may develop into other content in future primary mathematics learning.

Keywords: Board game, Inquiry-based Learning, Mathematics, Attitude

#### 1. Introduction

Digital technology is now widely embraced in education, especially in subjects such as mathematics, aiming to boost students' understanding and enthusiasm. Educators are increasingly turning to digital technologies to create engaging learning environments that blend entertainment with education. For example, researchers have designed digital board games to enhance students' learning (Premthaisong & Srisawasdi, 2020; Nukprach, Chaipidech & Srisawasdi, 2023), and Chang and Panjaburee's research in 2022 demonstrated that Virtual Reality technology positively impacts learners' achievements, problem-solving abilities, and engagement. In the area of mathematics in primary school, the difficulty of effectively transferring new knowledge due to students accepted unfavorable views towards the subject is significant (Ardelenau, 2019). The study shown the potential for improved learning outcomes and improved attitudes when modern digital technology-infused learning environments are used to encourage investigative problem-solving (Ardelenau, 2019). One of the approaches that many researchers adapt for teaching in the classroom, is inquiry-based learning. For example, the complexity of classifying quadrilaterals, as demonstrated by Japanese sixth graders' tendency to perceive rhombuses differently from parallelograms. encourage educators to employ a variety of teaching approaches, such as inquiry-based learning (Okazaki & Fujita, 2007). Moreover, an inquiry-based learning fosters problemsolving skills and cognitive engagement by encouraging systematic knowledge exploration (Kohen, Schwartz-Aviaia, & Peleg, 2023). This strategy holds promise for facilitating students' solution discovery and fostering positive attitudes while confronting difficult mathematical problems, thereby contributing to a deeper appreciation for mathematics.

From the benefit of digital technology and the inquiry-based learning approach, many instructors utilize inquiry-based learning integrated with digital technology, such as Srisawasdi and Punjaburee (2019), which used a game-transformed inquiry-based learning approach to teach high school students about chemistry. In the area of mathematics, as reported by Lin and Cheng (2022), the use of digital technology and the inquiry-based learning approach in mathematics education has been found to have a positive impact on students' motivation and accomplishment of goals. Consequently, this study incorporates mobile-supported board game into an inquiry-based learning can be considered a pedagogical approach for technology-enhanced mathematics education, to promote students' understanding of concepts and positive attitudes toward mathematics learning.

#### 2. Literature Review

#### 2.1 Mobile-supported board game

Due to the benefits of game-based learning and technology, many educational researchers are interested in how to use smart phones to support students' learning. Numerous research studies have demonstrated that the integration of technology-supported game-based learning has the potential to enhance student motivation and foster improved conceptual comprehension across various academic disciplines, including mathematics and science. For example, Chao et al. (2018) posited that mobile-supported game-based learning in mathematics, can enhanced students learning. In addition, according to Lin and Cheng (2022), the utilization of technology-enhanced board games has been identified as a valuable approach for enhancing students' motivation and improving their academic performance in the context of primary mathematics. Moreover, according to the findings of Yeo and Campbell (2022), most a primary mathematics teacher in the USA agreed that the utilization of games that were supported by technology are an effective tool for teaching mathematics.

#### 2.2 Inquiry-based Learning in Mathematics

Inquiry-based learning (IBL) is a learning approach that emphasizes developing problemsolving abilities, by training learners to investigate knowledge by questioning students to use their thinking process and finding logical conclusions, or to self-correct solutions. Many researchers employed Inquiry-based Learning to integrate with various technologies. According to Kohen et al. (2022), the utilization of Inquiry-based learning combined with technology has been shown to be a successful approach in the instruction of mathematics as well as this finding, it corresponds to Pedersen and Haavold (2023), who showed that IBL has potential in fostering a positive attitude and attraction towards the mathematics learning for students in the age-range 11-16. In addition, Radmehr et al. (2023) presented the potential of inquiry-based learning as a more effective pedagogical approach to instructing undergraduate students. It was found to enhance students' conceptual understanding, particularly in complicated mathematical topics. Furthermore, the implementation of Inquiry-Based Learning in Mathematics classes promoted a positive learning environment and attitudes towards mathematics (Aldridge & Robinson, 2022).

#### 3. Research Methodology

#### 3.1 Research Design

The research design was one group pre-test and post-test design. The participants were selected by using a purposive sample, in this case fifth-grade students. The data collection instruments consist of an attitude questionnaire and a quadrilateral conceptual test.

#### 3.2 Participants

The study involved a sample of 25 fifth-grade participants from a public school located in the Northeastern region of Thailand. The chosen school was a technologically advanced institution with access to the Internet. In addition, prior knowledge about quadrilaterals were needed. That was why they were chosen to be the in this research.

#### 3.3 Research Instruments

In this study, there are two research instruments. The first instrument is a conceptual test involving the conception of the properties of quadrilaterals. The test was adapted from the study of Okazaki and Fujita (2007) about the properties of quadrilaterals. There are five questions in the test. Each question consists of six to eight statements. The participants had to read and fill in the correct symbol if they think that statement is correct, the cross symbol if they think it is incorrect and the question mark symbol if they are not sure. In the first question, there are eight quadrilateral figures, the students are asked, which are parallelograms. Second question, there are six quadrilateral figures, the students are asked which are rectangles. The third question, there are six quadrilateral figures, the students are asked which are squares. The fourth question, there are six statements about properties, the students are asked about the properties of parallelograms. The fifth question, there are five statements, regarding the properties of rectangles, and in the last question, there are five statements, regarding the properties of squares. The second instrument is an attitude questionnaire regarding the use of technology to learn mathematics. A 5-point Likert-scale survey is being used to examine student attitudes. It consists of 20 questions covering four each on mathematics confidence (MC), four each on technological confidence (TC), four each on attitude towards using technology to learn mathematics (MT), four each on affective engagement (AE), and four each on behavioral engagement (BE). The following subscale reliability levels are indicated by Cronbach's alpha values: MC, 0.87; MT, 0.89; TC, 0.79; BE, 0.72; and AE, 0.65. Pierce et al.'s (2007) guestionnaire was used to create this.

#### 3.4 Data Collection and Analysis

The data collection of this research is shown in the figure 1. Before the approach, the students were given 20 minutes to do the pre-test of conceptual understanding and attitude questionnaire. After completing the instruments, they were expected to interact with a mobile-supported board game with inquiry-based learning, which consists of three phases: First phase, Pre gaming phase begins with open-ended questions and provides basic information about quadrilateral that students can use in the next step. Second, Gaming phase is a learning process through a quadrilateral board game with mobile phone to find answers to the first open-ended questions. Third phase, Post gaming phase is a summary of knowledge gained in answering those open-ended questions together. After finishing the learning process, students were administered the same questionnaires again for 20 minutes to determine their conceptual understanding and attitude toward mathematics learning with technology.



Figure 1. An illustration of the experimental procedure of this study.

The data collection process took one week (5 periods, 250 minutes) of which the first two periods consisted of a pre gaming phase in which open-ended questions and background theory were provided.

#### 3.5 Learning Material

The game carried on round by round. Firstly, each student began by picking up a question card. Then, they answer the questions on that card and scan a QR code to check the answers. If they got the answer right, they would get one coin. The coins could be collected for winning the game in case nobody won in the game. After that, each student could decide to follow one of the three directions. To clarify, the first direction is to roll the dice to race the horse. The students had to press the button of the mystery box, which contained five different-colored dices. If the student got the dice in any color, they could move the horse chess which shared the same color forward on the betel table. Also, if the chosen dice showed any number, they could move their dice forward as the same step on the betel table. During the game, the student can choose to either trap or reward their friends by putting +1 or -1 on their track. +1 refers to a step forward, whereas -1 refers to a step backward. Additionally, while playing, the student must bet which horse will win as well. Following the figure 2, With the rules of this game, students will encounter questions for themselves as well as assess the responses of other students regarding the properties of the quadrilateral.



Figure 2. An illustration of gaming steps with the quadrilateral board game.

#### 4. Result and Discussion

#### 4.1 Conceptual understanding of quadrilateral

This study using parametric statistics analysis of paired t-test, a comparison of pre-test and post-test conceptual understanding scores revealed that the post-test score of students' conceptual understanding was significantly higher than the pre-test score (t=-5.580, p<.05). The statistics the paired t-test are displayed in Table 1.

Table 1. Paired T-Test Result for Conceptual Understanding on The Properties of Quadrilateral.

Conceptual Understanding test	Ν	Mean	S.D.	t
Pre-Test	24	18.17	5.435	-5.580*
Post-Test	24	24.58	6.021	
*n<0.05				

\*p<0.05

As shown in the table above, the results indicate that students' post-test scores are significantly higher than their pre-test scores after using the board game with mobile integrated game-transformed inquiry-based learning approach. Based on the result in this study, students

could improve their conceptual understanding of the properties of quadrilaterals. Regarding the basis of prior research, Nukprach et al. (2023) found that using eye tracking of students when they learnt through the technology-supported board game could promote students' cognitive exertion which increases while playing board games. Hence, integrating board game with technology have been found to effectively aid in the comprehension of mathematical concepts on the properties of quadrilaterals.

#### 4.2 Students' Attitude toward Mathematics Learning with Technology

As shown in the table 2, we can see that it has a statistically significant effect on both mathematics confidence (MC) (F (1, 24) = 11.565; p < .05; partial  $\eta$ 2 = .335) and affective engagement (AE) (F (1, 24) = 7.271; p < .05; partial  $\eta$ 2 = .240).

	Atti	tude	_		
Subscale	Pre-Test	Post-Test	F	Sig	$\eta^2$
	Mean (SD)	Mean (SD)			-
MC	12.71(2.216)	14.75(2.739)	11.565	.002*	.335
TC	13.50(2.359)	14.54(2.395)	3.968	.058	.147
MT	14.62(3.943)	15.04(3.470)	0.204	.656	.009
AE	13.71(3.277)	15.46(2.859)	7.271	.013*	.240
BE	13.46(2.553)	14.67(2.632)	3.631	.069	.136

Table 2. The sti	udents' subscale mean	s of attitude by time	e and univariate MANOVA.
------------------	-----------------------	-----------------------	--------------------------

#### \*p<0.05

As shown in the table 2, it refers that a board game with mobile is the key to gain students' confidence in learning Mathematics (MC). In other words, mobile-supported board games promote the student's happiness (AE) when they can solve difficult math problems. The findings of this study align with the research conducted by Premthaisong and Srisawasdi (2020) and Lin and Cheng (2022), which explored the integration of technological devices with board games as a pedagogical tool for teaching primary school students is a great tool for teaching students in primary school. Similarly, Hwang and Chang (2011) observed that this instructional approach not only fostered students' interest and attitude towards learning, but also positively impacted their academic achievement. The utilization of mobile learning as a means of evaluating the enhancement of students' attitude and academic performance.

However, some students, who are normally familiar with using physical board games rather than the digital kinds, did not have a great performance and lack confidence in playing mobile-supported board games. Consequently, it resulted the significance value of technological confidence of no less than 0.05. Nevertheless, there are numerous other technologies that can be applied to board games to create mobile-supported board games that promote positive attitudes and academic achievement for students in primary school.

#### 5. Conclusion and Limitation

The purpose of this study is to investigate the conceptual understanding and the attitude toward mathematics learning with technology, of primary school students by using mobile-supported mathematics board game. Based on the findings of this study, we can confirm that the strategy is an effective tool for teaching difficult issues in mathematics learning. Moreover, they have a positive attitude of technology-enhanced mathematics learning with game-based inquiry learning. According to the findings of this study, mobile-supported board games have the potential to foster greater innovation in mathematics classrooms. Due to this study being

conducted in Thailand, it can apply in other developing countries. However, a limitation of this study is the small number of samples, which is a single sample, which makes it impossible to compare conceptual and attitude scores to traditional teaching techniques and gamified inquiry-based learning with mobile learning. Therefore, in future work, there should be comparison group to ensure the learning gains because of this approach.

#### Acknowledgements

This contribution was funded by Khon Kaen University, Thailand. The authors would like to thank the school's principal and all study participants including those who have contributed to this study's accomplishment.

#### References

- Aldridge, J., & Robinson, J. (2022). Environment–attitude relationships: girls in inquiry-based mathematics classrooms in the United Arab Emirates. *Learning Environments Research*, 619-640. https://doi.org/10.1007/s10984-022-09409-x
- Ardeleanu, R. (2019). Traditional and modern teaching methods in mathematics. *Journal of Innovation in Psychology, Education and Didactics*, 133-140.
- Chaipidech, P., & Srisawasdi, N. (2020). A proposal for personalized inquiry-based flipped learning with mobile technology. *Workshop Proceedings 26th International Conference on Computers in Education*. Philippines: *Asia-Pacific Society for Computers in Education*, 344-353.
- Chang, C., Panjaburee, P. & Chang, S. C. (2022). Effects of integrating maternity VR-based situated learning into professional training on students' learning performances. *Interactive Learning Environment*, 2-15.
- Chao, W., Yang, C., Hsien, S., & Chang, R. (2018). Using mobile apps to support effective game-based learning in the mathematics classroom. Journal Name, *8*(5), 354-357.
- Hwang, G. J., & Chang, H. F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education*, *56*(4), 1023-1031.
- Kohen, Z., Schwartz-Aviad, L., & Peleg, T. (2023). Who moved my triangle? Pre- and in-service teachers inquiring in a mathematics lab. *International Journal of Mathematical Education in Science and Technology*, 2-29.
- Lin, Y.-T., & Cheng, C.-T. (2022). Effects of technology-enhanced board game in primary mathematics education on students' learning performance. *Molecular Diversity Preservation International*, 3-12.
- Nukprach, N., Chaipidech, P., & Srisawasdi, N. (2023). Investigation of learning attraction towards digital board game by using eye tracking technology. *The 6th IEEE Eurasian Conference on Educational Innovation* (pp.171-174). Singapore: IEEE.
- Okazaki, M., & Fujita, T. (2007). Prototype phenomena and common cognitive paths in the understanding of the inclusion relations between quadrilaterals in Japan and Scotland. *Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education*,41-48.
- Premthaisong, S., & Srisawasdi, N. (2020). Supplementing elementary science learning with multiplayer digital board game: a pilot study. In H. J. So (Ed.), *The 28th International Conference on Computer in Education* (pp.199-207). Online Asia-Pacific Society for Computers in Education.
- Radmehr, F., Tohidinasab, K., & Tavakoli, M. (2023). Teaching and learning of vertex coloring using an inquiry-based approach. *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 33*(7),781-795
- Srisawasdi, N., & Panjaburee, P. (2018). Implementation of game-transformed inquiry-based learning chemistry to promote the understanding of and motivation to learn. *Journal of Science Education and Technology*, 152-164.
- Yeo, S., Rutherford, T., & Campbell, T. (2022). Understanding elementary mathematics teachers' intention to use a digital game through the technology acceptance model. *Education and Information Technologies*, 11515-11536

### Application of Blended Learning with PhET Simulation to Encourage Learning in Mathematics of Fractions

Atcharaporn ASSAWAPHUM<sup>a</sup>, Sasivimol PREMTHAISONG<sup>a</sup>, & Pawat CHAIPIDECH<sup>b\*</sup>

<sup>a</sup>Khon Kaen University Demonstration School International Division, Khon Kaen University, Thailand

<sup>b</sup>Faculty of Education, Khon Kaen University, Thailand \*pawach@kku.ac.th

**Abstract:** Technology in education is essential for both teaching and learning. The purpose of this study is to correct mathematic misconception about the addition and subtraction of fractions, as well as explore students' attitudes toward using technology to learn mathematics. In this study, blended learning activities was used with PhET simulation. Both hands-on and digital technology activity were implemented. The findings of this study revealed better comprehension in students who learned with a blended learning activity with PhET simulation. In addition, there was a general improvement in students' attitudes toward using technology for learning mathematics. The results implied that using blended learning activity with technologies in mathematics classrooms could correct leaners' misconception as well as fostering positive attitudes toward learning mathematics in primary level.

Keywords: Mobile learning, attitude, blended learning, misconception.

#### 1. Introduction

Technology can potentially enhance the teaching and learning of mathematics, leading to gains in higher-order thinking skills (Wenglinsky, 1998), as well as student achievements and self-efficacy (Mistretta, 2005). The National Council of Teachers of Mathematics (2000) explained, "Technology is essential in teaching and learning mathematics; it influences the mathematics that is taught and enhances student learning". Adewuyi (2001) noted that the style of teaching employed by teachers is a potent factor in motivating learners to learn; mathematics is a subject that is very easy to make difficult and very difficult to make easy. The perennial methods of teaching mathematics through listening, looking, and learning have failed. If anything, it has made students dislike mathematics (Akinsola, 2002). The current teaching and learning of mathematics include technology, such as using applications to calculate answers or playing math games through various websites and applications to review knowledge or import into lessons before learning begins or can be used in organizing learning activities that excite students about learning. Learning mathematics through simulations is another way for students to learn about technology. Simulations, such as PhET (https://phet.colorado.edu), provide a dynamic nature to multiple representations by making abstract concepts concrete, supporting an inquiry environment, and allowing for multiple trials and rapid feedback cycles, while also being engaging and enjoyable for students and teachers (Meadows, 2019). Simulation-based learning, it can be applied to many learning management methods; for example, Akinsola et al. (2007) used a simulation-games environment that could improve students' achievements and attitudes toward mathematics. The study found that the poor academic achievement of students in mathematics is partly due to the teaching method used and that stimulating teaching methods such as simulation games can sustain and motivate students' interest in learning mathematics. The results of this research suggest that emphasizing the use of simulation games to teach mathematics can improve mathematics teaching and learning in schools. Both the use of technology and simulation in managing mathematics learning result in learning achievement and knowledge. The students'

understanding is going to improve. If the advantages of using various technologies are applied in mathematics classes, there could be many benefits. Blended learning is another method of teaching and learning that use technology and face-to-face teaching together.

Blended learning combines both face-to-face instruction and technologically based online instruction. It combines traditional and advanced online learning methods, such as chatrooms, podcasts, and live online lectures (Barrett & Sharma, 2007). By using blended learning, learners take lectures in a face-to-face class and can interact with each other through online platforms. These online learning experiences may replace classroom instruction, but interaction with peers and instructors is also possible on these platforms (Pandya and Shroff, 2022). The use of blended learning in education can enhance student learning and make the whole learning experience more enjoyable. Moreover, Indrapangastuti et al. (2021) demonstrated that blended learning is more effective than conventional learning in improving students' achievement of mathematical concepts.

In mathematics learning, most students had issues for understanding mathematical concepts. For example, Retnawati et al. (2011) reported that up to 88.57% of students have trouble understanding the mathematical concepts that they are supposed to be learning. Moreover, Waskitoningtyas (2016) found that there were lots of reasons why students have difficulty learning mathematics, such as a lack of interest in the subject, the need to learn numerous formulas, boring teaching methods, and teachers not paying enough attention to students with low levels of understanding. One of the concepts that students have difficulties with is fractions. Alghazo and Alghazo (2017) found that If students have misconceptions about fractions now, they will continue to have misconceptions in the future. Thus, Aksoy and Yazlik (2017) mentioned that students should learn about fractions because fractions are related to many other math topics, such as percentages, decimals, probability, and algebra. Therefore, to correct misconceptions in mathematics related to addition and subtraction of fractions, it is necessary to employ teaching methods that engage students' interest in learning. This will enable students to perceive correct mathematics concepts better and allow them to share their thoughts, engage in discussions, and exchange opinions in the classroom. Hence, the researchers utilized blended teaching with simulation from PhET on fraction concepts in mathematics classrooms.

#### 2. Literature Review

#### 2.1 Technology-supported Learning in Mathematics

Over the past two decades, there has been an increase in the use of high-tech tools in classrooms. A new generation of information teaching techniques has emerged that combines digital media, human-computer interaction, and educational technology. For example, Verzosa et al. (2022) studied the development of an app and videos to support fractions. They found that videos immerse learners in a range of fraction representations that can extend their initial part-whole understanding of the fraction concepts. In addition, Hsu (2020) used Virtual Reality for mathematics classrooms. The results showed that VR has the effect of improving students' learning motivation and learning effectiveness in the digital teaching of mathematics. Furthermore, computer simulation is one of the technologies which could help students use multiple representations, support their efforts to create knowledge, concentrate on conceptual ideas. and offer quick feedback (D'Angelo et al., 2014). Simulation is helpful not only because it can provide believable numerical answers but also, perhaps more importantly because it requires a comprehensive comprehension of the problem to be solved (Reinhardt & Loftsgadeen, 1979). Problems are often unique or inaccessible to students. For such cases, students may be able to use interactive simulations as a supplement or even a substitute for an experiment. For example, the study of Srisawasdi et al. (2015) examined the effect of the teaching method of simulation-based inquiry with a dual-situated learning model (SimIn-DSLM) on students' conceptual understanding. This finding suggests that the SimIn-DSLM method could be used. To induce a mechanism of change within students' conceptual knowledge of sound wave phenomena, and the change of their conceptions could place them into the meaningful conceptual framework of basic scientific knowledge. Moreover, Arifin et al. (2022) studied the effectiveness of using PhET Simulation in primary school students' mathematical understanding of fractions. The results showed that using PhET Interactive Simulation significantly improved students' mathematical understanding of fractions. From the above, researchers focus on the technology used in simulation in mathematics teaching and learning. The use of simulations in teaching and learning can develop the students' competency.

#### 2.2 Blended Learning in Mathematics

Traditional classroom instruction and online learning have been modified and progressively replaced with "Blended Learning." The concept of blended learning, which combines multiple instructional models, has garnered considerable attention in recent years. The Blended Learning approach has been applied to mathematics education at the elementary school level. Research by Indrapangastuti et al. (2021) demonstrated the effectiveness of using blended learning to enhance students' mathematics concepts. The study's results suggested that the blended learning model is significantly more effective than the conventional learning model for enhancing students' mathematical concept achievement. These findings match Jemakmun (2022) research on blended learning to improve mathematics abilities. This study found that the blended learning model with formative and summative assessments improved students' mathematics learning outcomes, especially for those with high early math abilities. Blended learning methods can be combined with other teaching methods. Yennita and Zukmadini (2019) studied blended learning and Problem-based learning to improve critical thinking skills and student learning activities. This study concludes that applying problem-based learning (PBL) using the blended learning method can improve critical thinking skills and student learning activities.

Due to recent technological advancements, Marsh and Drexler (2001) and Willett (2002) claimed that blended learning represents all teaching models integrated with technology, such as e-mails, streaming media, and the Internet, and can be combined with traditional teaching methods. The blended learning concept has been defined and developed alongside internet technology. Therefore, researchers attempted to use blended learning, combining teaching technologies with face-to-face (F2F) instruction to assist students in archiving their learning objectives and practicing both learning and working (Driscoll, 2010).

#### 2.3 Misconceptions In Mathematics

Mathematics is one of the subjects taught at all educational levels. Students were typically introduced to mathematics at a young age. This is because mathematics has indirect effects on student activities. (Trivena et al., 2017) However, many students keep considering mathematics as a difficult and unattractive subject. If students believe that mathematics is relaxing, enjoyable, significant, and essential for their academic success and future, this will have a positive effect on their motivation to learn. Misconceptions and errors are possible. Other factors may include inexperience, difficulties in comprehending or interpreting inquiries, and a requirement for numerical literacy. A misconception, on the other hand, is the result of a misunderstanding or incorrect application of a rule or mathematical concept (Spooner, 2002). Errors are one method for measuring students' comprehension of concepts, problems, and procedures.

Fractions are one of the concepts with which students need assistance. In order to comprehend the concept of fractions, students must have an established understanding of fractional properties. Trivena et al. (2017) conducted a study on misconceptions about addition and subtraction of fractions among primary school students to determine the level of fraction addition and subtraction mastery among primary school students. Students' understanding of addition and subtraction was dominated by misconceptions, indicating a limited comprehension of the concepts, according to the findings. This is consistent with the research conducted by Pulungan and Suhendra (2019), who investigated misconceptions and errors involving fractions; based on the study's findings, it was discovered that some elementary

school pupils continue to have misconceptions regarding the concept of fractions. Various categories of errors are caused by these misconceptions when solving fraction problems. As a result of the student's misunderstanding of the actual concept of fractions, these misconceptions appear. The study emphasizes the importance of addressing misconceptions when teaching fractions to enhance students' understanding.

## 3. An Example of Teaching Mathematics with Blended learning Integrated Simulation on Fractions

In this section, the researchers intend to show the application of technologies in the mathematics classroom in an effort to enhance students' conceptual comprehension and foster positive attitudes towards utilizing technology for the purpose of learning mathematics. The researchers employed the Problem-based learning (PBL) approach as a pedagogical tool for delivering instruction in the present study. Furthermore, the present study also employed a blended learning approach that incorporated both hands-on learning materials and computer simulations. This pedagogical strategy aimed to enhance students' comprehension of the subject matter by facilitating their visualization of the concepts covered during class sessions. In particular, learners consistently regarded this instructional method as an essential part in their acquisition of fractional concepts. (See Table 1)

Components	Learning materials	Description of learning process	Examples of learning activity
1.Identify the problem	Video	Teachers present problems to students using familiar everyday situations.	CLEANING DU
2.Brainstorm possible solutions	Hands-on	Students used real materials and predicted answers through learning materials.	
3.Research the problem	PhET Simulation	The students search the problem to learn more about it using PhET.	
4.Develop a solution	PhET Simulation	The students doing another problem using PhET.	

Table 1. An Example of Teaching Mathematics with Blended Learning Integrate Simulation on Fraction

Components	Learning materials	Description of learning process	Examples of learning activity
5.Present the solution	Nearpod	The students present their solution to the class	
6.Reflect on the learning	Nearpod	The students apply their knowledge and reflect on the learning process	

#### 4. Methodology

#### 4.1 Participants

The participants in this study consisted of 25 primary students who are studying in the tenth grade, ranging in age from 9 to 10 years old, from a local public school located in the Northeastern part of Thailand. A majority of the students in the class, specifically 64%, demonstrated misconceptions regarding the concepts of addition and subtraction when applied to fractions. The students had already completed a unit on fractional concepts and were afterwards instructed on the operations of addition and subtraction using fractions. Furthermore, the participants possessed the capacity and had experience in utilizing technology for educational purposes on a consistent basis within the classroom setting prior to their involvement in this research.

#### 4.2 Research Instruments

In this study, there are two research instruments as measuring tools of conceptual understanding and attitude toward using technology to learn mathematics. The first tool is a conceptual test involving the conception of addition and subtraction fractions. The test is designed with open-ended questions to measure the ability to understand the topic. It consists of 16 test items converting the concept of addition and subtraction fractions. It has four concepts. There is addition with the same denominator (4 items), addition with a different denominator (4 items), subtraction with the same denominator (4 items), and subtraction with a different denominator (4 items). The second tool is an attitude to learning mathematics with a technology questionnaire. It is a 5-point Likert-scale questionnaire to investigate the student attitude. It consists of 20 items covering mathematics confidence (MC) (4 items), confidence with technology (TC) (4 items), attitude towards the use of technology for learning mathematics (MT) (4 items), affective engagement (AE) (4 items) and behavioral engagement (BE) (4 items). Cronbach's alpha values indicate the following subscale reliability for MC, MT, TC, BE, and AE are 0.87, 0.89, 0.79, 0.72, and 0.65, respectively. The questionnaire was obtained from Pierce et al. (2007).

#### 4.3 Data Collection and Analysis

Students took 30 minutes to complete a pre-test on fraction addition and subtraction. Then, spend another 20 minutes completing a questionnaire about pre-learning mathematics. After completing a pre-test, students received a simulation-based integrated learning arrangement. During data collection, students must discover solutions through PhET simulation, showcase

how to solve problems with Nearpod and reflect on each other. The teacher explains how to use the simulation and the Nearpod application and how to find answers and share them with the class through discussion. The students received a post-test and a questionnaire. (See Figure 1)



*Figure 1.* Examples of teaching materials: (A) is hands-on material, (B) is interactive PhET simulation.

The statistical data techniques selected for analyzing students' conception and attitude toward using technology for learning mathematics was Wilcoxon matched pairs signed-ranks test in SPSS 22.0.

#### 4.4 Learning Materials

In the part of the learning activity, technology materials that support the learning process are interactive PhET simulation and interactive presentation from the Nearpod application. Students could also use all applications on their mobile devices. First, the simulation can easily visualize fractional transformations and increase understanding of fraction content meaning of fractions and comparing fractions as well. Moreover, it can change numbers immediately. Whether it's the numerator or the denominator, the picture will also change. This is good for students who want to use their understanding more. Second, the Nearpod application was an interactive classroom tool that enabled teachers to create engaging learning experiences by providing interactive presentations, collaboration, and real-time assessment tools into one integrated solution. Students joined writing solutions on fractions through the Nearpod application. Then, students shared each of the students' writing solutions on fractions with the rest of the class and discussed together along with a summary of mathematical concepts via a collaborative board as show in Figure 2.



*Figure 2.* Examples of teaching materials: (A) is PhET simulation, (B) and (C) are Nearpod applications show students' solution and discussion about the answer.

#### 5. Result and Discussion

#### 5.1 Student's Conceptual Understanding in Addition and Subtraction of Fraction

The present study employed nonparametric statistics analysis, specifically the Wilcoxon signed-ranks test, to examine the difference between pre-test and post-test conceptual understanding scores. The results showed a significant increase in students' post-test scores for conceptual understanding of addition and subtraction of fraction concepts compared to their pre-test scores (Z=-4.385, p<.05) as shown in Table 2.

Table 2. Statistics Of Wilcoxon Matched Pairs Signed-Ranks Test for Students' Conceptual Understanding Scores

		Ν	Mean Rank	Sum of Ranks	Z	Asymp. Sig.(2-tails)
Post-Pre	Negative Ranks	0 <sup>a</sup>	0	0		
	Positive Ranks	25 <sup>b</sup>	13	325	-4.385	<.001*
	Equal	0 <sup>c</sup>	-			

\* p= <.05

The quantitative result indicates the post-test score of students' conceptual understanding of addition and subtraction fractions was significantly higher than the pre-test score. This finding revealed that the conceptual understanding of students who learn with blended learning with PhET simulation is better after learning. The results related to Arifin et al. (2022) studied the effectiveness of using PhET simulation in primary school students' mathematical understanding of fractions. The results showed that using simulation significantly improved students' mathematical understanding of fractions.

#### 5.2 Student's Attitude toward Using Technology for Learning Mathematics

The Wilcoxon signed-rank test analysis found that there were significant differences between the pretest and post-test for attitude to learning mathematics with technology is mathematics confidence (MC) (Z = -3.733, p < .001), attitude towards the use of technology for learning mathematics (MT) (Z = -4.121, p < .001), confidence with technology (TC) (Z = -3.907, p < .001), affective engagement (AE) (Z = -3.856, p < .001) and behavioral engagement (BE) (Z = -3.805, p < .001) as shown in Table 3

		Ν	Mean Rank	Sum of Ranks	Z	Asymp. Sig.(2-tails)
MC	Negative Ranks	3 <sup>a</sup>	5.17	15.5		
	Positive Ranks	20 <sup>b</sup>	13.03	260.5	-3.733	<.001*
	Equal	2 <sup>c</sup>				
MT	Negative Ranks	1 <sup>d</sup>	6	6		
	Positive Ranks	23 <sup>e</sup>	12.78	294	-4.121	<.001*
	Equal	1 <sup>f</sup>				
TC	Negative Ranks	2 <sup>g</sup>	6.75	13.5		
	Positive Ranks	22 <sup>h</sup>	13.02	286.5	-3.907	<.001*
	Equal	1 <sup>i</sup>				

Table 3. Statistics of Wilcoxon matched p	pairs signed-ranks test for attitude to learning
---	--

		Ν	Mean Rank	Sum of Ranks	Z	Asymp. Sig.(2-tails)
AE	Negative Ranks	2 <sup>j</sup>	4	8		
	Positive Ranks	20 <sup>k</sup>	12.25	245	-3.856	<.001*
	Equal	3 <sup>1</sup>				
BE	Negative Ranks	3 <sup>m</sup>	5.67			
	Positive Ranks	<b>21</b> <sup>n</sup>	13.48		-3.805	<.001*
	Equal	1°				

\* p= <.05

From the study, the overall outcome was students' attitude toward using technology for learning mathematics. There are good results from using technology in mathematics classrooms, as simulations keep students excited, interested. The students liked it and gave it a lot of attention. As well as showing how to do the problem, students can write on iPads. Easy to write, edit the answer, present students' ideas, type, and write along with seeing classmates' answers for talk and discussion. In the meantime, Emata (2023) examined the relationship between math self-efficacy, technology attitude, and attitudes toward mathematics among students. The study found a moderate level of math self-efficacy and attitudes towards math among the students, while the level of technology attitude is high. So, from the study's results, technology should be adapted to teaching and learning.

## 5.3 Student's Interview on the Attitude toward Using Technology for Learning Mathematics

This study found that students' attitudes toward using technology for learning mathematics. The researchers chose to explain each of the five types of explanations with an example from a student.

#### 5.3.1 Mathematics Confidence: Student A said:

"I'm not afraid to make a math problem. Because it's easy to edit from the iPad without a pencil and eraser."

5.3.2 Attitude Towards Use of Technology for Learning Mathematics: Student B said:

"I love to use apps for math. because normal learning is boring, I like math, but I don't like writing numbers and answers on books. I like math games and anything else that can't just be learned from books"

5.3.3 Confidence with Technology: Student C said:

"I like to use iPad because it's very easy to use, easy to write, easy to click. I want to use iPad in every class."

#### 5.3.4 Affective Engagement: Student D said:

"I'm so excited to use iPad to learn math because I love to see and share my solution with my classmates through the app."

#### 5.3.5 Behavioral Engagement: Student E said:

"I am very focused when studying math. Because I have to keep up with the teacher, I wish I had more time to do the problem. because the teacher gives little timer"

That means students are not worried about showing how to do or responding to answers through technology. Because technology makes it easy to write and answer, students

do not enjoy learning solely from books. They enjoy using technology to learn, need to talk to their classmates about the answers they have received, and are determined to learn, making efforts to listen and focus on their studies to keep up with both the teacher and their classmates in the classroom. This qualitative study showed students' attitudes to learning mathematics with technology. The result indicated that many students described an attitude related to mathematics, confidence, attitude towards the use of technology for learning mathematics, confidence in technology, affective engagement, and behavioral engagement.

#### 6. Conclusion and Limitation

Based on the research findings, it can be concluded that blended learning with technology can better resolve misconceptions about addition and subtraction of fractions. Furthermore, PhET simulations can help students visualize and increase their comprehension. Additionally, students interact through the Nearpod application that could exchange their ideas with classmates immediately. Moreover, students have positive attitude to learning mathematics with technology. It's implied that teaching mathematics with blended learning toward technologies had an effect on students' conceptual understanding and attitude toward technology in fraction concept in mathematics classroom.

Based on the results of the research, technology plays an essential role in the correction of mathematical misconception. Technologies that facilitate learning, such as PhET and Nearpod, are widely available and compatible with various devices, including smartphones, tablets, and laptops. These technologies have the potential to enhance student achievement by improving teaching and learning processes. Conceptual understanding can be attained through all six stages, with the most pronounced expression occurring throughout the stage of Brainstorming possible solutions. This stage involves the utilization of physical materials, including the stage of Researching the problem and developing a solution, where the use of PhET is employed. Consequently, this process facilitates the formation of precise concepts. One potential limitation of this study is its reliance on a single sample, hence resulting in a limited sample size. Consequently, researchers are unable to make comparisons between scores, behaviors, and students' attitudes towards learning in relation to both traditional teaching methods and blended teaching incorporating simulation. The sample utilized in the study did not yield results that can be generalized to the general population. In future investigations, it is recommended that researchers conduct trials on a substantial cohort of students.

#### Acknowledgements

This work was supported by Khon Kaen University, Khon Kaen, Thailand. The authors would like to express sincere thanks to the research team for all the big help.

#### References

 Adewuyi, M. A. (2001). Teacher and student related variables as correlates of achievement in mathematics in Oyo State of Nigeria. Unpublished M.Ed. dissertation. University of Ibadan, Ibadan.
 Akinsola, M. K. (2002). Instructional methods employed by mathematics teachers: A managerial

approach. *African Journal of Educational Planning and Policy Studies*, 3(1), 25-32. Akinsola, M. K. (2007). The effect of simulation-games environment on students' achievement in and

attitudes toward mathematics in secondary schools. *The Turkish Online Journal of Educational Technology*, 6(3), 11.

Aksoy, N., & Yazlik, D. (2017). Student errors in fractions and possible causes of these errors. *Journal of Education and Training Studies*, *5*(11), 219-233. https://doi.org/10.11114/jets.v5i11.2679

Arifin, S., Febby, F., Nugraha, A., & Prasetyo. (2021). Mathematics Learning Design using Phet Interactive Simulation to Support Students Mathematical Understanding. *Journal of Physics: Conference Series*, 1775(1), 012019. doi: 10.1088/1742-6596/1775/1/012019.

- Barrett, B., & Sharma, P. (2007). *Blended Learning: Using technology in and*. Oxford, England: Macmillan Publishers Limited.
- D'Angelo, C., Rutstein, D., Harris, C., Bernard, R., Borokhovski, E., & Haertel, G. (2014). *Simulations for STEM learning: Systematic review and meta-analysis.* Menlo Park, CA: SRI International.
- Debbie, M., Verzosa, M. L., De Las Peñas, A., Aberin, Q., Garciano, A., Sarmiento, J., Mallari, J. C., & Tolentino, M. A. C. (2021). Development of an app and videos to support the fraction learning trajectory from grades 1-7. *Proceedings of the 30th International Conference on Computers in Education* 336-344. Asia-Pacific Society for Computers in Education, 11(2), 1-16.
- Emata, C. (2023). The moderating effect of technology attitude on the relationship between math selfefficacy and attitudes towards mathematics among Science, Technology and Engineering (STE) students. *Unnes Journal of Mathematics Education*, *12*(1), 1-12.
- Hsu, Y. (2021). Exploring the learning motivation and effectiveness of applying virtual reality to high school mathematics. *Journal of Educational Technology & Society*, *24*(1), 1-12.
- Indrapangastuti, D., Surjono, H., & Yanto, B. (2020). Effectiveness of the blended learning model to improve students' achievement of mathematical concepts. *Journal of Physics: Conference Series*, 1462(1), 012031. <u>https://doi.org/10.1088/1742-6596/1462/1/012031</u>
- Jemakmun. (2022). Application of Blended Learning to Increase Mathematics Learning Outcomes from Beginning Mathematics Ability. *International Journal of Education, Teaching, and Social Science*, 2(4), October 2022.
- M. Driscoll. (2002). Consultants point of view. Retrieved from https://www-07.ibm.com/services/pdf/blended learning.pdf
- Marsh, J., & Drexler, P. (2001). *How to design effective blended learning*. Sunnyvale, CA: Brandon-Hall.
- Meadows, Caniglia. (2019). Using PhET Simulations in the Mathematics Classroom. *Mathematics Teacher*, *112*(5), 386.
- Mistretta, R. M. (2005). Integrating technology into the mathematics classroom: The role of teacher preparation programs. *The Mathematics Educator*, *15*(1), 18–24.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.
- Pandya, M. D., Shroff, N., & Hamzah, M. I. (2021). Blended Learning Readiness: Understanding Attitude of Undergraduate Students Towards Blended Learning Modalities. *Journal of Educational Technology & Society*, 24(1), 1-14.
- Pierce, R., Stacey, K., & Barkatsas, A. (2007). A scale for monitoring students' attitudes to learning mathematics with technology. *Computers & Education*, *48*(2007), 285–300.
- Pulungan, R. (2017). Analysis of Student's Misconceptions and Mistakes in Learning Fraction. *Journal* of *Physics: Conference Series*, 943(1), 012119. doi: 10.1088/1742-6596/943/1/012119.
- Reinhardt, E., & Loftsgardeen, D. O. (1979). Using simulation to resolve probability paradoxes. International Journal of Mathematics Education in Science and Technology, 10(2), 241–250.
- Retnawati, H., Kartowagiran, B., Hadi, S., & Hidayati, K. (2011). Identification of students' difficulties in learning mathematics and science in elementary schools. *Journal of Education*, *41*(2), 162-174.
- Spooner, M. (2002). *Errors and misconceptions in maths at key stage 2. Working towards a successful SATs.* London: David Fulton Publishers.
- Srisawasdi, N., Junphon, S., & Panjaburee, P. (2015). Effect of Simulation-based Inquiry with Dualsituated Learning Model on Change of Student's Conception. *Procedia-Social and Behavioral Sciences*, 174, 3279-3286. https://doi.org/10.1016/j.sbspro.2015.01.1049
- Trivena, V., Ningsih, A., & Jupri, A. (2017). Misconception on addition and subtraction of fraction at primary school students in fifth-grade. *Journal of Physics: Conference Series*, 895, 1-7. https://doi.org/10.1088/1742-6596/895/1/012139
- Waskitoningtyas, R. S. (2016). Analysis of learning difficulties in mathematics for fifth grade elementary school students in Balikpapan in the time unit material for the 2015/2016 academic year. *Scientific Journal of Mathematics Education*, 5(1), 24-32.
- Wenglinsky, H. (1998). Does it compute? The relationship between educational technology and student achievement in mathematics. Princeton, NJ: Educational Testing Service.
- Yennita, Y., & Zukmadini, A. (2020). Problem-based learning (PBL) and blended learning in improving critical thinking skills and student learning activities in biochemistry courses. *Journal of Physics: Conference Series*, 1467(1), 012031.

## Design and Development of a Personalized Recommender System of Student Question-Generation Exercises for Programming Courses

Chih-Hung LAI<sup>a</sup>, Pham-Duc THO<sup>b\*</sup>

<sup>a</sup>Department of Computer Science and Information Engineering, National Dong Hwa University, Hualien, Taiwan, Republic of China <sup>b</sup>International School, Vietnam National University, Hanoi, Vietnam \*mrthoedu@gmail.com

**Abstract:** Computer programming courses often face the challenge of high dropout rates, prompting educators to seek effective solutions to improve student engagement and retention. Student Question-Generation (SQG) have been explored as potential remedy to this issue, enabling advancements in computer science education. However, the majority of existing SQG support systems lack personalized features to enhance student learning. In response to this gap, the study introduces a personal recommender system of SQG exercise tailored for programming courses. The system leverages individual student preferences and SQG exercise complexities to personalize learning experiences within SQG activities. This paper presents the design of the system and its user interface and delves into motivations and technical bases underlying its development.

**Keywords:** programming course, Student Generated Question, Recommender System, collaborative filtering

#### 1. Introduction

Computer programming is a foundational part of computer science education, and many educational institutions are updating their curricula to address the challenges in programming courses. High student failure and dropout rates in traditional classrooms are attributed to factors like limited class time, resource constraints, and more. Scholars propose remedies, including collaborative learning, problem-solving techniques (Wang et al., 2017), visualization tools (Minjie, 2015), psychological analysis, visual programming environments (Minjie, 2015), and student-generated questions (Andrew et al., 2011; Hsu & Wang, 2018; Lai et al., 2017).

Among these solutions, SQG have emerged as a critical pedagogical strategy with long-standing recognition (Barak & Rafaeli, 2004). SQG has been noted to enhance comprehension of acquired knowledge, and cultivate skills in algorithmic thinking (Hsu & Wang, 2018), motivation, engagement and learning performance (Crogman & Crogman, 2018; Yeong et al., 2019). Through SQG, students are empowered to apply learned concepts to novel challenges and bridge newly acquired knowledge/skills with prior understanding (Lai et al., 2017). This approach encourages students to adopt an alternative cognitive mode and enhance their repertoire of learning strategies (Hsu & Wang, 2018; Yu & Chen, 2014).

Despite the positive outcomes observed, there were some noteworthy concerns surrounding the implementation of this approach. Specifically, it came to light that certain students expressed a lack of motivation when engaging with assignments involving SQG, as proposed by the instructor throughout the course. This lack of motivation was evident in cases where high-achieving students were presented with relatively straightforward SQG exercises, or conversely, where less experienced students were confronted with overly complex ones.

To address this challenge and fully capitalize on the advantages of SQG, this research introduces an innovative system known as PERS (Personalized Exercise Recommender System). PERS builds upon the SQG concept by incorporating a Recommender System (RS) that suggests appropriate SQG exercises or problem statements to individual students. Operating on the principles of collaborative filtering, PERS intelligently recommends SQG exercises based on students' interests and their assessed proficiency level.

#### 2. Related work

#### 2.1 Question generation and programming learning

Question Generation, also referred to as Question Posing, involves the active process wherein students formulate examination questions based on the content they have read. Even in the face of unfamiliar scenarios or exceptional circumstances, individuals, including children, naturally pose questions. Constructivist educational theories emphasize the significance of this innate curiosity as a driving force for learning and advocate for instructors to facilitate the emergence of such spontaneous inquiry. This facilitative role of the instructor is seen as pivotal in nurturing the inherent questioning process.

Question types span diverse formats, encompassing coding exercises, multiplechoice, matching, short-answer, true-false, fill-in-the-blank, and word puzzles (Barak & Rafaeli, 2004; Lai et al., 2017). When generating questions, students are required to engage deeply with the pertinent information in their texts and subsequently formulate correct answers and distractors (Lai et al., 2017). The ability of students to elucidate the reasoning behind their crafted answer options serves as an indicator of their grasp of the reading material (Lan & Lin, 2011;). Furthermore, question generation allows instructors to identify students' comprehension challenges, thereby enabling targeted instructional interventions (Lan & Lin, 2011; Yeh & Lai, 2012).

Numerous web-based learning systems have been developed to foster student question generation, including QPPA (Yu et al., 2002), Question-Authoring and Reasoning Knowledge System (QuARKS) (Yu, 2009), CodeWrite (Denny et al., 2011), StudySieve (Andrew et al., 2011), and PIPLS (Lai & Tho, 2016). While many of these systems are domain-agnostic, accommodating various question types and multimedia content, and supporting anonymous interactions, only a limited few are specifically tailored for programming courses (Denny et al., 2011; Lai & Tho, 2016; Reilly, 2012) and none of them incorporate automated methods to recommend SQG exercises to students.

#### 2.2 Recommender systems and learning environments

A RS is a software tool designed to discern and propose content that a specific user would find valuable (Ricci et al., 2011). It falls within the realm of information filtering systems, which exploit user data to predict ratings or preferences a user might assign to particular items. Consequently, the primary advantage of an RS is its ability to identify the most suitable array of items for a given user, optimizing rating predictions.

Ricci et al. (2011) delineated five types of RSs: content-based, knowledge-based, demographic, community-based, collaborative, and hybrid. Collaborative Filtering RSs (CFRSs) have seen extensive utilization (Elahi et al., 2016). CFRSs operate on the premise that users may favor items endorsed by other users with similar preferences in the past. This principle underpins the RS employed in this study. CFRSs consist of two main approaches: user-user and item-item (Elahi et al., 2016), both commonly employing the Nearest Neighbors algorithm (Nikolakopoulos et al., 2021).

RSs have found broad application in e-learning environments within the scope of Technology-Enhanced Learning (TEL) to enhance students' self-directed learning (Manouselis et al., 2011). In e-commerce, RSs suggest products; in e-learning, they recommend educational resources (such as papers, books, or courses) to participants like students and teachers (Liu et al., 2022).

RSs have been analyzed for their application in e-learning systems (Santos & Boticario, 2011), presenting three technological prerequisites for developing semantic education RSs.

Enhancing RSs in collaborative learning environments, Anaya et al. (2013) introduced an influence diagram employing machine learning to assess user collaboration, leading to the development of an automatic RS with a pedagogical decision tree. A cloud-based architecture for recommending learning elements based on the learner's emotional state was also proposed (Leony Arreaga et al., 2013).

An extensive survey on RS evaluation in the context of TEL highlighted the need for better evaluation strategies (Erdt et al., 2015). In conclusion, the review underscores the popularity of RS in education, especially technology-driven approaches for enhancing learning. Furthermore, it notes the absence of RSs supporting programming learning with SQG, which forms the rationale for the proposed system.

#### 3. PERS: An integrated approach

In the traditional SQG approach, instructors utilize classroom lectures and student-created SQGs, which are added to a shared Question Bank. Challenges involve overseeing student interaction and addressing dissatisfaction, especially among advanced students who may find exercises too simple. Another approach is more lab practice with instructor oversight, offering better control over SQG exercise interaction. Yet, it risks overemphasizing formal learning, conflicting with broader goals. The challenge is enhancing the current method while reducing direct instructor involvement, maintaining its advantages.

In response to these considerations, this study introduces PERS, a web application that integrates SQG as a teaching tool along with a RS for peer exercises. It is important to note that this proposal not only facilitates interactions similar to the existing approach but also personalizes students' learning journeys through SQG exercises. This empowers the instructor to guide students' learning paths by creating exercises integrated into the system. Moreover, students are afforded the opportunity to evaluate and indicate their preferences for SQG exercises, enabling the system to gain insights into their individual preferences. Using this feedback, the system generates suggestions for new SQG exercises tailored to each student, capitalizing on the notion that students with shared preferences and perceptions of exercise complexity are prime sources for recommendations.

From an educational perspective, this proposition emerges as a novel intermediary between instructors and students. In the ensuing section, the proposal is presented in meticulous detail, outlining the developed Web Application's principal modules and features. Subsequently, the technical elements of the implemented RS are elucidated.

The primary interaction workflow between students and the proposed system is illustrated in Figure 1. Upon logging in, the system assesses whether the student has previously encountered SQG exercises. If they have, the system employs a collaborative filtering approach to recommend 10 SQG exercises. If not, the system selects the most approachable and engaging SQG exercises from the available pool. Additionally, the system assembles a list of 10 SQG exercises intended for presentation to students.

The subsequent stages can be readily comprehended by referring to the diagram. However, it's noteworthy that following the submission of evaluations for the SQG exercises, the system employs collaborative filtering to furnish the student with a fresh set of recommended SQG exercises. As a result, students may engage with varied SQG exercises based on their individual interactions.

Consequently, the student profile is developed based on their evaluations (Favorites and Complexity) of the completed SQG exercises. The current version of PERS was crafted using Python and employs the PostgreSQL Server as its underlying database system.

Among its noteworthy features, PERS encompasses:

- User Authentication Area: This area permits the entry of previously registered students and teachers.
- Student Interface: Students benefit from an informative panel (as depicted in the right panel of Figure 2), which empowers them to manage their profiles,



access a summary of their completed SQG exercises, and review the evaluations of those exercises.

*Figure 1.* The main workflow of the interaction between the student and the proposed system

<i>~</i>	C Fundame	ntals	ŵ	Course Contents	
	3 Lesson 1h 30min			2/5 COMPLETED	ë
kercise				Introducing C © 1 Hour	¥
te a program to put their value. F input: 12 10 8 output must be	compute the volume o or example, if length is shown in the following	f a box. You must declare three vari 12, width is 10, height is 8, volume is format: The volume is 960	ables for length, wiath, neight, and volume of a box and 960.	C Fundamentals © 2 Hour	C 3 Lessons
Your Answer				1. Formotted input/Output	35:00
				2. Selection Statements	25:00
				3. Expressions	30:00
				Basic Types ③ 1 Hour	V 🖾 4 Lessons
				Loops (3) 12:54	✓ []] 5 Lessons
			Submit Now	Arrays (0) 59:00	V
Your Feed	back		Your Recommended Exe	rcises	Ð
Question Fo Question Di	ivourite fficulty	**** ****	Add fractions	Compute the vol	ume of a

Figure 2: PERS front-end

- Central Display Area: The central region (as shown in Figure 2) is tailored to exhibit the description of the chosen SQG exercise along with its evaluation criteria.
- Bottom-Right Area: The section located at the bottom-right (depicted in Figure 2) provides either a roster of recommended SQG exercises or an exhaustive list of all SQG exercises.
- Teacher Profile: In the teacher's profile, educators can seamlessly integrate new SQG exercises and gain access to statistical insights regarding both the SQG exercises and student interactions.

Importantly, the inclusion of the RS within PERS was realized through the utilization of Surprise. This Python Scikit, abbreviated as SciPy Toolkits, facilitates the creation and evaluation of recommender systems, underscoring PERS's technological foundation.

#### 4. Conclusion and future work

In this study, PERS, a user-friendly Web App is introduced seamlessly combining SQG with a robust recommender system to support students in learning Programming more effectively compared to other systems. In the forthcoming stages, this study will improve the functionalities of the existing systems and assessing the impact of the system on students' academic performance. The insights gleaned from this study will stimulate heightened interest among researchers in this domain.

#### Acknowledgements

This research was funded by National Science and Technology Council, Taiwan, R.O.C. (Contract NO: MOST 111-2410-H-259-009-)

#### References

- Anaya, A. R., Luque, M., & García-Saiz, T. (2013). Recommender system in collaborative learning environment using an influence diagram. *Expert Systems with Applications, 40*(18), 7193-7202.
- Andrew, L. R., Denny, P., Plimmer, B., & Bertinshaw, D. (2011). Supporting student-generated freeresponse questions. 16th annual joint conference on Innovation and technology in computer science education,
- Barak, M., & Rafaeli, S. (2004, 7//). On-line question-posing and peer-assessment as means for webbased knowledge sharing in learning. *International Journal of Human-Computer Studies*, 61(1), 84-103. https://doi.org/10.1016/j.ijhcs.2003.12.005
- Berry, J. W., & Chew, S. L. (2008). Improving Learning Through Interventions of Student-Generated Questions and Concept Maps. *Teaching of Psychology*, *35*(4), 305-312, Article Pii 905075619. https://doi.org/10.1080/00986280802373841
- Chang, S. B., Huang, H. M., Tung, K. J., & Chan, T. W. (2005). AGQ: a model of student question generation supported by one-on-one educational computing. 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!,
- Crogman, H., & Crogman, M. T. (2018, May). Modified generated question learning, and its classroom implementation and assessment [Article]. *Cogent Education, 5*(1), 1-41, Article 1459340.
- Denny, P., Luxton Reilly, A., Tempero, E., & Hendrickx, J. (2011). CodeWrite: supporting student-driven practice of java. 42nd ACM technical symposium on Computer science education,
- Elahi, M., Ricci, F., & Rubens, N. (2016, 2016/05/01/). A survey of active learning in collaborative filtering recommender systems. *Computer Science Review, 20*, 29-50. https://doi.org/https://doi.org/10.1016/j.cosrev.2016.05.002
- Erdt, M., Fernandez, A., & Rensing, C. (2015). Evaluating recommender systems for technology enhanced learning: a quantitative survey. *leee Transactions on Learning Technologies*, *8*(4), 326-344.
- Hsu, C. C., & Wang, T. I. (2018, 2018/06/01/). Applying game mechanics and student-generated questions to an online puzzle-based game learning system to promote algorithmic thinking skills. *Computers & Education, 121*, 73-88. https://doi.org/10.1016/j.compedu.2018.02.002
- Lai, C.-H., Tho, P.-D., & Liang, J.-S. (2017). Design and evaluation of question-generated programming learning system. 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI),

- Lai, C. H., & Tho, P.-D. (2016). Development of a Programming Learning System Based on a Question Generated strategy. 24th International Conference on Computers in Education, India.
- Lan, Y. F., & Lin, P. C. (2011). Evaluation and improvement of student's question-posing ability in a web-based learning environment [Article]. *Australasian Journal of Educational Technology*, 27(4), 581-599.
- Leony Arreaga, D. A., Parada Gélvez, H. A., Muñoz Merino, P. J., Pardo Sánchez, A., & Delgado Kloos, C. (2013). A generic architecture for emotion-based recommender systems in cloud learning environments.
- Leovy, E., Ruth, C., Liliana, M., & Ivan, C. (2017). Using collaborative learning scenarios to teach programming to non-CS majors. *Computer Applications in Engineering Education*, 25(5), 719-731. https://doi.org/10.1002/cae.21832
- Linxiao, M. (2007). *Investigating and Improving Novice Programmers' Mental Models of Programming Concepts* [Dissertation, University of Strathclyde]. United Kingdom.
- Liu, T., Wu, Q., Chang, L., & Gu, T. (2022, 2022/12/01). A review of deep learning-based recommender system in e-learning environments. *Artificial Intelligence Review*, *55*(8), 5953-5980. https://doi.org/10.1007/s10462-022-10135-2
- Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., & Koper, R. (2011). Recommender systems in technology enhanced learning. *Recommender systems handbook*, 387-415.
- Minjie, H. (2015). *Teaching Novices Programming: A Programming Process Using Goals & Plans with a Visual Programming Environment* University of Otago]. Dunedin, New Zealand.
- Nikolakopoulos, A. N., Ning, X., Desrosiers, C., & Karypis, G. (2021). Trust your neighbors: a comprehensive survey of neighborhood-based methods for recommender systems. *Recommender systems handbook*, 39-89.
- Piteira, M., Costa, C. J., & Aparicio, M. (2018). Computer programming learning: how to apply gamification on online courses? *Journal of Information Systems Engineering & Management, 3*(2), 11-20.
- Reilly, A. L. (2012). *The Design and Evaluation of StudySieve: A Tool That Supports Student-Generated Free-Response Questions, Answers and Evaluations.* [Dissertation, The University of Auckland]. NewZealand.
- Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to Recommender Systems Handbook. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), *Recommender systems handbook* (pp. 1-35). Springer US. https://doi.org/10.1007/978-0-387-85820-3\_1
- Rosenshine, B., Meister, C., & Chapman, S. (1996, Sum). Teaching students to generate questions: A review of the intervention studies [Review]. *Review of Educational Research*, 66(2), 181-221.
- Santos, O. C., & Boticario, J. G. (2011). Requirements for semantic educational recommender systems in formal e-learning scenarios. *Algorithms*, *4*(3), 131-154.
- Vladimiras, D., Tatjana, J., & Valentina, D. (2018). From Android games to coding in C—An approach to motivate novice engineering students to learn programming: A case study. *Computer Applications in Engineering Education*, *26*(1), 75-90. https://doi.org/10.1002/cae.21862
- Wan, X., Jamaliding, Q., & Okamoto, T. (2011). Analyzing Learners' Relationship to Improve the Quality of Recommender System for Group Learning Support. *J. Comput., 6*(2), 254-262.
- Wan, X., & Okamoto, T. (2011). Utilizing learning process to improve recommender system for group learning support. *Neural Computing and Applications, 20*, 611-621.
- Wilson, E. V. (2004). ExamNet asynchronous learning network: augmenting face-to-face courses with student-developed exam questions. *Computers & Education, 42*(1), 87-107.
- Yeh, H. C., & Lai, P. Y. (2012). Implementing online question generation to foster reading comprehension [Article]. Australasian Journal of Educational Technology, 28(7), 1152-1175. <Go to ISI>://WOS:000314400200006
- Yeong, F. M., Chin, C. F., & Tan, A. L. (2019). Use of a competency framework to explore the benefits of student-generated multiple-choice questions (MCQs) on student engagement. *Pedagogies: An International Journal*, 1-23. https://doi.org/10.1080/1554480X.2019.1684924
- Yu, F. Y., & Chen, Y. J. (2014, Mar). Effects of student-generated questions as the source of online drill-and-practice activities on learning. *British Journal of Educational Technology*, *45*(2), 316-329.
- Yu, F. Y., Wu, W. S., & Huang, H. C. (2018, Oct). Promoting Middle School Students' Learning Motivation and Academic Emotions via Student-Created Feedback for Online Student-Created Multiple-Choice Questions [Article]. Asia-Pacific Education Researcher, 27(5), 395-408. https://doi.org/10.1007/s40299-018-0398-x

## A Chatbot for Image Recommendation in Mobile Language Learning

#### Mohammad Nehal HASNINE<sup>a\*</sup>, Huyen T. T. BUI<sup>a</sup> & Hiroshi UEDA<sup>a</sup>

<sup>a</sup>Research Center for Computing and Multimedia Studies, Hosei University, Japan \*nehal.hasnine.79@hosei.ac.jp

**Abstract:** Mobile language learning applications have already shown great promise and have played a significant role in how language learners learn foreign vocabulary. Many mobile language learning applications such as mobile apps, computer programs, chatbots, and electronic dictionaries are developed to support vocabulary learning. Finding images to represent a word, considering the learning contexts, is challenging. Therefore, this article describes the development of a chatbot to provide foreign language learners with images to represent a word to memorize using Telegram messenger platforms. Using Python, Microsoft Azure services, and the Telegram API, this chatbot has been designed to provide five relevant still images from the image search engine to get the images to represent a word immediately. This minimizes learners' time spent searching images and optimizes the planning for their learning activities. This chatbot, AIVAS-iBOT (Appropriate Image-based Vocabulary Acquisition System- Image BOT), was also deployed in Skype and hence could be accessed from Telegram and Skype.

**Keywords:** AIVAS-iBOT, chatbot for language learner, image recommendation, vocabulary acquisition

#### 1. Introduction

Mobile language learning, known under the umbrella terms mobile learning or ubiquitous learning, has shown promise in delivering close integration of language learning with learners' communication needs and cultural experiences. Because of this, learners can use smartphone functionalities to record experiences and listen to audio at any time, which encourages spontaneous interaction.

Foreign vocabulary learning using mobile language learning has been prevalent in recent years, especially in Asian countries such as Japan, China, Hong Kong, Thailand, and Taiwan. Vocabulary learning research indicated that foreign words are easy to memorize when it is represented with an image. It is because images have a significant role to play in our memory. Hence it is said that a picture is worth a thousand words. Many language learning textbooks extensively use images to create the learning content. Many language teachers and language learning apps use a combination of images to improve language learning. However, finding relevant images for a word that could be used as educational resources takes time and effort for the learners.

Educational chatbots have gained interest in the past years due to the advanced artificial intelligence algorithms used in searching content and recommending relevant content to learners. Typically, a chatbot is a program supporting user interaction via conversation in natural language, and it is accessible through the web or social networks (Pérez-Soler et al., 2021).

This paper aims to develop an intelligent chatbot, AIVAS-iBOT (Appropriate Imagebased Vocabulary Acquisition System- Image BOT), for mobile language learners using natural language processing and Telegram. This chatbot is also deployed on Skype and can be downloaded from Skype. This research aims to assist foreign language learners in getting familiar with an educational chatbot environment and minimize image search time.

#### 2. Literature Review

#### 2.1 Chatbots in Education

In education, chatbots are becoming an increasingly popular option to interact with learners, and their popularity and adoption are rapidly spreading (Smutny & Schreiberova, 2020). For example, a review article by (Smutny & Schreiberova, 2020) listed 47 educational chatbots using the Facebook Messenger platform. Consequently, a systematic review by (Kuhail et al., 2023) reported 36 papers to understand, compare, and reflect on recent attempts to utilize chatbots in education. The main emphasis of using chatbots in education is to improve learners' learning experiences and learning outcomes. Some examples of educational chatbots are ELIZA, ALICE, and SmartChild. At the Botlist directory, over 2000 chatbots built and deployed on various platforms (e.g. Amazon Echo, Skype, and Slack) could be found.

#### 2.2 Image Recommendation Research in Language Learning

Image recommendation systems have became popular in various domains, such as ecommerce and product recommendation. However, few research studies have examined the scope of image recommendation for language learning. One research by (Hasnine et al., 2018) built an image recommendation environment for language learners. This study proposed a model for Feature-based Context-specific Appropriate Image (FCAI) recommendation to mobile language learners. However, this model has yet to be implemented in a chatbot environment. Also, a few chatbots are built to support image recommendation for word memorization.

Therefore, we introduce AIBAS-iBOT, a chatbot for image search and recommendation to mobile language learners in this study. This educational tool can be downloaded from Telegram and Skype messages.

#### 3. AIRS-iBOT

#### 3.1 Design Process and Methods



Figure 1 shows the flowchart of the AIVAS-iBOT.



In the design process, Question 1 defines the language for the chatbot's user interface. Question 2 indicates the country where the results come from, and Question 3 indicates what you want to search for. Each question has some button to guide user action; it enhances the user experience by selecting the available button or directly entering the question/answer.

After getting a list of results, AIRS-IBOT will give the information on the next step. Here a learner will have some choices for continuing using the chatbot. They are:

- Continue Learning: It will go back to the first question, and you will continue to use the chatbot
- Developer's Information: It gives the information of the developer of this chatbot.
- Quit: Quit the operation of the chatbot

#### 3.2 Integration with Telegram and Skype

Telegram (<u>https://telegram.org</u>) and Skype (<u>https://www.skype.com/en/</u>) are popular instant messaging platforms. According to the Telegram website (<u>https://telegram.org/apps</u>), Telegram apps are open source and support reproducible builds. Anyone can independently verify that the Telegram apps you download from the App Store or Google Play were built using the same code that we publish. Skype website (<u>https://www.skype.com/en/features/</u>) suggests that hosting a video or an audio conference with up to 100 people, recording your calls, enabling live captions and subtitles or just simply talking over their smart chat platform is possible at Skype.

Due to these benefits, many apps and chatbots are built and integrated into Telegram and Skype. Telegram and Skype have 550 million and 2.1 billion users, respectively. Many of them use these platforms for educational purposes. Therefore, we aimed to integrate AIVAS-iBOT with Telegram and Skype. Figure 1 shows the integration.

#### 3.3 Deploy in Telegram

At first, we created a Telegram account to create our chatbot on Telegram Bot. To start a new chatbot in Telegram, search for "BotFather" – a bot for managing all Telegram Bots and start registering the new bot.

- Send /start to start a new conversation
- Send /newbot to create a new Telegram bot.
- Configure Telegram in the Azure portal with an access token field



Figure 1. AIVAS-iBOT Integration with Telegram and Skype.



Figure 2. Access Token to Configure Telegram in Azure portal.

#### 3.4 Image Recommendation

AIVAS-iBOT generates results upon answers to three questions by a learner, as follows:

First, *Where are you from*? Answers to this question give the chatbot an idea of a learner's cultural background.

Second, *Which culture are you interested in?* Answers to this question give the chatbot an idea about the language the learner is interested in.

Third, *What are you looking for*? Answers to this question give the chatbot an idea about the topic the learner is looking for.



Figure 3. Image Recommendation in AIVAS-iBOT.

For each question, the AIVAS-iBOT suggests some options. So, a learner can choose from those options without typing his/her option if there answer matches with AIVAS-iBOT suggested options.

#### 4. Summary

In this project, we developed a chatbot that returns appropriate images relevant to the language learners' search query. We use the Azure CLI to create AIRS-iBOT and use Bing Search API v7 as the function for searching images. For developing this Chatbot, we used Python as the programming language and Bot Framework Emulator to test and debug the bot. With Bing Search API, we can: AnswerCount. This means we can set the number of answers that we want the response to include. Also, we can Count Code. This means understanding the country where the results come from. In addition, Bing Search API can support many languages, the language to use for user interface strings.

Using AIVAS-iBOT, a mobile language learner gets appropriate images to represent a word he/she intends to memorize. The search and ranking result is swift in our chatbot. This chatbot is a new development to our image recommendation reseach for foreign vocabulary acquisition (Hasnine et al., 2017, 2019).

Future work will address the efficacy of this chatbot. We aim to investigate how learners interact with this chatbot environment, whether the recommended images help learners in short- and long-term memory retention, and what image properties are essential to recommend images for abstract nouns or adjectives. We also aim to investigate where generative AI could be used in improving the performance of AIVAS-iBOT.

#### Acknowledgements

This project is supported by JSPS Grant-in-Aid for Early-Career Scientists grant no. 21K13651.

#### References

- Hasnine, M. N., Flanagan, B., Ishikawa, M., Ogata, H., Mouri, K., & Kaneko, K. (2019). A platform for image recommendation in foreign word learning. *Companion Proceedings of the 9th International Conference on Learning Analytics and Knowledge (LAK'19)*, 187–188.
- Hasnine, M. N., Ishikawa, M., Hirai, Y., Miyakoda, H., & Kaneko, K. (2017). An algorithm to evaluate appropriateness of still images for learning concrete nouns of a new foreign language. *IEICE Transactions on Information and Systems*, 100(9), 2156–2164. https://doi.org/10.1587/transinf.2016edp7463
- Hasnine, M. N., Mouri, K., Flanagan, B., Akcapinar, G., Uosaki, N., & Ogata, H. (2018). Image recommendation for informal vocabulary learning in a context-aware learning environment. *Proceedings of the 26th International Conference on Computer in Education*, 669–674.
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1), 973–1018. https://doi.org/10.1007/s10639-022-11177-3
- Pérez-Soler, S., Juárez-Puerta, S., Guerra, E., & de Lara, J. (2021). Choosing a Chatbot Development Tool. *IEEE Software*, 38(4), 94–103. https://doi.org/10.1109/MS.2020.3030198
- Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, 151, 103862. https://doi.org/10.1016/j.compedu.2020.103862

# W12

## The 1st International Workshop on Computational Thinking and Programming Education (CTPE)

## The Impact of an Educational Board Game on Students' Learning Achievement and Flow State in Media and Information Literacy

#### Ya-Xuan ZHANG, Ting-Chia HSU\*

Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taiwan. \*ckhsu@ntnu.edu.tw

**Abstract:** Given the increasing international emphasis on the importance of identifying and critiquing media content and the proper use of information and communication software, it is imperative to enhance students' media and information literacy education. In this study, an educational board game for media and information literacy was developed to cultivate the basic knowledge of media and information literacy among university school students in Taiwan. A quasi-experimental study was conducted to investigate the effects of the educational board game on students' learning achievement and flow state. Results showed that the educational board game increased students' learning achievement and flow state, but no difference was found between the collaborative-team gaming group and the competitive-multiplayer gaming group in terms of learning achievement and flow state.

**Keywords:** Media and Information Literacy, Educational Board Game, Flow State, Learning Achievement

#### 1. Introduction

With the development of information technology, people no longer limit their sources of knowledge to traditional media such as newspapers and magazines or television news. The younger generation is inclined to access all kinds of information online using mobile devices. However, the vast amount of information available on social media is full of inaccuracies, some of which were created with the intention of falsifying real information. These articles often have the plausible presentation of real figures and information, which is known as "fake news" (Jones-Jang et al., 2019). Fake news spreads rapidly through the Internet, causing negative impacts on various sectors such as politics, business, and even philanthropy (Vafeiadis et al., 2019). For instance, intentional dissemination of fake news during a specific period led to the occurrence of the Brexit referendum in the UK, or decreased people's willingness to wear masks during the COVID-19 pandemic (Isaak & Hanna, 2018; Kouzy et al., 2020). It is evident that the public's behavior can easily be misled by misinformation. Children in the digital age are surrounded by a plethora of trustworthy and untrustworthy information on a daily basis. This places a great deal of stress on the development of their emotional and cognitive abilities, which may increase their anxiety, damage their self-esteem or distort their worldview (Howard et al., 2021). Although these young people believe that they are confident in the use of these tools in an era when they are dealing with digital products, the fact is that they are watching violent pornography and distorted news messages while using mobile devices on their own (Ofcom, 2020). Meanwhile, problems like cyber addiction and cyber bullying are growing every year (Ofcom, 2020). Therefore, it is vital that alternative media and information literacy education material be further developed (Alexander & Galina, 2020).

The term "gamification" refers to the application of game mechanics to non-game contexts. This new approach to teaching and learning has been widely accepted in recent years in education, and has been integrated into business training, as it can lead to greater motivation and engagement of participants (Caponetto et al., 2014). Gamification is now widely used in various kinds of courses, such as traditional physical classrooms, MOOCs, or other online courses, as well as mixed courses and platforms where gamified content can be applied (Dicheva et al., 2015). Gamification manifests itself in the form of, but is not limited to digital games, virtual classes, and competitive activities. Compared to simply incorporating game mechanics into the curriculum, game-based learning integrates instructional goals into a game to support teaching or student learning. It is another form of applying games in educational settings (Cardinot & Fairfield, 2022). Board games have a long history, and the integration of educational content into board games is not a recent development. Subjects such as mathematics, social sciences, language, and more have educational board game products available. Particularly in academic contexts, educational board games are a recent hit. BoardGameGeek, the world's leading board game website, has classified educational board games as a separate type of board game (BoardGeekGame, 2023). As cognitive skills are valued in the 21st century, game-based learning that incorporates elements of gameplay into the classroom provides learners with a space to explore, be creative, explore topics of interest in groups, and promote opportunities for inquiry and questioning (Pinedo et al., 2022). Learning tasks in game-based learning can reduce the complexity of cognitive learning, clarify thinking strategies, and direct learners' attention to the learning outcomes, thereby giving learners a greater opportunity to further develop and control the problemsolving learning process and thus achieve their learning goals. Learners can have a positive impact on their emotions and persistence in learning through teacher guidance, encouragement, questioning of learners and sharing of knowledge and experience while reducing the cognitive load of learning (Cai et al., 2022).

The study aimed to investigate the effects of using board games in a media and information literacy course on students' learning achievement and flow state, and provided suggestions for teachers who plan to use educational board games for media and information literacy education in the future.

The research questions are as follows:

(1). Does the expanded contextual education board game have any effect on learners' learning achievement of media and information literacy?

(2). Does the educational board game have any effect on the flow of learners?

#### 2. Literature Review

#### 2.1 Media and information literacy

As technology advances, diverse media from computer devices and mobile phones are taking over the traditional press's dominant position in society. When individuals watch television, play games, or browse the web, they may be potentially interfered with by government authorities, policies or rules. These include possible issues such as cultural diversity, visual media presentation or aesthetic design, or the complexity of media presentations or text composition. Nowadays, multimedia presentation highly contrasts with the traditional linear media of the past (Livingstone, 2004). Students' media literacy in the early years is not sufficient to cope with the large volume of digital information. Therefore, other literacies, such as information literacy, digital literacy and information and computer literacy, have been proposed to help people develop robust media literacy skills.

Although there is a distinction between overlapping and independent literacy in media literacy, many scholars have suggested that an integration of the two can attempt to address the issues which are considered in the field of media and information literacy education today (Leaning, 2019; Mackey & Jacobson, 2011). As new social media and online interaction models have emerged, independent literacy skills have failed to address the challenges. In 2013, the United Nations Educational, Scientific and Cultural Organization

(UNESCO) pointed out that the ability to distinguish evidenced data and information from various information and dissemination platforms is a competency that should be possessed by digital natives in the 21st century(UNESCO, 2013).

Based on the scope of information and media literacy defined by UNESCO, critical thinking can effectively enhance students' ability to identify the accuracy of information (Al-Zou'bi, 2021). The improvement of information and media literacy can also benefit the general public in other domains. Currently, in addition to traditional lecture-based instruction, the education sector has adopted various technological tools to teach information and media literacy. For example, in the field of economics education, the integration of information and media literacy can serve as a fundamental component for students to understand the digital economy, helping them discern which economic information is usable and trustworthy, thereby avoiding financial risks and developing sound economic literacy (Mukhtar & Putri, 2021).On the other hand, games are also effective tools for assisting children in developing critical thinking about media. They promote student collaboration and discussions with peers in the classroom, aiding in the development of literacy. Games can also integrate and reflect their daily media experiences (Costa et al., 2018).

#### 2.2 Game-based learning

The use of educational board games is more limited than that of general gameplay. The main reason is that educational board games have a specific learning objective and a specific appeal, especially for primary school students, university students and even adults (Liu et al., 2020). Their design places emphasis on balancing the learning content with playfulness. It also examines whether the knowledge retained by learners in the gameplay can be appropriately applied to real-world situations. An effective game-based learning environment should support learners in their choice of actions, in moving towards their defined goals, and in experiencing the learning outcomes of their chosen actions. Even if a mistake is made by choosing the wrong option, the risk-free setting of the game environment provides learners with the challenge of failing, inspiring them to revise their ideas and re-plan their actions to succeed in the gameplay after they figure out the correct approach (Tay et al., 2022).

Compared to traditional lecture-based lessons, game-based learning can lead to better learning outcomes and higher motivation. It is also more engaging for students to improve their problem-solving skills (Al-Azawi et al., 2016). The benefits of game-based learning are not limited to the results of tests (Sailer & Homner, 2020). Learning outcomes are generally categorized as cognitive outcomes, motivation and learning behaviors (Sailer & Homner, 2020). Cognitive knowledge involves conceptual knowledge of the intended learning objectives, including facts, principles, as well as application-oriented procedural, strategic and contextual knowledge. The result of learning motivation includes the learner's intrinsic motivation, dispositions, preferences, attitudes, confidence and self-efficacy. Learning outcomes refer to technical skills and motor skills (Sailer & Homner, 2020). Gamebased learning has a significant impact on improving cognitive outcomes, and it is better able to help students improve their learning performance than traditional methods (Chen et al., 2020; Ou et al., 2021). Furthermore, another advantage of game-based learning is that it is not limited to classroom settings. It continues to have a positive impact on learners in informal educational environments as well. By setting some basic expectations, game-based learning can yield unexpected outcomes (Alam, 2022). In the 21st century, which emphasizes cognitive skills, game-based learning provides learners with opportunities to explore, unleash their creativity, engage in group inquiries on topics of interest, and foster learning spaces that promote discussion and questioning (Pinedo et al., 2022).
#### 2.3 Flow state

A flow state is a state in which an individual focuses on an activity without noticing external things (Csikszentmihalvi, 1991). Therefore, the flow state can be influenced when learners participate in, attend, or exert mental effort during an activity (Kiili et al., 2021). The design of an educational board game must take into account the elements of participants, tasks and objects. This means that it includes clear objectives, immediate feedback, potential control over fun, ease of use and a sense of control over the activity. Flow consists of three stages, namely flow antecedents, flow experience and flow consequences. When educational games incorporate the above elements, they make it easier for learners to enter a state of mind-flow (Kiili, 2005). Flow has also been found to be strongly correlated with learning achievement. When learners have a more focused flow state, they have better learning results. On the other hand, the use of competitive elements in the game also makes a difference in the flow of learners' minds, and group game designs are more immersive than individual game designs. The relationship between flow and learning achievement is also influenced by the mediating effect of students' learning satisfaction (Chan et al., 2021; Kim & Park, 2021; Rvu et al., 2022). Moreover, flow has been found to be strongly correlated with learning achievement. When learners experience a more focused state of flow, their learning outcomes tend to be superior. The presence or absence of competitive elements in games also influences the occurrence of flow among learners. Group game designs are more conducive to immersion in activities, and the relationship between flow and learning achievement is influenced by the mediating effect of learner satisfaction (Chan et al., 2021; Kim & Park, 2021; Ryu et al., 2022). Currently, there are various scales available for measuring flow. Examples include the Flow State Scale (FSS), the Dispositional Flow Scale (DFS-2) (Jackson & Eklund, 2002; Jackson & Marsh, 1996), as well as scales designed specifically for certain activities, such as Kiili's (2006) Flow Scale for Games to understand flow experiences during gameplay and Pearce et al.'s (2005) Flow State Scale. In this study, considering the pre-test and post-test design as well as the experimental timeframe, we chose to use an 8-item flow state scale.

#### 3. Method

Hi-MyLife, an educational board game for media and information literacy, was developed as the teaching material for this study. The board game offers two distinct gameplay modes: Faction Battle and Multiplayer Competition. In the Faction Battle mode, players team up in pairs, working collaboratively to tackle card-based challenges. Conversely, the Multiplayer Competition mode requires all four participants to independently brainstorm solutions to these challenges. The core gameplay involves drawing cards that depict various story scenarios. Players then choose answer cards from their hand to decipher the challenges and accumulate points. Following each round, learners revisit any story scenario cards they initially answered incorrectly, getting another chance to tackle the challenges. Ultimately, the player with the highest score emerges as the victor of the game. The participants were first-year students taking a digital learning theory course at a national university in northern Taiwan. A total of 28 participants took part in the study. The participants were randomly assigned by the researcher to groups of four, with 16 participants in a collaborative-team (CT) game group, while 12 participants were in a competitive-multiplayer (CM) game group.

The experimental process is shown in Figure 1. After a brief introduction and explanation of the rules of the board game, the pre-test, which lasted 15 minutes, was conducted. The pre-test consisted of a scale of flow state and a learning achievement test of media and information literacy. The scale was developed by Pearce et al. (2005) using a 5-point Likert scale with eight questions on enjoyment, engagement and control, with options of strongly agree, agree, neutral, disagree and strongly disagree. The learning achievement test on media and information literacy was developed by the research team based on the current IT textbooks for primary and secondary schools, and was reviewed by two experts. A group is assigned to a set of board games with the same game accessories. However, the

game rules vary depending on the group. The collaborative-team gaming group had four players divided into two teams to compete against each other, whereas the competitivemultiplayer (CM) gaming group was designed for four players to compete against each other. The session lasted 90 minutes. After the intervention, a post-test was administered to the participants, which consisted of an 8-item Flow State Scale and a 20-item Media and Information Literacy Learning Achievement Test. The items in the Media and Information Literacy Learning Achievement Test were different from those in the pre-test. The post-test session lasted for 20 minutes.



Figure 1. The experimental process.

#### 4. Results

#### 4.1 Learning achievement of media and information literacy

Table 1 shows the homogeneity of the CT gaming group and the CM gaming group. The results conformed to the assumption of homogeneity (F = 0.476, p > .05); however, the test of normal distribution was not found. Therefore, the Wilcoxon Signed Rank Test was used to examine whether there was a significant difference in the learning achievement of the two groups after participating in the educational board game.

The overall pretest of learning achievement (M = 29.64, SD =11.86) was significantly different (M = 37.32, SD = 11.16) from the overall posttest of learning achievement (Z = -2.27, p < .01) in media and information literacy learning.

Table 1. Levene's test results for learning achievement in media and information literacy

F	df1	df2	р
0.476	1	26	0.496

tests	Ν	М	SD	Mean Rank	Sum of Ranks	Z
Pretest	28	29.64	11.86	9.61	86.50	2 27**
Posttest	28	37.32	11.16	15.56	264.50	-2.21
**p < .01						

Table 2. The Wilcoxon Signed Rank Test of all participants

To investigate whether there were any significant differences in the learning achievement of the two groups after using different modes of the educational board game, the Mann-Whitney U test was used. Table 3 shows that there was no significant difference (Z = -0.917, p > .05) in the posttest of the learning achievement of the CT gaming group and the CM gaming group.

<b>T</b> I I A	1 4 (")						,	
I ahle 3	Wilcoxon	signed_rank	( test tor	learnind	achievement	usina hoar	d dame	arounina
Tuble 0.	VVIICOXOII	Signed runn	1001101	icuning	acriteventerit	using bour	a guine	grouping

board game grouping	Ν	М	SD	Mean Rank	Sum of Ranks	U	W	Ζ
CT gaming groups	16	39.38	11.96	15.72	251.50	76 50	154 50	_0 917
CM gaming groups	12	34.58	14.69	12.88	154.50	10.50	104.00	-0.517

#### 4.2 Flow state for media and information literacy learning

Table 4 presents the homogeneity of the CT gaming group and the CM gaming group according to the Levene's test, which meets the homogeneity assumption (F = 1.134, p > .05) and normal distribution for the data. Therefore, paired samples t tests were performed to examine whether there was a significant difference in the flow state of learning of the two groups after playing the educational board game. The CT gaming group showed a significant difference (t = -2.23\*, p < .05) between the pretest (M = 3.41, SD = .684) and the posttest (M = 3.64, SD = .714) for flow state. The CM gaming group showed a significant difference (t = -2.32\*, p < .05) between the pretest (M = 3.36, SD = 0.64) and the posttest (M = 3.79, SD = 0.60) for flow state. There was also a significant difference (t = -3.19\*\*, p < .01) between the pretest (M = 3.71, SD = 0.66) for the overall flow state.

F	df1	df2	p
1.134	1	26	0.297

Table 4. Levene's test results for flow state when accessing media and information literacy

Table 5. <i>Learnin</i>	g achievement fo	or media and	information I	literacy usir	ng paired sar	nples t tests
	J					

Croups	Ν.	Mear	n(SD)	+	2
Groups	<i>IN</i> —	Pre-test	Post-test	l	ρ
CT gaming groups	16	3.408(.684)	3.643(.714)	-2.23 <sup>*</sup>	.042
CM gaming groups	12	3.357(.636)	3.794(.599)	<b>-</b> 2.32 <sup>*</sup>	.041
The overall result	28	3.386(.650)	3.708(.660)	-3.19**	.004

\**p* < .05, \*\**p* < .01

To investigate whether there were differences in the flow state of students in the two groups after playing the educational board game, analysis was conducted using an independent sample t test. Table 6 shows that no significant difference was found in the flow state of the CT gaming group and the CM gaming group (t = -0.59, p > .05).

Table 6. Flow state in media and information literacy using independent sample t tests

Groups	Ν	Mean	SD	t	р
CT gaming groups	16	3.64	0.71	-0 59	559
CM gaming groups	12	3.36	0.60	0.00	.000

To further understand the differences in the qualitative results, the qualitative analysis presents the results based on the students' feedback. The categorized data are shown in Table 7. The data can be categorized into three main categories, namely gaming rules, learning materials and game components.

Table 7. Student fee	dback
----------------------	-------

Dimensions	Feedback
Game rules	The game rules are too difficult and it is not clear how to play the game.
	A bit more diversity can be added to the characters.
Learning	Some questions are vague, while some narrative answers are
materials	ambiguous. Keywords for each question can be annotated.
Game	Function cards are too many, while props are also abundant. Simplifying
components	them would be better.

#### 5. Discussion and Conclusions

The purpose of this study was to investigate the difference in the learning achievement and flow state of students in the CT and CM gaming groups when playing an educational board game about media and information literacy. The research findings revealed that the educational board game designed for promoting media and information literacy had a significant impact on overall student learning achievement. This discovery aligns with the findings of Li and Hou (2022), underscoring the effectiveness of utilizing board games as educational tools in enhancing student academic performance. Compared to traditional lecture-based teaching, board games offer a more motivating and engaging learning experience. Their game mechanics, coupled with the ability to seamlessly integrate visuals and text, contribute to the establishment of effective learning pathways. The factors of rewards, tokens, and competition in gameplay increase the enjoyment of learning. In addition, this study also found that the educational board game improved students' flow experience, which is in line with Lin and Wang's (2022) finding that the integration of board game materials into academic learning activities can enhance students' flow states.

However, this study also found that collaborative-team gaming did not affect students' perceptions of flow state compared to competitive-multiplayer gaming, in contrast to Hou and Keng (2021) who found that collaborative-team gaming resulted in greater immersion in a state of flow. Our result demonstrated that the possibility of having a more knowledgeable learner who led or dominated the game play, whereas a lack of discussion of the content of the materials during gameplay prevented students from engaging in the game. In addition, no differences were found in the learning achievements of the collaborative-team gaming group and the competitive-multiplayer gaming group. The results differ from those of Azhar et al. (2022) who found that collaborative-team gaming learning was associated with better learning achievement than competitive-based learning. This may be due to the fact that the grouping in this study did not take into account the variables of students' backgrounds such as gender, personality traits, and cognitive style, which may have prevented the collaboration from being effective (Liao et al., 2019).

According to the findings of this study, it is recommended that teachers who wish to use educational board games in the future should consider the differences in the selection of teaching materials according to the students' personalities. Regarding the teaching of media and information literacy courses, teachers may consider incorporating elements of board games into the course design. This could be used as an alternative approach for certain learning topics or as a replacement for traditional lecture-based classroom assessments. Those who develop educational board games in the future can conduct tests with similar target groups in stages and collect possible feedback so that the content and game design can be further revised according to their needs. For future educators or researchers who wish to explore the content of educational board games, the variables of students' backgrounds or other technological media, such as augmented reality, can be incorporated into the study for further discussion.

#### Acknowledgements

This study is partially supported by the National Science and Technology Council under the contract number NSTC 111-2410-H-003 -168 -MY3.

#### References

- Al-Azawi, R., Al-Faliti, F., & Al-Blushi, M. (2016). Educational gamification vs. game based learning: Comparative study. *International journal of innovation, management and technology*, 7(4), 132-136.
- Al-Zou'bi, R. (2021). The impact of media and information literacy on acquiring the critical thinking skill by the educational faculty's students. *Thinking Skills and Creativity*, 39, 100782.
- Alam, A. (2022). A digital game based learning approach for effective curriculum transaction for teaching-learning of artificial intelligence and machine learning. *2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, 69-74.
- Alexander, F., & Galina, M. (2020). Current trends in media and information literacy in research and scientific publications of the early 21st century. *International Journal of Media and Information Literacy*, 5(2), 153-163.
- Azhar, S. A. F. J., & Jalil, H. A. (2022). Comparison of Individual and Collaborative Game-Based Learning Using Tablet In Improving Students' Knowledge In Primary Classroom Environment. *Asian Journal of University Education*, *18*(1), 205-216. <u>https://doi.org/10.24191/ajue.v18i1.17188</u>
   BoardGeekGame. (2023). *Board game categories*.
- https://boardgamegeek.com/browse/boardgamecategory
- Caponetto, I., Earp, J., & Ott, M. (2014). Gamification and education: A literature review. Proceedings of the European Conference on Games-based Learning, Acad Conferences Ltd.
- Cardinot, A., & Fairfield, J. A. (2019). Game-based learning to engage students with physics and astronomy using a board game [Article]. *International Journal of Game-Based Learning*, *9*(1), 42-57. <u>https://doi.org/10.4018/IJGBL.2019010104</u>
- Chan, K., Wan, K., & King, V. (2021). Performance over enjoyment? Effect of game-based learning on learning outcome and flow experience. *Frontiers in Education*, *6*, 185.
- Chen, C.-Y., Huang, H.-J., Lien, C.-J., & Lu, Y.-L. (2020). Effects of multi-genre digital game-based instruction on students' conceptual understanding, argumentation skills, and learning experiences. *IEEE Access*, *8*, 110643-110655.
- Costa, C., Tyner, K., Henriques, S., & Sousa, C. (2018). Game creation in youth media and information literacy education. *International Journal of Game-Based Learning (IJGBL)*, 8(2), 1-13.
- Csikszentmihalyi, M. (1991). *Flow: The Psychology of Optimal Experience*. Harper Perennial Modern Classics.
- Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Journal of Educational Technology & Society*, *18*(3), 75-88.
- Hou, H.-T., & Keng, S.-H. (2021). A dual-scaffolding framework integrating peer-scaffolding and cognitive-scaffolding for an augmented reality-based educational board game: An analysis of learners' collective flow state and collaborative learning behavioral patterns. *Journal of Educational Computing Research*, 59(3), 547-573.
- Isaak, J., & Hanna, M. J. (2018). User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer*, *51*(8), 56-59.
- Jackson, S. A., & Eklund, R. C. (2002). Assessing flow in physical activity: The flow state scale–2 and dispositional flow scale–2. *Journal of sport and exercise psychology*, 24(2), 133-150.

- Jackson, S. A., & Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The Flow State Scale. *Journal of sport and exercise psychology*, *18*(1), 17-35.
- Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *The Internet and higher education*, *8*(1), 13-24.
- Kiili, K. (2006). Evaluations of an experiential gaming model. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, 2(2), 187-201.
- Kiili, K. J. M., Lindstedt, A., Koskinen, A., Halme, H., Ninaus, M., & McMullen, J. (2021). Flow experience and situational interest in game-based learning: Cousins or identical twins. *International Journal of Serious Games*, 8(3), 93-114.
- Kim, S.-H., & Park, S. (2021). Influence of learning flow and distance e-learning satisfaction on learning outcomes and the moderated mediation effect of social-evaluative anxiety in nursing college students during the COVID-19 pandemic: A cross-sectional study. *Nurse Education in Practice*, *56*, 103197.
- Leaning, M. (2019). An approach to digital literacy through the integration of media and information literacy. *Media and Communication*, 7(2), 4-13.
- Liao, C.-W., Chen, C.-H., & Shih, S.-J. (2019). The interactivity of video and collaboration for learning achievement, intrinsic motivation, cognitive load, and behavior patterns in a digital game-based learning environment. *Computers & Education*, *133*, 43-55.
- Li, C. T., Hou, H. T., & Lee, L. H. (2022). Design of a dual-hierarchy scaffolding board game-based learning activity for EFL reading comprehension. *Language Teaching Research*, https://doi.org/10.1177/1362168822112590.
- Lin, Y.-T., & Wang, T.-C. (2022). A Study of Primary Students' Technology Acceptance and Flow State When Using a Technology-Enhanced Board Game in Mathematics Education. *Education Sciences*, *12*(11), 764.
- Liu, Z.-Y., Shaikh, Z., & Gazizova, F. (2020). Using the concept of game-based learning in education. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(14), 53-64.
- Livingstone, S. (2004). Media literacy and the challenge of new information and communication technologies. *The communication review*, 7(1), 3-14.
- Mackey, T. P., & Jacobson, T. E. (2011). Reframing information literacy as a metaliteracy. *College & research libraries*, 72(1), 62-78.
- Mukhtar, S., & Putri, K. (2021). Technology integrated on media literacy in economic studies on higher education. *Journal of Social Studies Education Research*, *12*(1), 95-123.
- Ofcom. (2020). *Children and Parents: Media Use and Attitudes Report.* https://www.ofcom.org.uk/research-and-data/media-literacy-research/childrens/children-and-parents-media-use-and-attitudes-report-2022
- Ou, K.-L., Liu, Y.-H., & Tarng, W. (2021). Development of a virtual ecological environment for learning the Taipei tree frog. *Sustainability*, *13*(11), 5911.
- Pearce, J. M., Ainley, M., & Howard, S. (2005). The ebb and flow of online learning. *Computers in human behavior*, 21(5), 745-771.
- Pinedo, R., García-Martín, N., Rascón, D., Caballero-San José, C., & Cañas, M. (2022). Reasoning and learning with board game-based learning: A case study. *Current Psychology*, *41*(3), 1603-1617.
- Ryu, E. J., Jang, K. S., & Kim, E. A. (2022). Influence of Learning Presence of Non-Face-to-Face Class Experience in Nursing Students on Academic Achievement: Mediating Effect of Learning Flow and Moderated Mediation of Digital Literacy. *Journal of Korean Academy of Nursing*, 52(3), 278-290. <u>https://doi.org/10.4040/jkan.21241</u>
- Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review*, 32(1), 77-112.
- Tay, J., Goh, Y. M., Safiena, S., & Bound, H. (2022). Designing digital game-based learning for professional upskilling: A systematic literature review. *Computers & Education*, 104518.
- UNESCO. (2013). Global Media and Information Literacy Assessment Framework: country readiness and competencies. https://unesdoc.unesco.org/ark:/48223/pf0000224655.page=22
- Vafeiadis, M., Bortree, D. S., Buckley, C., Diddi, P., & Xiao, A. (2019). Refuting fake news on social media: nonprofits, crisis response strategies and issue involvement. *Journal of Product & Brand Management*, 29(2), 209-222. <u>https://doi.org/10.1108/jpbm-12-2018-2146</u>

## Game-Based Learning of Al Image Recognition on Computational Thinking and Self-Efficacy of Undergraduates

Hui-Wen HSUEH, Ting-Chia HSU\*

Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taiwan, ROC \* ckbsu@ptpu.edu.tw

\* ckhsu@ntnu.edu.tw

Abstract: With the rise of artificial intelligence (AI), cultivating the AI abilities of the new generation of learners has become crucial. However, currently, there is a lack of tools in education found to effectively enhance students' AI abilities. To discover effective ways to foster students' understanding of AI concepts and skills, this research developed an Artificial Intelligence Image Recognition Robot unit, which was conducted through a series of AI courses in conjunction with educational robots. There are total of 41 undergraduates participating the instructional experiment. The first half of the experiment focused on introducing AI concepts and implementing image recognition models, while the second half involved robot programming and a robot board game competition. At the beginning and the end of the nit, students completed the AI Learning Achievement Test, the Programming-Oriented Computational Thinking scale and the Robot Self-Efficacy Scale to assess differences before and after the unit. The research found that by taking the AI Image Recognition Robot unit, students' AI learning achievements, programming-oriented computational thinking, and robot self-efficacy were significantly enhanced. This suggests that students will be better equipped to evaluate and understand the effectiveness of algorithms in programming and problem solving in the future. It is recommended that robots be incorporated as educational tools when engaging in Al learning.

**Keywords:** Artificial Intelligence, Educational Robot, self-efficacy, Computational Thinking

#### 1. Introduction

With the rise of artificial intelligence (AI), it is being widely used in various fields and industries, bringing great changes to people's lives. With its rapid development, the demand for personnel with AI skills is gradually increasing, and it is crucial to cultivate the ability of AI for the new generation of learners (Su et al., 2022). However, how to effectively train students to learn AI concepts and technologies is lack of empirical study evidences. Relevant studies have mentioned that learning the CT concept and implementation from learning by doing of educational robots plays a key role in AI education, and allowing students to participate in the development and production of robots can effectively promote their interest and understanding in the field of AI (Dodds et al., 2006). Therefore, this study used educational robots with the hope of improving students' learning achievements in learning AI.

Self-efficacy is a theoretically and empirically robust motivational belief, and people with high self-efficacy put in more effort and persist for longer than those with low self-efficacy (Klassen & Klassen, 2018). This study used the Robot Self-Efficacy Scale to test whether this unit can improve students' robot self-efficacy and enable them to explore robot learning more actively, so as to help them obtain deeper help and inspiration in the field of AI in the future. Computational thinking, a way of solving problems and understanding computational processes, plays an important role in the modern digital generation (Witherspoon et al., 2017). There is a close relationship between computational thinking and

programming. Writing programs helps to improve students' computational thinking ability, so that they can solve real-life problems using this ability (Selby, 2015). This unit used the Programming-Oriented Computing Thinking scale to test whether the writing of robot programs can improve the performance of students' computing thinking, which will help students to effectively solve problems in AI or everyday life in the future.

Both aspects play an important role in learning to apply AI technology. Therefore, through robot self-efficacy and programming-oriented computational thinking, we wanted to detect the influence of students' mental state when learning about AI, and thus we proposed the following two research questions:

- Does the artificial intelligence image recognition robot unit affect students' programmingoriented computing thinking?
- Does the artificial intelligence image recognition robot unit affect students' robot selfefficacy?

#### 2. Literature review

#### 2.1 Educational robots

In Papert's constructivism, it is mentioned that if learners can engage in hands-on learning themselves, it will be more effective for understanding knowledge (Papert, 1980). Integrating educational robots into the classroom allows students to have complete control over the entire learning process, engaging them in hands-on activities that enhance their active participation and focus. This approach makes the teaching content more meaningful and interesting (Chaidi et al., 2021). Additionally, involving students in the development and creation of robots can effectively stimulate their interest and understanding in the field of artificial intelligence (Dodds et al., 2006). It not only improves learners' motivation to learn and their ability to collaborate in teams, but also enables them to learn from mistakes. Various students (Castro et al., 2018). Through constructing and programming robots, students' computational thinking is significantly enhanced, making educational robots a highly effective teaching tool for promoting computational thinking development (Chaidi et al., 2021; loannou & Makridou, 2018; Qu & Fok, 2021).

#### 2.2 Computational Thinking

Computational thinking is an important concept that is of great interest to the international education community, and cultivating students' computational thinking has become an important task to promote the development of IT talents (Yang et al., 2020). Wing (2006) mentioned that computational thinking is the ability to solve problems, construct systems and understand human behavior based on computer science. Computational thinking emphasizes thinking at different levels and in different fields, and can cultivate learners' critical thinking skills so that they can be applied to solve other daily life problems. Computational thinking has also been identified as an indispensable and important skill in learning AI knowledge (Kim et al., 2021).

#### 2.3 Self-efficacy

Self-efficacy refers to an individual's estimation of or belief about his or her own ability in a specific situation or task. It is one of the factors that affect human activities. It directly or indirectly affects human behavior, such as personal motivation, thinking mode, and the amount and time of personal effort to complete an activity (Bandura, 2006). Self-efficacy is a theoretically and empirically robust motivational belief that has been shown to play an important role in learning and developing new skills (Klassen & Klassen, 2018). This study measures students' robot self-efficacy, which is used to evaluate learners' self-efficacy in

learning robotics, that is, learners' motivational beliefs when learning about robots (Tsai et al., 2021).

#### 3. Method

#### 3.1 The Artificial Intelligence Image Recognition Robot Unit

For this study, we designed an artificial intelligence image recognition robot unit. The topics include the concept of artificial intelligence and machine learning, the implementation of image recognition model training and robot programming, and so on. We used the Teachable Machine platform launched by Google for image recognition implementation, as it is easy to operate and does not require programming or any professional knowledge. (Figure 1). In this study, each student was assigned the task of utilizing the platform, teachable machine, to implement image recognition. The students were given specific assignments to hone their skills in training image recognition models comprehensively. Upon completion of recognition, learners could import the trained image recognition model into the robot program.



Figure 1. The Interface of the Teachable Machine Image Recognition Platform

The programming unit utilizes the BlocklyDuino platform, through which students learned how to write robot programs online. The platform allowed them to experiment with importing code into the robots and successfully operating the robots (as shown in Figure 2).

BIOCKIYDUINO Editor F2	🔯 Al Thinker ESP3	2-CAM - 4 COM4 -
* XML ARDUINO		V 🗈 🕂 X 📕 🛇 🗖 🎙
#頁模組 Linkt76973#1177461		程式碼 🗘
Link Provident Statistics	S 下載初體 📕	
EDD00464月1日日 ESP32	Writing at 0x00078000 (90 %)	
ESP32111号模组	Writing at 0x0007c000 (93 %)	
ESP32網頁藍穿模組	Writing at 0x00080000 (96 %)	
ESP32籃芽鍵盤模組	Writing at 0x00084000 (100 %)	
Telegram Bot模组	Wrote 829136 bytes (479772 compressed) at 0x00010000 in 11.1 seconds (effective 597.8 k	
ESP32-CAM编审程组	Hash of data verified.	
Pixel-Bitler William	Compressed 3072 bytes to 119	
1 100-0 100 00 00 00 00 00 00 00 00 00 00 00 0	Writing at 0x00008000 (100 %)	
1/2/11/9/F (BK	Wrote 30/2 bytes (119 compressed) at 0x00008000 in 0.0 seconds (effective 4096.1 kbit/s)	
全身辨識	Hash of data verified.	
手夠辨識	Invine	
機器學習(google)	Hard resettion via PTS on	
人臉偵測 (face-api)	naid reserving the Kra prise	
人脑辨識 (face-api)	Finish	
(+) · · · · · · · · · · · · · · · · · · ·	1	
.R/L♠	期間 下載状態 程式碼 下載	

Figure 2. The Interface of the BlocklyDuino Programming Platform

The post-competition used the "Shopping Robot Game" developed by Professor Ting-Chia HSU's team from National Taiwan Normal University. This board game combines the concepts of artificial intelligence and the Internet of Things. Players can control the robots in the game using their own trained image recognition models and programmed code. The game aims to apply the knowledge learned in the unit and enhance learners' motivation through gameplay. Following the principles of the Internet of Things, players roll the dice on the webpage to move the robots on the map (Figure 3) and collect materials. By utilizing image recognition, players add the collected materials to their shopping carts strategically, aiming to gather the required items from the shopping list in the board game.



Figure 3. The Map of the Shopping Robot Game

#### 3.2 Research Framework

In this study, the research framework was divided into independent and dependent variables. The independent variable was the artificial intelligence image recognition robot course, while the dependent variables were learning achievement, programming-oriented computational thinking, and robot self-efficacy, as shown in Figure 4.



*Figure 4.* Research Framework

#### 3.3 Experimental Design

In this study, the curriculum was divided into two parts, with the first part focusing on introducing the concepts of artificial intelligence, machine learning, and the implementation of image recognition. The second part emphasized robot programming and the implementation of a board game involving image recognition robots. Figure 5 illustrates the unit design.

	First half of the lesson	Second half of the lesson
Topic	<ol> <li>AI and Machine Learning Concepts</li> <li>Image Recognition Model Implementation and Testing</li> </ol>	<ol> <li>Robot Programming</li> <li>Al Image Recognition Shopping Robot Game</li> </ol>

Figure 5. Instructional Design

The first part of the course introduced the fundamentals of artificial intelligence and machine learning. Participants were assigned the tasks of implementing and testing image recognition models. In the second part, hands-on learning was emphasized. Learners were introduced to the concept of robot motor movement and provided with partially completed programming code in a block-based format. Participants were responsible for completing the missing sections in the code. After finishing the code, they used BlocklyDuino to write the robot board game program and subsequently tested the written program.

Towards the end of the curriculum, participants were grouped as necessary to engage in an educational board game titled "Shopping Robot Game." Following a brief introduction from the teacher, the game commenced, accommodating students in pairs or groups of four. They utilized the programs and models they had developed earlier to manipulate the robots within the board game. After the game concluded, a summary session was held to recap the experience.

The curriculum consisted of two parts: Al image recognition and robot programming, with each part lasting for 3 hours, for a total of 6 hours to complete this Al image recognition unit. After completing the learning unit, participants were asked to fill out the learning achievement test, the robot self-efficacy scale, and the programming-oriented computational thinking scale to assess the differences in participants' abilities before and after the unit. The experimental process is illustrated in Figure 6.



Figure 6. Experimental Process Flowchart

#### 3.4 Experimental Participants.

The participants in this study were first-year students from an engineering-related department at a national university in the northern region of Taiwan. The study adopted a quasi-experimental research design and was conducted during the Computer Introduction

course offered by the department. The two-week instructional content focused on artificial intelligence image recognition. The total number of participants in the study was 41.

#### 3.5 Experimental Tools.

#### 3.5.1 Artificial Intelligence and Computational Thinking Test

To investigate whether there was a significant difference in learning achievement after completing this unit, the Artificial Intelligence and Computational Thinking Test was used to assess learners' performance. This test includes both multiple-choice questions and matching questions. There are 14 multiple-choice questions and three matching questions, with a maximum score of 100 points.

#### 3.5.2 Programming-Oriented Computational Thinking Scale

In this study, students completed the Programming-Oriented Computational Thinking Scale to understand whether their programming-oriented computational thinking was influenced after completing the artificial intelligence image recognition robot unit. This scale was adapted from the scale developed by Kılıç et al. (2021) and consisted of three dimensions: Conceptual Knowledge, Algorithmic Thinking, and Evaluation. The original scale had a Cronbach's alpha reliability coefficient of .97. For this study, the dimensions of Conceptual Knowledge and Algorithmic Thinking were used for measurement, comprising a total of 19 questions. The Cronbach's alpha reliability coefficients for these two dimensions were .89 and .94, respectively. The scale utilizes a Likert-type 5-point scale for evaluation, with higher scores indicating better programming-oriented computational thinking abilities. The example of the items are shown as follows. I am clear about the learning concepts of robotics. I can link robotics concepts with other learning subjects. I can evaluate several robotic solutions for solving a problem.

#### 3.5.3 Robot Self-Efficacy Scale

In this study, students completed the Robot Self-Efficacy Scale to understand their level of robot self-efficacy after completing the artificial intelligence image recognition robot unit. The scale used in this study is based on the one developed by Tsai et al. (2021), which consists of five dimensions: Conceptual Understanding, Technical Skills, Application, Analysis, and Collaboration. The full scale contains a total of 16 items, and has a Cronbach's alpha reliability coefficient of .97. Each item is evaluated using a Likert-type 5-point scale, with higher scores indicating higher levels of robot self-efficacy. The example of the items are shown as follows. I can use the loop structures (for,while, etc.) appropriately. I can write a program that can do more than one operation at the same time. I can create simple algorithms without using unnecessary codes in the program.

#### 4. Research results

#### 4.1 Learning Achievement in Artificial Intelligence and Computational Thinking

As shown in Table 1, the paired-samples *t* test was used as the statistical method to examine the difference in students' learning achievements in artificial intelligence and computational thinking concepts before and after the course. The research results revealed a significant difference between students' pretest (M = 73.22, SD = 9.13) and posttest (M = 83.29, SD = 7.48) scores in artificial intelligence and computational thinking learning achievements ( $t = -7.11^{***}$ , p < .001). This indicates a noticeable improvement in students' learning achievements after taking the artificial intelligence image recognition robot course.

Table 1. Paired Samples t Test for Learning Achievement in Artificial Intelligence and Computational Thinking Concepts

Mean	+	NI	
Pretest	Posttest	L	IN
73.22(9.13)	83.29(7.48)	-7.11***	41
***p < .001			

#### 4.2 Programming-Oriented Computational Thinking

As shown in Table 2, paired-samples *t* tests were conducted to examine the difference in students' programming-oriented computational thinking before and after the course. The results revealed significant improvements in both dimensions, namely Conceptual Knowledge ( $t = -5.44^{***}$ , p < .001) and Algorithmic Thinking ( $t = -6.89^{***}$ , p < .001). This indicates that students' programming-oriented computational thinking showed a noticeable improvement after completing the artificial intelligence image recognition robot unit.

Table 2. Paired Samples t Test for	r Programming-Oriented	Computational	Thinking
------------------------------------	------------------------	---------------	----------

Factoro	N	Mear	Mean(SD)		
Faciors	IN	Pre-test	Post-test	l	
Conceptual	41	2.93	3.46	-5.44***	
knowledge		(0.81)	(0.58)		
Algorithmic	41	2.54	3.22	-6.89***	
thinking		(0.82)	(0.67)		
***p < .001		· · ·	• •		

#### 4.3 Robot Self-Efficacy Scale

As shown in Table 3, the paired samples *t*-test was used to detect the difference in the robot self-efficacy of the students before and after the unit. The results showed that practice, cooperation, comprehension and analysis all reached significant levels: practice ( $t = -3.52^{***}$ , p < .001), collaboration ( $t = -2.55^*$ , p < .05), comprehension ( $t = -3.61^{***}$ , p < .001), and analysis ( $t = -3.56^{***}$ , p < .001). It means that after the students took the artificial intelligence image recognition robot unit, the practice, cooperation, understanding and analysis of the robot self-efficacy of the students showed a clear trend of improvement.

Factors	N	Mear	4	
Factors	<i>I</i> V —	Pre-test	Post-test	L
Practice	41	2.73	3.16	-3.52***
		(0.92)	(0.76)	
Collaboration	41	2.93	3.30	-2.55*
		(0.94)	(0.78)	
Comprehension	41	2.75	3.23	-3.61***
		(0.86)	(0.76)	
Analysis	41	2.81	3.36	-3.56***
-		(0.82)	(0.74)	

#### Table 3. Paired Samples t Test for Robot Self-Efficacy

\*p < .05, \*\*\*p < .001

#### 5. Discussion and conclusions

The research findings indicate that after the two-week artificial intelligence image recognition robot unit, students' artificial intelligence learning achievement, programming-oriented computational thinking, and robot self-efficacy were significantly improved. Their performance after the unit was superior to their performance before the unit. The results demonstrate that integrating educational robots into the artificial intelligence image recognition classroom effectively enhances students' programming-oriented computational thinking. This suggests that students who learned artificial intelligence through educational robots were more effective in evaluating and understanding the effectiveness of algorithms when programming and problem solving.

Furthermore, concerning robot self-efficacy, the unit indeed boosted students' belief in their own abilities when working with robots. This means that students who participated in this unit had higher levels of self-efficacy in learning about robots, leading them to set more challenging goals for themselves and to strive to achieve them. Therefore, if educational robots can effectively enhance students' self-efficacy in learning, it will greatly benefit their personal motivation, thought processes, and effort levels when acquiring artificial intelligence skills.

Therefore, if educational robots can effectively enhance students' self-efficacy in learning, it will greatly benefit their personal motivation, thought processes, and effort levels when acquiring artificial intelligence skills.

#### Acknowledgements

This study is partially supported by the National Science and Technology Council under contract number NSTC 111-2410-H-003 -168 -MY3. We would like to express our gratitude to the research assistant, Mr. Mu-Sheng CHEN.

#### References

- Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-efficacy beliefs of adolescents, 5*(1), 307-337.
- Castro, E., Cecchi, F., Salvini, P., Valente, M., Buselli, E., Menichetti, L., Calvani, A., & Dario, P. (2018). Design and impact of a teacher training course, and attitude change concerning educational robotics. *International Journal of Social Robotics, 10*, 669-685. https://doi.org/10.1007/s12369-018-0475-6
- Chaidi, E., Kefalis, C., Papagerasimou, Y., & Drigas, A. (2021). Educational robotics in Primary Education. A case in Greece. *Research, Society and Development, 10*(9), e17110916371-e17110916371. https://doi.org/10.33448/rsd-v10i9.16371
- Dodds, Z., Greenwald, L., Howard, A., Tejada, S., & Weinberg, J. (2006). Components, curriculum, and community: Robots and robotics in undergraduate ai education. *Ai Magazine, 27*(1), 11-11. https://doi.org/10.1609/aimag.v27i1.1860
- Ioannou, A., & Makridou, E. (2018). Exploring the potentials of educational robotics in the development of computational thinking: A summary of current research and practical proposal for future work. *Education and Information Technologies*, 23, 2531-2544. https://doi.org/10.1007/s10639-018-9729-z
- Kılıç, S., Gökoğlu, S., & Öztürk, M. (2021). A valid and reliable scale for developing programmingoriented computational thinking. *Journal of Educational Computing Research*, 59(2), 257-286. https://doi.org/10.1177/0735633120964402
- Kim, S., Jang, Y., Kim, W., Choi, S., Jung, H., Kim, S., & Kim, H. (2021). Why and what to teach: Al curriculum for elementary school. *Proceedings of the AAAI Conference on Artificial Intelligence*. https://doi.org/10.1609/aaai.v35i17.17833
- Klassen, R. M., & Klassen, J. R. (2018). Self-efficacy beliefs of medical students: a critical review. *Perspectives on medical education, 7*(2), 76-82. https://doi.org/10.1007/s40037-018-0411-3
- Papert, S. (1980). " Mindstorms" Children. Computers and powerful ideas.

- Qu, J. R., & Fok, P. K. (2021). Cultivating students' computational thinking through student–robot interactions in robotics education. *International Journal of Technology and Design Education*, 1-20. https://doi.org/10.1007/s10798-021-09677-3
- Selby, C. C. (2015). Relationships: computational thinking, pedagogy of programming, and Bloom's Taxonomy. *Proceedings of the workshop in primary and secondary computing education.* https://doi.org/10.1145/2818314.2818315
- Su, J., Zhong, Y., & Ng, D. T. K. (2022). A meta-review of literature on educational approaches for teaching AI at the K-12 levels in the Asia-Pacific region. *Computers and Education: Artificial Intelligence, 3*, 100065. https://doi.org/10.1016/j.caeai.2022.100065
- Tsai, M.-J., Wang, C.-Y., Wu, A.-H., & Hsiao, C.-Y. (2021). The development and validation of the robotics learning self-efficacy scale (RLSES). *Journal of Educational Computing Research*, *59*(6), 1056-1074. https://doi.org/10.1177/0735633121992594
- Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33-35.
- Witherspoon, E. B., Higashi, R. M., Schunn, C. D., Baehr, E. C., & Shoop, R. (2017). Developing computational thinking through a virtual robotics programming curriculum. ACM Transactions on Computing Education (TOCE), 18(1), 1-20. https://doi.org/10.1145/3104982
- Yang, K., Liu, X., & Chen, G. (2020). The influence of robots on students' computational thinking: A literature review. *International Journal of Information and Education Technology, 10*(8), 627-631. https://doi.org/10.18178/ijiet.2020.10.8.1435

## Using a four-step learning activity in a programming course: classroom participation, learning performance, and attitude

Liang-Yi Lla\*, Wen-Lung HUANG<sup>b</sup> & Chien-Ping CHUNG<sup>c</sup>

 <sup>a</sup>Program of Learning Sciences, Institute for Research Excellence in Learning Sciences, National Taiwan Normal University, Taiwan
 <sup>b</sup>Department of Communication, Fo Guang University, Yilan, Taiwan
 <sup>c</sup> Department of Information and Finance Management, National Taipei University of Technology, Taiwan
 \*lihenry12345@ntnu.edu.tw

**Abstract:** In a programming course, a teacher has to teach not only programming concepts and syntax but also how to implement programs. However, teachers do not have enough time to complete the tasks. To address the problem, this study developed a four-step learning activity in a flipped classroom, which consists of testing, task assigning, collaborative problem-solving, and presentation. This study also conducted an instructional experiment for examining the effects of the four-step learning activity on students' classroom participation, learning performance, and course satisfaction. The results showed that the students with the four-step learning activity and the students with the traditional instruction demonstrated the same level of classroom participation. While the students with the four-step learning activity had higher learning performance than the students with the traditional instruction. The students perceived that the four-step learning activity can facilitate learning motivation, classroom interaction, learning performance, and help-seeking. They hoped that the four-step learning activity can be used in other programming courses in the future.

Keywords: flipped classroom, collaborative problem-solving, programming

#### 1. Introduction

Computational thinking (CT) refers to a set of problem-solving skills, such as solving problems, designing systems, and understanding human behavior, by drawing on concepts fundamental in computer science (Wing, 2006). It can be applied to solve everyday problems and has been advocated as a fundamental twenty-first century skill that students need to develop. Therefore, researchers, educators, and policy makers have concentrated on CT in the recent few years. In addition, the development of artificial intelligence (AI) technology has increased the accuracy of prediction and recognition. Some AI libraries have made programmers develop AI applications easier. More and more people and companies have begun to invest in research and industrial applications related to artificial intelligence. Therefore, AI-related applications are gradually appearing, such as image recognition, voice assistants, autonomous driving, and data science. Programming is considered to be a key tool for acquiring CT experience and developing CT skills and is the foundation of artificial intelligence to solve domain problems and design innovative applications, it is most important to cultivate students' programming concepts and abilities.

In programming courses, students need to understand the programming concepts (such as function, passing by value, and addressing) and syntax of programming languages (such as

the syntax of control structures, function declaration and call, and object-oriented), Students also need to acquire the abilities to use the learned programming concepts and syntax to solve computational problems. Therefore, programming is not only a course for conceptual understanding but also a subject that needs to be practiced. In general, programming courses may be two credits or three credits a semester. In a traditional lecture-based classroom, teachers may not have enough time to teach programming concepts and syntax and guide students to implement programming codes. To solve the problem, the flipped classroom may be a good approach. In a flipped classroom, teachers can record a video that introduces the programming concepts and syntax taught this week; and students can watch the video to learn these concepts and syntax before the class. In the classroom, teachers can assign several computational problems for the students to practice the programming concepts and syntax. When students implement their codes, teachers can provide appropriate guidance and suggestions for students and can also have more interactions with the students in the classroom.

However, it may not be enough to just upload the videos on the Web for students to watch before class, and to assign several programming problems for students to solve in the classroom. Students may not actively watch the pre-class videos, and may not have the ability and motivation to solve the computational problems in the classroom. Ultimately, let the flipped classroom approach failed. In order to enhance the flipped classroom approach, teachers should design suitable instructional activities for facilitating pre-class video viewing and inclass activities. This study designed a four-step activity, including quizzes, task assignments, cooperative problem-solving, and code explanation to enhance the flipped classroom approach. In addition, this study also conducted a four-week experiment for understanding the impact of the four-step activity on classroom participation, learning performance, and course satisfaction.

#### 2. Flipped classroom instruction

The traditional lecture-based approach uses a teacher-centered method. Students passively accept the knowledge taught by teachers and have fewer opportunities to interact with teachers and classmates. This approach cannot effectively cultivate students' self-discipline and high-level thinking skills (Lai & Hwang, 2016). The flipped classroom instruction responds to this issue by transforming teaching into a student-centered approach. It is divided into two activities: pre-class activities and in-class activities (Long, Cummins, & Waugh, 2017). In the pre-class activities, students build up their basic knowledge through different online learning activities, such as video viewing activities. In the in-class activities, the teacher conducts different learning activities, such as group cooperative learning, problem-based learning, and project-oriented learning. Students apply the knowledge they have acquired before class to develop cooperation and higher-level thinking skills.

Although flipped classroom instruction offers these benefits, previous studies have shown mixed results. Karabulut-Ilgu et al. (2018) analyzed 30 papers that compared the traditional lecture-based approach with flipped classroom instruction and found that 15 papers reported that the students in flipped classroom instruction demonstrated better learning outcomes than the students in the traditional approach, 5 papers reported mixed results, 8 papers reported no difference, and 2 papers reported that flipped classroom instruction had worse learning outcomes than the traditional approach. The mixed results may be due to the pre-class learning activities and in-class learning activities. Therefore, more studies are needed to help us understand how to conduct pre-class and in-class activities in flipped classroom instruction for maximizing students' learning outcomes.

In-class activities are very important for flipped classroom instruction because they help students apply, integrate and evaluate the knowledge learned in the pre-class activities. In the classroom, teachers generally conducted cooperation, problem-solving, or project-based

learning activities. For example, Chen , Chen, & Chen (2015) used the flipped classroom approach in a statistics course. They used a cooperative learning activity in the classroom. The students were divided into groups. Group members discussed the pre-class videos. After the discussion, the teacher gave feedback for the students' discussion. Finally, each group completed a group report. Sun, Xie, & Anderman (2018) used the flipped classroom approach in a calculus course. In the in-class activity, the teacher first explained the content of the pre-class videos with some examples and then divided the students into groups. Group members have to collaboratively solve several problems. Finally, the teacher selected several groups to report their answers.

The flipped classroom approach has also been applied to programming courses. In programming courses, the in-class activities were similar to those used in other courses, including cooperation, problem-solving, and testing. For example, Chiang (2017) used the flipped classroom approach in an object-oriented analysis and design course, the students went to edX to watch instructional videos before class. They needed to solve problems encountered in edX during class. Durak (2018) used the flipped classroom approach in a programming course. In the classroom, the students solved several computational problems and then collaboratively complete a group project. Chis, Moldovan, Murphy, Pathak, & Muntean (2018) used the flipped classroom approach in a programming course. In the classroom, the students first solved several computational problems and then completed a test. Cakiroglu & Ozturk (2017) used the flipped classroom approach in a programming course. In the classroom, the teacher divided the students into groups and then gave the students several computational problems. The students then collaboratively solved them. Finally, They reported their answers. These studies used different in-class activities for the flipped classroom approach. Therefore, this study refers to the activities used in the previous studies to develop a four-step in-class activity.

#### 3. Four steps in-class activity

This study used the flipped classroom approach in a programming course. The flow of the weekly flipped classroom activity is as follows:

- 1. The teacher published a video lecture on a learning management system before one week of the class.
- 2. Students watch videos within a week.
- 3. In the classroom, the teacher uses the four-step in-class activity. The four steps are testing, task assignment, collaborative problem solving, and code explanation.

Step 1 is testing. In this step, students conducted an online test at the beginning of the class. A testing system was developed. Teachers can build up a test sheet by the testing system. The testing system allowed teachers to design three types of questions: multiple choice, filling, and open-ended questions. When students conduct a test, the test sheet was displayed. The students then answered the questions in order. The order of the questions is random, so the students answered the same questions, but the order of the questions was different, which can reduce the chance of students' cheating.

After the testing step, the students then worked in a group of two to solve several computational problems (step 2 task assignment). The grouping of the students considered the students' preferences and their academic performance. The problems assigned by the teacher considered the students' academic performance and testing outcomes. After the grouping and problems assigned, the students collaboratively solved the problems in groups (step 3 collaborative problem solving). If they cannot solve the problems, they can ask the teacher for assistance. Finally, the teacher selected a group for each problem. The selected groups explained their code about the computational problems (step 4 code explanation).

#### 4. Method

This study conducted an experiment. The students in the control group were 37 undergraduate students. They took part a programming course for 16 weeks to learn python programming language. The course teacher used a traditional lecture-based approach in the course. The teacher first introduced the programming concepts in the classroom. He then assigned several computational problems related to the introduced programming concepts and syntax, and the students solved the problems in the classroom. Therefore, for a 3-hour in-class activity, the teacher taught programming concepts and syntax for about 1.5 hours, and the students wrote the code for the problems for 1.5 hours. The teacher also assigned 1-3 computational problems as weekly homework assignments. The students must complete the assignments and upload their assignments to the learning management system after class.

The students in this experimental group were 33 undergraduate students. They took part a programming course to learn python programming language. The course teacher and learned content were the same as the control group. However, because of the course schedule, the teacher used the traditional lecture-based approach before mid-term exam and used flipped classroom approach with the four-step in-class activity after the mid-term exam. Therefore, our analysis focuses on the learning performance after the mid-term exam.

#### 5. Results

#### 5.1 Classroom participation

In this study, the number of roll calls after the midterm exam was used as the degree of classroom participation. There were 5 roll calls after the midterm exam of the control group, and 8 roll calls after the midterm exam of the experiment group. We first calculated the percentage of roll calls, and then used the independent sample t-test to compare the difference in the percentage of roll rolls between the two groups. The results showed that the percentage of roll calls of the students in the control group (N=37, mean=88.65, SD=21.36) was not significantly different (t=-1.315, p=0.193) from the students in the experiment group (N=33, mean=94.32, SD=14.35). The result meant that the four-step activity has no significant impact on students' course participation.

#### 5.2 Learning performance

This study compared the students' final exam scores. An independent sample test was conducted to compare the final exam scores of the two groups. The results showed that the students in the control group had significantly lower (t=-6.010, p<0.000) final exam scores (N=37, mean=24.54, SD=19.42) than the students in the experiment group (N=33, mean=53.45, SD=20.83), which indicates that the four-step in-class activity can significantly improve the learning outcomes of the students.

#### 5.3 Students' attitude to the four-step activity

At the end of the course, the students in the experiment group must answer an attitude questionnaire. The questionnaire contains 8 questions, using a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree), filled in anonymously. The results are shown in Table 1, which shows that students believe that the flipped classroom approach with the four-step activity had better effectiveness on learning motivation, learning performance, classroom interaction, and help-seeking than the traditional lecture-based approach. They hoped that the flipped classroom approach can be used in future programming courses.

Table1: Students' attitude to the four step approach	
--	--

question	Mean	SD

Compared with the traditional lecture-based approach, I think the flipped classroom	3.83	0.91
can improve my learning motivation		
Compared with the traditional lecture-based approach, I think the flipped classroom approach with the four-step in-class activity can improve my learning performance	3.90	0.84
Compared with the traditional lecture-based approach, I think the flipped classroom approach with the four-step in-class activity increase my interaction with the teacher.	3.41	1.00
Compared with the traditional lecture-based approach, I think the flipped classroom approach with the four-step in-class activity can increase my interaction with my classmates.	4.03	0.96
Compared with the traditional lecture-based approach, I think the flipped classroom approach with the four-step in-class activity is more interesting	3.86	0.83
When I have problems in class, I can get help more easily in the flipped classroom approach with the four-step in-class activity than in the traditional lecture-based classroom.	3.79	0.92
Compared with the traditional lecture-based approach, I prefer the flipped classroom approach with the four-step in-class activity.	3.66	0.80
Compared with the traditional teaching method, I hope that the future programming courses can use the flipped classroom approach with the four-step in-class activity	3.62	0.72

#### 6. Conclusion

This study designed a four-step in-class activity for the flipped classroom approach. An experiment showed that the students in the flipped classroom approach with the four-step activity and the students in the traditional lecture-based approach had the same classroom participation. However, the students in the flipped classroom approach with the four-step activity had better learning performance than the students in the traditional lecture-based approach. The students also believed that the flipped classroom approach with the four-step activity can promote learning motivation, classroom interaction, learning performance, and help-seeking. They also hoped that the flipped classroom approach with the four-step activity can be used in the future programming courses. These results showed that the flipped classroom approach with the flipped classroom approach with the flipped classroom approach with the four-step activity can be used in the future programming courses.

#### References

Cakiroglu, U., & Ozturk, M. (2017). Flipped Classroom with Problem Based Activities: Exploring Selfregulated Learning in a Programming Language Course. *Educational Technology & Society*, *20*(1), 337-349. Chen, L. W., Chen, T. L., & Chen, N. S. (2015). Students' perspectives of using cooperative learning in a flipped statistics classroom. *Australasian Journal of Educational Technology*, *31*(6), 621-640.

- Chiang, T. H. C. (2017). Analysis of learning behavior in a flipped programing classroom adopting problem-solving strategies. *Interactive Learning Environments, 25*(2), 189-202. doi:10.1080/10494820.2016.1276084
- Chis, A. E., Moldovan, A. N., Murphy, L., Pathak, P., & Muntean, C. H. (2018). Investigating Flipped Classroom and Problem-based Learning in a Programming Module for Computing Conversion Course. *Educational Technology & Society, 21*(4), 232-247.
- Durak, H. Y. (2018). Flipped learning readiness in teaching programming in middle schools: Modelling its relation to various variables. *Journal of Computer Assisted Learning,* 34(6), 939-959. doi:10.1111/jcal.12302
- Karabulut-Ilgu, A., Cherrez, N. J., & Jahren, C. T. (2018). A systematic review of research on the flipped learning method in engineering education. *British Journal of Educational Technology*, 49(3), 398-411. doi:10.1111/bjet.12548
- Lai, C. L., & Hwang, G. J. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education, 100*, 126-140. doi:10.1016/j.compedu.2016.05.006
- Long, T. T., Cummins, J., & Waugh, M. (2017). Use of the flipped classroom instructional model in higher education: instructors' perspectives. *Journal of Computing in Higher Education*, 29(2), 179-200. doi:10.1007/s12528-016-9119-8
- Sun, Z. R., Xie, K., & Anderman, L. H. (2018). The role of self-regulated learning in students' success in flipped undergraduate math courses. *Internet and Higher Education*, 36, 41-53. doi:10.1016/j.iheduc.2017.09.003

## Review of TAM used in Educational Technology Research: A Proposal

Kai-Yu TANG<sup>a\*</sup>, Chun-Hua HSIAO<sup>b</sup>

<sup>a</sup>Graduate Institute of Library & Information Science, National Chung Hsing University, Taiwan

<sup>a</sup>Innovation and Development Center of Sustainable Agriculture, National Chung Hsing University, Taiwan

<sup>b</sup>School of Business, Kainan University, Taiwan \*ky.nctu@gmail.com; kytang@dragon.nchu.edu.tw

Abstract: Since the research on user acceptance of information technology was introduced by Fred D. Davis in the 1980s, two critical constructs (perceived usefulness and ease of use) and the technology acceptance model (TAM) have generated significant impacts on information science and management information system literature. In the early 2000s, educational technology researchers began to adopt TAMrelated concepts to explain teaching and learning contexts. For example, Selim (2003) conducted research to empirically examine students' acceptance of course websites and published on Computers & Education. Accordingly, this proposal paper extends the authors' previous works (Hsiao & Yang, 2011; Hsiao, Tang & Liu, 2015; Tang, Hsiao & Chen, 2019; Tang, Hsiao & Hwang, 2022) to trace the latest development of educational technology research using TAM-related framework. In this proposal, we have identified a total of 366 articles with 20,242 citations. For the data analysis, we proposed a series of analyses (e.g., bibliometrics, main path analysis, document cocitation analysis, and social network analysis) to examine the knowledge trajectory in this field. This study is the first attempt to propose a research framework with a series of analyses to systematically analyze and review the longitudinal development of research using a TAM-related framework in educational technology. With this attempt, some significant turning points (e.g., theoretical converging and diverging points) can be further explored in the network diagram to identify the critical constructs, models, and theories in educational technology research. Overall, established on a large number of citation data, the results of this study aim to provide a platform for academic discussion and to make a prediction for the future trend of educational technology research.

**Keywords:** Technology acceptance, TAM, educational technology, bibliometric features, citation-based network analysis

#### 1. Introduction

Fred D. Davis introduced the user acceptance of information technology in 1989. He proposed the theoretical framework of the technology acceptance model (TAM) in the same year (Davis et al., 1989). Since then, TAM has become one of the dominant research models in the technology management discipline (Venkatesh, Davis, & Morris, 2007). Recently, according to the Web of Science (WoS), Davis's original paper titled "Perceived usefulness, perceived ease of use, and user acceptance of information technology" has been cited over 22,473 times (the search was conducted on April 29, 2023).

The theoretical model was originally developed for technologies in the context of organizations (Davis, 1989), which hypothesized that the actual use of a certain technology is directly influenced by a person's behavioral intention to use, which in turn, is determined by perceived usefulness (PU) and attitude toward the technology. Moreover, PU and attitude are affected by perceived ease of use (PEOU). Previous researchers have suggested that the main strengths of TAM are its parsimony and strong generalizability (Lee et al., 2003;

Plouffe et al., 2002; Hsiao & Yang, 2011; Hsiao et al., 2015). The model has evolved and been widely applied to various technology-related adoption behaviors.

Up to the date of this research, over three thousand journal papers (n = 3,298) are conducting technology acceptance-related research in the social science citation index (SSCI) database. Among which, a total of 366 articles published in core educational technology journals were selected for analysis. The top ten educational technology journals were *Computers & Education, Educational Technology & Society, British Journal of Educational Technology, Interactive Learning Environments, Education and Information Technologies, Australasian Journal of Educational Technology, Educational Technology <i>Research and Development, International Review of Research in Open and Distributed Learning, Journal of Educational Computing Research, Journal of Computing in Higher Education.* Each journal has published at least 14 TAM-related articles and has been cited more than 493 times.

Many review papers conducted meta-analysis and content analysis for the literature. While meta-analysis is useful in distinguishing the interrelationships among TAM factors across different settings, content analysis is based on coding schemes from existing literature to identify and categorize patterns and themes in the research of interest. This present study extends and modifies the authors' previous works (Hsiao & Yang, 2011; Hsiao, Tang & Liu, 2015; Tang, Hsiao & Chen, 2019; Tang Hsiao & Hwang, 2022) by proposing a series of bibliometric and citation-based network analyses to map the intellectual structure and synthesize the trends of technology adoption research.

Specifically, based on extensive literature citation data, this research proposal aims to: (1) identify key bibliometric features of the literature in the educational technology domain, including keywords, core journals, productive countries, and authors. (2) identify the most critical paths on the citation network of various adoption theories using main path analysis. (3) based on the result, further classification and exploration within subfields and important theories will be examined using co-citation network analysis. To obtain the above goals, a series of following analyses will be conducted: bibliometric and citation analysis, main path analysis, co-citation analysis, social network analysis, and statistical analysis.

This proposal paper offers valuable contributions. First, this present study is a few of the studies that apply bibliometric techniques to the technology acceptance research literature. Second, the integrated analyses, including citation-based network, co-citation analysis, and statistical clustering analysis, complement and improve the findings of other studies that have approached the subject from both qualitative and quantitative perspectives.

## 2. Past bibliometrics and citation-based analyses on TAM and TAM-related research in educational technology domain

The previous meta-research did help researchers revalidate the importance of PU and PEOU in the model and distinguish the interrelationships among other TAM factors across different settings. However, only a few provided citation-based evidence to support the development and progression of TAM to respond to Venkatesh et al.'s (2007) concerns. Using citations can show how knowledge disseminates within scientific disciplines (Garfield et al., 1964). Furthermore, the count of citations is currently treated as one common means to demonstrate the general acceptance of an academic research article.

Hsiao and Yang (2011) adopt a co-citation analysis and other statistical analyses to identify core documents within the same discipline or field by analyzing the references. Their results concluded three main trends of TAM: task-related systems, e-commerce systems, and hedonic systems. Hsiao et al. (2015) conducted a citation-based analysis of technology acceptance research. They used main path analysis and edge-betweenness clustering analysis. Results identified five distinct research fronts, including e-learning, mobile-commerce, e-health, e-tourism, and technology post-acceptance research, and provides a research-based platform for further scholarly discussions on the theoretical development trajectories and most active research fronts in the field. Following previous studies, Tang et al. (2019) reviewed the characteristics of electronic commerce innovations using bibliometric and keyword analysis. The findings show that in the first decade (2000-2009), the top 10

papers in electronic commerce innovation focused on traditional IT adoption, while in the recent decade (2010-2018), researchers showed more diverse interest in innovative EC applications such as RFID, cloud computing, and crowdsourcing. Along with these innovation contexts, researchers adopted more theories, such as signaling theory, in addition to general attitude-intention theories. More recently, Tang et al. (2022) continued the survey to explore the scholarly network of artificial intelligence (AI)-related research in the information science (IS) domain using co-citation network analysis. The findings indicate that both researchers from Global North (GN) and Global South (GS) focused on technology adoption research throughout the investigated period (2010-2020). While GN researchers in the IS domain focused on applied research involving intelligent systems between 2010 and 2015, GS researchers focused on big data applications between 2016 and 2020.

According to the above review studies of TAM, this paper proposed a series of bibliometric and citation-based network analyses to analyze the extensive bibliographic citations of TAM research papers published in well-recognized journals. The finding is expected to provide insights into tracing the trajectory of TAM literature development and visualize the most critical citation routes in the citation network of technology acceptance research.

#### 3. Proposed Methodological Framework

#### 3.1 The process of data inclusion

The data inclusion of this research was as follows. First, an initial search using the full model name of the technology acceptance model is conducted on the Web of Science (WoS) to have an overview of technology adoption studies. The WoS is one of the reputed sources for searching academic literature, covering the most high-quality journals and research papers in the technology management field, including education and educational research. A series of key terms, including perceived usefulness, perceived ease of use, technology acceptance, and technology adoption, were used for the topics search (title, keywords, and abstract) in the WoS.

Up to the date of this research, over three thousand journal papers (n = 3,298) are conducting technology acceptance-related research in the social science citation index (SSCI) database. Among which, a total of 366 articles published in core educational technology journals were selected for analysis (Chang et al., 2022; Hwang & Tsai, 2011; Hwang & Wu, 2014). The top ten educational technology journals were *Computers* & *Education, Educational Technology & Society, British Journal of Educational Technology, Interactive Learning Environments, Education and Information Technologies, Australasian Journal of Educational Technology, Educational Technology Research and Development, International Review of Research in Open and Distributed Learning, Journal of Educational Computing Research, Journal of Computing in Higher Education.* Each journal has published at least 14 articles related to TAM theories, constructs, and research models in educational contexts. In total, 366 articles have received over 20,242 citations, demonstrating their significant impact in the field.

#### 3.2 Bibliometrics and citation analysis

Bibliographic analysis is a popular research approach that involves quantitative evaluation of various scientific publications like books, conference proceedings, and journal articles. It provides crucial insights into the research productivity and impact of researchers, institutions, and countries, as well as the trend of the field (Tang et al., 2023). Policymakers, funding agencies, and researchers use bibliographic analysis to evaluate research impact and productivity. Citation analysis is a crucial component of bibliographic analysis as it helps to identify highly cited or influential documents within a specific field. This is based on the assumption that heavily cited papers have a greater impact on the subject matter than those less frequently referenced. Citation analysis is considered "the field's view of itself" as it traces the links and interactions between researchers and different research fields (White &

Griffith, 1981). Despite some limitations, such as negative citation, citation analysis is a valuable and cost-effective tool to evaluate scientific performance.

#### 3.3 Main path analysis

Main path analysis has been proposed as a useful quantitative and citation-based approach for analyzing significant trajectories, identifying significant papers, and depicting recent active research areas. This method represents scientific papers as nodes in a citation network, with links established based on citation information. The 'source' nodes are papers cited but do not cite any other papers, while the 'sink' nodes are papers that cite others but are not themselves cited. Sources represent the origins of knowledge flow, while sinks represent the end nodes of current knowledge dissemination.

Hummon and Doreian (1989) introduced the 'local main path' concept, which uses a 'priority first search' algorithm to identify the only path representing the main path of a citation network. Liu and Lu (2012) proposed the 'global main path,' which determines the path with the most significant accumulated SPC value. While the local main path emphasizes progressing significance, the global main path emphasizes overall importance. Liu and Lu also introduced the concept of a 'key-route main path' to identify structural patterns, particularly the divergence-convergence pattern of knowledge diffusion flows. The key-route main path often includes local and global main paths and has several merits for large citation networks. Several recent studies have demonstrated that MPA is a well-valid approach applied in various scientific disciplines, such as in technology management (Cho et al., 2021), public health (Huang et al., 2021), and corporate governance (Hung et al., 2022).

#### 3.4 Co-citation network analysis

Co-citation network analysis is based on the calculation of citation times that two documents or authors are cited together in the same work (Small, 1973). In particular, periodical papers with peer review have shown their reliability after careful evaluation, which further confirms the value of co-citation analysis (Ramos-Rodríguez & Ruíz-Navarro, 2004). co-citation analysis is often used to measure the network or the degree of relationships between documents according to the counts of their joint citations (Small, 1973). Co-citation counts can be statistically analyzed and processed to produce maps showing the relative closeness between documents. The more counts of co-citation two documents receive, the higher their co-citation strength, and the more likely they are to be bibliographically related. In other words, these highly co-cited papers are assumed to have a higher degree of similarity.

Network researchers suggested that social network analysis provides a novel approach to visualizing a network's most prominent documents (Wasserman & Faust, 1994). This visualization technique helps realize and map interdisciplinary scholarly communications in various academic disciplines. Recently, researchers have combined the methods (i.e., cocitation network analysis) to explore the relationships in multiple domains, such as information literacy (Chen et al., 2022), virtual reality education research (Cheng et al., 2022), STEM education (Hsu et al., 2023), and AI in e-learning (Tang et al., 2023).

To conclude, bibliometric and citation-based network analyses are effective methods to identify the most influential documents, analyze the relational links between them, and identify the intellectual structure of documents that belong to the same discipline within the TAM literature. A proposed analytic framework is summarized and illustrated in Figure 1.

#### Acknowledgements

This work was financially supported by National Science and Technology Council, Taiwan under contract numbers NSTC 112-2410-H-005-024. This work was also financially supported by the "Innovation and Development Center of Sustainable Agriculture" from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan.

#### References

- Chang, C. C., Hwang, G. J., & Tu, Y. F. (2022). Roles, applications, and trends of concept mapsupported learning: a systematic review and bibliometric analysis of publications from 1992 to 2020 in selected educational technology journals. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2022.2027457
- Chen, C. C., Wang, N. C., Tang, K. Y., & Tu, Y. F. (2022). Research issues of the top 100 cited articles on information literacy in higher education published from 2011 to 2020: A systematic review and co-citation network analysis. *Australasian Journal of Educational Technology*, 34-52.
- Cheng, K. H., Tang, K. Y., & Tsai, C. C. (2022). The mainstream and extension of contemporary virtual reality education research: Insights from a co-citation network analysis (2015–2020). *Educational Technology Research and Development*, 70(1), 169-184.
- Cho, R. L. T., Liu, J. S., & Ho, M. H. C. (2021). The development of autonomous driving technology: perspectives from patent citation analysis. *Transport Reviews*, 41(5), 685-711.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science*, 35(8), 982-1003.
- Davis, Fred D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- Hsiao, C. H., & Yang, C. (2011). The intellectual development of the technology acceptance model: A co-citation analysis. *International Journal of Information Management*, 31(2), 128-136.
- Hsiao, C. H., Tang, K. Y., & Liu, J. S. (2015). Citation-based analysis of literature: a case study of technology acceptance research. *Scientometrics*, 105(2), 1091-1110.
- Hsu, Y. S., Tang, K. Y., & Lin, T. C. (2023). Trends and hot topics of STEM and STEM education: a coword analysis of literature published in 2011–2020. *Science & Education*. https://doi.org/10.1007/s11191-023-00419-6
- Huang, C. H., Chou, T. C., & Liu, J. S. (2021). The development of pandemic outbreak communication: A literature review from the response enactment perspective. *Knowledge Management Research* & *Practice*, 19(4), 525-535.
- Hummon, N. P., & Dereian, P. (1989). Connectivity in a citation network: The development of DNA theory. *Social Networks*, 11(1), 39-63.
- Hung, S. C., Lai, J. Y., & Liu, J. S. (2022). Mapping technological trajectories as the main paths of knowledge flow: Evidence from printers. *Industrial and Corporate Change*, 31(3), 863-889.
- Hwang, G. J., & Tsai, C. C. (2011). Research trends in mobile and ubiquitous learning: A review of publications in selected journals from 2001 to 2010. *British Journal of Educational Technology*, 42(4), E65-E70.
- Hwang, G. J., & Wu, P. H. (2014). Applications, impacts and trends of mobile technology-enhanced learning: a review of 2008–2012 publications in selected SSCI journals. *International Journal of Mobile Learning and Organisation*, 8(2), 83-95.
- Selim, H. M. (2003). An empirical investigation of student acceptance of course websites. *Computers & Education*, 40(4), 343-360.
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for information Science*, 24(4), 265-269.
- Tang, K. Y., Chang, C. Y., & Hwang, G. J. (2023). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, 31(4), 2134-2152.
- Tang, K. Y., Hsiao, C. H., & Chen, M. C. (2019). A research survey of electronic commerce innovation: evidence from the literature, *Advances in Technology Innovation*, 4(4), 247–259.
- Tang, K. Y., Hsiao, C. H., & Hwang, G. J. (2022). A scholarly network of AI research with an information science focus: Global North and Global South perspectives. *PLOS One*, 17(4), e0266565.
- Venkatesh, V., Davis, F., & Morris, M. G. (2007). Dead or alive? the development, trajectory and future of technology adoption research. *Journal of the Association for Information Systems*, 8(4), 267-286.
- Wasserman, S. Faust, K. (1994). Social network analysis: Methods and applications. Cambridge, UK: Cambridge University Press.
- White, H. D., & Griffith, B. C. (1981). Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for Information Science*, 32(3), 163-171.

## Exploring Motivational Differences in Competitive and Cooperative Game-Based Learning through Educational Board Games

Yuan-Yao Chiang, Ting-Chia HSU<sup>\*</sup>

Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taiwan, ROC Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taiwan, ROC \*ckhsu@ntnu.edu.tw

**Abstract:** The purpose of this study was to determine whether educational board games have a significant impact on learning motivation by developing the information media literacy educational board game Hi-Life. It also examined the differences in learning motivation between the competitive group and the cooperative learning group through a randomized experimental design. A total of 26 students from a national university in northern Taiwan participated in the study, 14 students in the control group for cooperative game-based learning and 12 students in the experimental group for competitive game-based learning. Through the analysis of the Mann-Whitney U test, the results showed that the dimension of Relevance was significant in both the experimental and control groups. Although the rest of the dimensions are not significant, the overall data showed that the average values of the three dimensions of Relevance, Confidence and Satisfaction were all higher than 3.5. The students demonstrated a favorable learning motivation in the ARCS model after playing the educational board game.

**Keywords:** Game-based learning, competitive game-based learning, cooperative game-based learning, learning motivation

#### 1. Introduction

Because technology enhances the pace of information exchange, people often receive the latest information about their lives through smart mobile devices such as computers and cell phones. As technology advances, people can post their own comments or share information on social media, which has raised concerns about the reliability of information. Thus, information media literacy has become an important issue that people need to be aware of. Therefore, there is an urgent need to develop innovative instructional material of information media literacy nowadays.

Game-based learning refers to the reimagining of learning tasks to make them more engaging and meaningful, ultimately leading to effective learning outcomes (Plass et al., 2020). Moreover, game-based learning incorporates various motivational elements for learners, such as enjoyment, competition, cooperation, and challenges (Olejniczak et al., 2020). In the realm of competitive and cooperative game-based learning, certain studies indicate that both competition and collaboration can provide beneficial learning effects for students, fostering a harmonious overall teaching atmosphere and promoting active engagement in the classroom (Rodríguez et al., 2023).

Collaborative game-based learning enhances the positive atmosphere in the classroom and encourages students to actively participate in learning activities (Naumoska et al., 2023). Competitive game-based learning makes learning more enjoyable and increases

student engagement and motivation. However, it may also lead to negative teaching environments and detrimental effects if not well-designed, emphasizing the importance of effective instructional design (Altawalbeh et al., 2023).

The sample of this study was drawn from students who took the theories and practice of digital learning at a national university in Taiwan. They were randomly divided into two groups: the cooperative group as a control group and the competitive group as an experimental group. Participants will learn information media literacy through educational board games. The study attempted to investigate the effect of game-based learning on learning motivation. It will also explore whether there was a difference in learning motivation between the cooperative group and the competitive group. Therefore, the research questions are as follows: (1).Did educational board games have a significant impact on learning motivation of information media literacy? (2).Was there a difference in learning motivation of information media literacy between competitive game-based learning and cooperative gamebased learning?

#### 2. Literature

#### 2.1 Educational Board Games

Recently, game-based learning (GBL) has become popular in both teaching and learning. Compared to traditional learning methods, GBL can increase the effectiveness and efficacy of learning (Lin, 2019) and is more likely to increase motivation to learn (Plass et al., 2015). Among the many types of GBL, the design of a board game is the most commonly used method. Educational board game refers to use the teaching materials related to educational themes to design the board game. Students are immersed in the contexts or story situation to explore and learn through an event or storyline. In order to improve students' learning effectiveness and stimulate their learning interest and motivation while transforming the traditional lecture-based teaching approach, educational board games have been recently integrated into the curriculum to enable students to learn through playing the games.

#### 2.2 Competitive vs cooperative GBL

Apart from playing board games, some studies have incorporated other elements into course design such as competition and cooperation in GBL. For example, researchers like Kirschner et al. (2011) gave students the opportunity to work individually or collaboratively to solve biology-related problems. Results showed that students were better able to handle complex and cognitively demanding problems when working cooperatively. Fu et al. (2009) helped students learn about computers through the web-based game-based learning; and findings suggested that both cooperative and competitive learning elements can boost students' enjoyment of learning and stimulate different types of knowledge growth. Students are more likely to exhibit stronger analytical skills when competition is introduced. On the other hand, cooperative learning helps create a greater degree of generalized abilities.

#### 3. Method

The developed instructional board game was employed in this empirical study, with four information media literacy-related themes of copyright, personal assault, fake news, and fraud. A total of 26 students from a national university in northern Taiwan participated in the study, with 14 students in the control group and 12 students in the experimental group. The three-lesson experiment period with a total of 120 minutes was planned for the study. Prior to the study, the instructor explained information media literacy and rules of board games in 20 minutes. Then, the intervention was conducted using the designed themes in the board game. After the game, students completed the questionnaires on the learning motivation.

The study used the ARCS Learning Motivation Scale developed by Keller (2010), which consists of four dimensions: Attention (with an alpha value of .83), Relevance (with an alpha value of .81), Confidence (with an alpha value of .92), and Satisfaction (with an alpha value of .90). In this study, 4 to 5 questions were selected from each dimension, with a total of 17 questions. The questions were scored on a 5-point Likert scale, with a score of 1 for "strongly disagree" to 5 for "strongly agree". The definitions and sample questions are shown in the table below.

Element	Definition	Questionnaire contents
Attention	Attract students' interest and curiosity	The application of this educational board game appeals to my curiosity in many ways.
Relevance	Make the needs of the students relevant to their objectives and foster a positive attitude toward learning	I can connect the board game's themes to things I've discovered, accomplished, or considered in my own life.
Confidence	Support students in believing that they can successfully learn and apply what they have learned	I am confident to be able to learn this course through this educational board game.
Satisfaction	Enabling students to feel a sense of satisfaction and accomplishment after learning	I feel satisfied when I have completed the activities in this educational board game.

Table 1. ARCS model

#### 4. Result

#### 4.1 The post test of ARCS learning motivation

Table 2 shows the scores of ARCS after playing board games. According to finding, the mean value of attention is 3.10, relevance 3.58, confidence 3.50, and satisfaction 3.93. The three dimensions' scores—relevant, confident, and satisfied—are all greater than 3.5. The three demonstrates of the students' learning motivations were achieved after playing educational board games.

	N	Min.	Max.	М	SD		
Attention	26	2.50	4.25	3.10	0.47		
Relevance	26	2.25	4.75	3.58	0.51		
Confidence	26	2.80	4.20	3.50	0.41		
Satisfaction	26	3.00	5.00	3.93	0.57		

Table 2. Result of descriptive statistics on ARCS

#### 4.2 The post test of ARCS learning motivation using Mann-Whitney U test

Table 3 presents the difference in posttest scores between the control group and the experimental group using Mann-Whitney U test. The results revealed that the only dimension which differed significantly was Relevance (U = 37.50, p = .015). All other dimensions were not significantly different. This finding shows that the story of board games is relevant to the real-life or actual contexts in the competitive scenario. The results also suggests that when designing the game, students can be given real-life examples to play that are relevant to their

own lives or that they have studied, considered, or reflected on in their own lives with competitive pressures.

	Control group		Expe	erimental	group	_Mann-Whitney	р	
	Ν	М	SD	Ν	М	SD	U	
Attention	14	2.92	0.27	12	3.29	0.52	47.50	.051
Relevance	14	3.36	0.49	12	3.83	0.42	37.50	.015*
Confidence	14	3.46	0.40	12	3.45	0.44	83.50	.979
Satisfaction	14	3.79	0.63	12	4.10	0.46	62.50	.255

Table 3. Result of Mann-Whitney U tests for the experimental and control groups on ARCS

\*p<.05

#### 5. Discussion and conclusions

The sample of this study consisted of 26 students who were enrolled in a class on the theory and practice of digital learning. They were randomly divided into two groups: the cooperative group for the control group and the competitive group for the experimental group. Through Hi-Life, an instructional board game created in this study, students can access information media literacy relevant to life-related information. In this study, the impact of educational board game learning on learning motivation was investigated, along with exploring whether there was a difference in learning motivation between the cooperative group and the competitive group.

Through the analysis of the Mann-Whitney U test, the results showed that the dimension of Relevance was significant in both the experimental and control groups and the rest of the dimensions of ARCS are not significant. This suggests that information media literacy learning through board games have no significant effect on the ARCS learning motivation. The overall mean values of the three dimensions of relevance, confidence, and satisfaction are all higher than 3.5, despite the fact that there is no significant effect on each ARCS learning motivation dimension. This indicates that the students' ARCS learning motivation is still positive after playing the board game. This also aligns with the research conducted by Rodríguez et al., where both competition and collaboration can provide meaningful learning outcomes for students. Furthermore, collaborative game-based learning and competitive game-based learning motivation (Naumoska et al., 2023; Altawalbeh et al., 2023).

In addition, only Relevance was found to be significant in the Mann-Whitney U test results, perhaps because the story situations of the board games were more relevant to students' daily lives. When students played the board games, they were able to associate them with situations that were related to their daily lives or experiences that had happened to them. In addition, the mean values of the experimental group were partially larger than those of the control group, which could be inferred from the fact that the competitive feature was more motivating to the students than the cooperative feature. This finding also echoes Fu et al.'s (2009) finding that students show higher analytical ability when competition is present. In other words, during the process of playing the board game, students compare their own experience with the story situation of the board game and analyze it to obtain scores.

The study aimed to investigate whether educational board games have an impact on the ARCS learning motivation, with the limited focus on the educational themes of information media literacy. Therefore, the research limitation is the challenging to determine whether the results of this study can be applied to other domains as well. In the future, it is expected to increase the comparison of pre- and post-tests results between the control group and the experimental group in the future. The sample size can also be increased to ensure greater confidence in the results of the learning difference in the ARCS motivation.

#### Acknowledgements

This study is partially supported by the National Science and Technology Council under contract number NSTC 111-2410-H-003-168-MY3.

#### References

- Altawalbeh, K., & Irwanto, I. (2023). Game-Based Learning: The Impact of Kahoot on a Higher Education Online Classroom. *Journal of Educational Technology and Instruction*, *2*(1), 30-49.
- Fu, F. L., Wu, Y. L., & Ho, H. C. (2009). An investigation of coopetitive pedagogic design for knowledge creation in web-based learning. *Computers & Education*, *53*(3), 550-562.
- Keller, J. M. (2010). Motivational Design for Learning and Performance: The ARCS Model Approach. Springer. http://dx.doi.org/10.1007/978-1-4419-1250-3
- Kirschner, F., Paas, F., Kirschner, P. A., & Janssen, J. (2011). Differential effects of problem-solving demands on individual and collaborative learning outcomes. *Learning and Instruction*, *21*(4), 587-599.
- Lin, T. J., Duh, H. B. L., Li, N., Wang, H. Y., & Tsai, C. C. (2013). An investigation of learners' collaborative knowledge construction performances and behavior patterns in an augmented reality simulation system. *Computers & Education*, *68*, 314-321.
- Naumoska, A., Dimeski, H., & Stojanovska, M. (2023). Using the Escape Room game-based approach in chemistry teaching. *Journal of the Serbian Chemical Society*, (00), 88-88.
- Olejniczak, K., Newcomer, K. E., & Meijer, S. A. (2020). Advancing evaluation practice with serious games. *American Journal of Evaluation*, *41*(3), 339-366.
- Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational psychologist*, *50*(4), 258-283.
- Plass, J. L., Homer, B. D., MacNamara, A., Ober, T., Rose, M. C., Pawar, S., & Olsen, A. (2020). Emotional design for digital games for learning: The effect of expression, color, shape, and dimensionality on the affective quality of game characters. *Learning and instruction*, *70*, 101194.
- Rodríguez-Ferrer, J. M., Manzano-León, A., Fernández-Jiménez, C., Luque de la Rosa, A., Fernández-Campoy, J. M., & Aguilar-Parra, J. M. (2023). Shall we play together? Game-based learning for engagement and classroom climate in Spanish socially deprived communities. *Frontiers in Psychology*, *14*, 1163441.

## Developing Computational Thinking: Using TurtleStitch and Physical Computing

#### Peter SEOW <sup>a\*</sup>, Chin-Lee KER<sup>a</sup> & Haoran XUE<sup>b</sup>

<sup>a</sup>National Institute of Education, Nanyang Technological University, Singapore <sup>b</sup>New Life Community Services, Singapore \*peter.seow@nie.edu.sg

**Abstract:** In this paper, we describe how the use of physical computing devices like an embroidery machine and TurtleStitch software can be used to engage learners in coding and developing their Computational Thinking (CT) skills. Two lessons are described on how novice learners can create geometric patterns with code while applying CT skills. In the first lesson, students learn to create polygon shapes such as square, triangle, pentagon, and hexagon. Through abstraction, pattern recognition, and algorithmic thinking, they must develop a modularized code block to create a polygon with sides and length of the polygon as input parameters. In the second lesson, they must create pinwheel geometric patterns through decomposition, pattern recognition and algorithmic thinking. The accompanying approaches such as tinkering, creating, debugging persevering, and collaborating were used to develop CT skills as learners generate geometric patterns with block-based codes in TurtleStitch. The use of embroidery machines and coding on TurtleStitch can provide opportunities for novice learners to develop coding and computation thinking skills as they creatively generate patterns with codes that can be embroidered in a tangible textile form.

**Keywords:** Coding, Computational Thinking, Physical Computing, Embroidery, Novice Leaners

#### 1. Introduction

Embroidery machines have been used for centuries to create intricate and beautiful designs on fabric. However, with the advent of physical computing and software, these machines can now be used to design patterns with code. We explore the use of embroidery machines as a physical computing device with TurtleStitch to develop computational thinking skills such as abstraction, decomposition, pattern recognition, and algorithmic thinking. We present two lessons on how novice learners of coding can develop Computational Thinking skills using the TurtleStitch to create geometric patterns. In the first lesson, learners can learn about algorithmic thinking, pattern recognition and abstraction to create a function for creating a nsided polygon. In the second lesson, learners use decomposition, pattern recognition and algorithmic thinking to create pinwheel patterns. We used the approaches of having learners explore and experiment the codes through tinkering, creating their own pattern designs, find out errors in the code through debugging, preserving through their experienced failures, and collaborating with others in the process of learning coding. The generated pattern designed by the learner can be used to create an embroidery on a textile with the embroidery machine. The purpose of this paper is to share the potential of using embroidery machines as a tool with software like TurtleStitch to develop computational thinking in novice learners.

#### 2. Background

#### 2.1 Physical Computing

Physical computing is a recent growth area in computing education that involves the use of the hardware and software to build interactive physical systems that sense and respond to

the real world (O'Sullivan & Igoe, 2004). It has been shown to result in the broad engagement of users in learning and address shortcomings in established approaches (Sentence, Waite, Hodges, MacLeod & Yeomans, 2017). An example of a tool used for physical computing in the classroom is the popular BBC micro:bit (Cápay, & Klimová, 2019). Physical computing can provide an engaging and hands-on experience to learning computing, and can help simulate digital creativity, collaboration, and broader technology skills such as the use of sensors alongside coding and computational thinking.

A physical computing system uses sensors and microcontroller to translate analogy signals to a software system, and control electro-mechanical devices such as motors, lighting, servos, and other hardware devices. Many of the technologies that we experienced today are examples of physical computing such as the automatic lighting systems, voice-activated home control, or robot vacuum cleaner. Another class of physical computing examples are those that enable the creation of physical or tangible objects using the hardware and software. This is commonly found in computer-aided manufacturing (CAM) systems where pre-programmed computer software dictates the movement of factory tools and machinery such as robotic arm on an assembly line or Computer Numerical Control (CNC) machines. Closer to home, some sewing embroidery machines can function as embroidery patterns that can be used by the machine to create the embroidery. This approach could offer students a creative opportunity to learn coding and apply computational skills to create embroidery on sewing machines. Also, learners can see the application of codes and computers to creation of objects in real-world contexts.

#### 2.2 TurtleStitch

TurtleStitch is a free and open-source platform for generating and sharing patterns for embroidery machines (Wolz, Auschauer, Mayr-Stadler, 2019). It is derived from the educational block-based programming platform Snap!(Harvey et al., 2013). While TurtleStitch itself is not physical computing, it can be used in conjunction with physical computing devices such as embroidery machines or electronic paper cutters to create physical objects based on the designs generated by the program. In this sense, TurtleStitch can be considered a part of the physical computing process. TurtleStitch bridges the programming with textiles, enabling the creation of stitched patterns of materials, for novice and experienced coders. TurtleStitch is built upon the ideas of Seymour Papert's (1999) Logo where the microworld is the controlling the turtle. In TurtleStitch, the turtle represents the needle of the embroidery machine where it can be given instructions to create novel patterns for embroidery. By using TurtleStitch, novice learners can develop computational thinking (Wing, 2008) skills by working through the process of creating designs for embroidery machines. This involves breaking down complex designs into smaller, more manageable parts (decomposition), looking for similarities and patterns in the designs (pattern recognition), focusing on the important information to make generalisations (abstraction), and developing step-by-step instructions for the embroidery machine to follow (algorithms). Through this process, learners can develop their computational thinking skills while creating objects that are tangible and beautiful. After creating the patterns on TurtleStitch, the design can be exported to the embroidering sewing machine to create the embroidery on textile.



Fig 1. Shows the interface of the TurtleStitch programming platform.

#### 3. TurtleStitch Lesson

In the following sub-sections, we describe the lessons that we designed and enacted to introduce novice learners to coding with the TurtleStitch and applying Computational Thinking in the process of generating their geometric patterns with codes.

#### 3.1 Creating simple polygons to N-sided polygons

After introduction to the interface and basics of using TurtleStitch, learners code to create simple polygons such as a square, triangle, pentagon, and hexagon. We used Parson's puzzle approach by providing them the blocks to use to create a square. Along with the blocks, we provided prompts for them to think about how the turtle will move to create a square (see fig 2). Learners would need to put the blocks together and fill in the appropriate values such as the repeat loops and degrees to turn to create a square.



Create a square

Fig 2. Creating a square on TurtleStitch

Next, learners modify the code to create the triangle, pentagon, and hexagon. To guide the learners, prompts were provided on how many turns the turtle will make or angle the turtle needs to turn to create the polygons. After creating the polygons, learners are to observe the patterns in the codes and created polygons (see fig. 3). They would notice the number of repeats corresponded to the number of turns or sides of the polygon and turn degrees was 360/number of turns. Also, they observed the number of sides was the important in creating the polygon in the codes. This is a form of abstraction by removing the details and focusing on the essential information.

# What is common? Triangle Square Pentagon Hexagon Triangle International Internationa



Fig 3. Creating different polygons

Based on their observations of patterns in the codes and the polygon sides, learners must create an n-sided polygon given the number of sides using the Block feature in TurtleStitch. Creating the block encapsulates the details of creating the polygon which is a form of abstraction of hiding the details. Learners are guided to create a Block with input arguments on the number of the sides and length of polygon (See Fig. 4).

Block Editor		draw text hello with size 21	
Polygon Sides Len repeat Sides move Len steps turn (* 1000 / Sides) degrees		text length of hello with size 21 if on edge, bounce x position y position direction	Polygon 8 80
OK Apply Cancel	li.	Make a block	Input your values – sides and length

Fig 3. Creating a block for creating N-sided polygons

From the process of applying Computational Thinking in pattern recognition, learners can see the relationship between the codes, created shapes and values to identify the important information by abstraction to develop a generalized code to create polygons. In generating the polygons, they would have to use algorithmic thinking in sequencing the blocks, using the loop statements and appropriate values such as the angles to turn.

#### 3.2 Pinwheel Patterns

The pinwheel is commonly used pattern in TurtleStitch to create embroidery. Based on the previous lesson on creating polygons with codes, learners are introduced to the creation of pinwheel patterns. Given several pinwheel patterns, learners are asked to decompose and recognize polygon patterns (Fig 5).

### Can you create these patterns?

What basic shapes do you see?



Fig 5. Decomposing and pattern recognition in pinwheels

After they can identify the basic polygon, the learners think about how the pinwheel pattern is formed with the polygon shape and the codes required to create. Applying decomposition and algorithmic thinking the learners code to create the pinwheel patterns.

#### 4. Approaches to developing Computational Thinking

In developing novice leaners' Computational Thinking skills in algorithmic thinking, pattern recognition, abstraction, and decomposition, it is important to consider the approaches to support the development. In our lessons, we adopted the approaches of tinkering, creating, debugging, persevering, and collaborating advocated by UK's Barefoot Computing (UK Stem Learning, 2023) Through tinkering, we encouraged learners to develop computational thinking and coding skills by exploring and experimenting codes. They could change parameter values and sequenced the blocks differently to see how their designs are changed. We encouraged creativity in our learners to create the designs that are meaningful to them and giving them choices. On many occasions, the codes do not produce the intended design and produce errors. Learners will debug their codes to find and rectify their errors. If not, they would have to create a solution as a work around. Debugging codes or creating complex designs require the learners to be persevere and keep going with their projects. We support the learners by working with them to solve problems in collaborative approach. Students experience challenges when their designs are embroidered on the textiles with the sewing machines. For example, their designs may have too many stitches tearing the textile or appear to be sparse with the wrong stitching. They would need to revise their codes for the design to be embroidered again. Fig 6a-c shows the processes where learners are engaged in learning coding, collaborating with peers, and creating the embroidery. From our lessons with novice coders, the associated processes of coding coupled with the use of physical computing plays an important role in the development of computational thinking.



Fig 6a. Coding on TurtleStitch



Fig 6b. Learner collaborating on the machine



Fig 6c. Creating embroidery on the machine
#### 5. Conclusion

In this paper we describe the use of TurtleStitch with an embroidery machine as a physical machine to develop computational thinking skills of pattern recognition, decomposition, algorithmic thinking, and abstraction. Through the design of lessons, a novice learner of coding can apply computational thinking to design and create geometric-based patterns such as polygons and pinwheels that can be embroidered on textiles on a machine. The approaches of tinkering, creating, debugging, persevering, and collaborating used in our lessons support the development of computational thinking in learners. Thus, it is important to consider the processes and experiences of learners that would enable to develop computational thinking.

#### Acknowledgements

The study was funded by SMF project "It takes a village 2: Mentoring through Tinkering." The authors would like to thank the TurtleStitch users' group for the many ideas in enabling creativity through coding and embroidery.

#### References

- Cápay, M., & Klimová, N. (2019). Engage Your Students via Physical Computing!. In 2019 IEEE Global Engineering Education Conference (EDUCON) (pp. 1216-1223). IEEE.
- Harvey, B., Garcia, D. D., Barnes, T., Titterton, N., Armendariz, D., Segars, L., Lemon, E., Morris, S., & Paley, J. (2013). Snap!(build your own blocks). In *Proceeding of the 44th ACM technical symposium on Computer science education* (pp. 759-759).
- O'Sullivan, D., & Igoe, T. (2004). *Physical computing: sensing and controlling the physical world with computers*. Course Technology Press.
- Papert, S. (1999). What is Logo? Who needs it. Logo philosophy and implementation, 4-16.
- Sentance, S., Waite, J., Hodges, S., MacLeod, E., & Yeomans, L. (2017). "Creating Cool Stuff" Pupils' Experience of the BBC micro: bit. In *Proceedings of the 2017 ACM SIGCSE technical symposium on computer science education* (pp. 531-536).

Papert, S. (1999). What is Logo? Who needs it. *Logo philosophy and implementation*, 4-16. UK Stem Learning (2013). *The Computational Thinkers.* 

https://www.stem.org.uk/resources/elibrary/resource/465617/barefoot-computing-informationand-posters#&gid=undefined&pid=5

- Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1881), 3717-3725.
- Wolz, U., Auschauer, M., & Mayr-Stalder, A. (2019). Code crafting with turtlestitch. In ACM SIGGRAPH 2019 Studio (pp. 1-2).Mackay, W. E. (1995). Ethics, lies and videotape. In I. R. Katz, R. Mark, L. Marks, M. B. Rosson, & J. Nielsen (Eds), Proceedings of CHI'95 (pp. 138-145). ACM Press.

## Applying cooperative learning with role division to learn recursion

#### YungYu ZHUANG<sup>\*</sup>, Hong-Wei CHEN, Jen-Hang WANG & Gwo-Dong CHEN

Department of Computer Science and Information Engineering, National Central University,

Taoyuan City, Taiwan \*yungyu@ncu.edu.tw

**Abstract:** Nowadays, most programming languages support multiple programming paradigms, enabling the choices between them for different application scenarios and even using them together, for example, object-oriented programming for system modeling, event-driven programming for graphical user interface, and functional programming for wrapping actions. However, learning different programming paradigms is quite difficult, especially when they are given as a whole. Recursion is a typical example since it originated in functional programming but was also included in most structured programming tend to think about the control flow of recursion and get confused by how the recursion can be implemented. To address this issue, we developed a methodology along with system software to help novices think of recursive solutions from the definition rather than the implementation.

**Keywords:** Programming, programming paradigm, recursion, cooperative learning, role division

#### 1. Introduction

Learning programming is learning how to appropriately use programming paradigms to implement software and organize program code (Floyd, 1979), for example, learning how to use objects, events, and functions. For a specific application scenario, programmers need to appropriately use the programming paradigms supported in the programming language they are using to implement software features. In other words, programming paradigms are the methodology to organize programs for specific application scenarios and the way to implement parts of a software system, which enables programmers to express their ideas with concrete program code. Without knowledge of programming paradigms, writing a good program to fit the system requirement is impossible.

Programming paradigms can be supported with either language constructs or design patterns. Language constructs are the essential elements in a programming language to form a programming paradigm, basically supported by reserved keywords and language syntax. For example, the "if-else" and "for" structured programming and "class" for object-oriented programming; programmers can easily use them for desired functionalities. On the other hand, design patterns are developed based on existing language constructs to achieve functionalities that are not directly supported by the language. In other words, following the developed patterns to write a piece of code to implement functionality such as "for" or "class." In languages that do not directly support event-driven programming, for example, C++ and Java, programmers can implement events with the Observer pattern (Gamma et al., 1993), while in a language like C#, programmers may simply use the language constructs named events and delegates.

Modern programming languages tend to support multiple programming paradigms to meet different programmers' needs and even allow programmers to use several paradigms together. For example, languages like C++ and Python support both procedural programming and object-oriented programming and allow programming with functional flavor. Scala supports object-oriented programming and functional programming. In fact, mainstream

languages in the industry, such as C language, also support functional-like things in addition to procedural programming, for example, recursion. Supporting multiple programming paradigms in a single programming language is a trend.

However, learning multiple programming paradigms at the same time is guite difficult. Learners may get confused about what programming paradigms actually are, how they are supported, and when they can be applied. We also found such a situation in classrooms where students have difficulty switching between programming paradigms and using them correctly. As Dijkstra mentioned, students usually struggle against learning recursion because they consider recursion based on their understanding of structured programming rather than from the viewpoint of definition (Diikstra, 1999). Once students start trying to think and implement recursion with structured programming, i.e., sequence, branch, and loop (Dahl, Dijkstra, & Hoare, 1972; Nassi & Shneiderman, 1973), they will get confused about the execution of every recursive step since recursion cannot be natively represented with structured programming and flowcharts. Although it is possible to convert a recursive call into an iteration with appropriate symbols and conditions, the representation will become iterative style---the recursive call will be represented by repetitive steps rather than a comprehensive step. Actually, students must recognize that recursion is a definition and use it without tracing each step. Therefore, how to help novices learn how to use programming paradigms is an important issue.

The observation led us to think about how to help students think and learn individual programming paradigms. So far as we know, most research activities on improving programming learning discuss programming as a whole or stay in discussing structured programming. In this paper, we take the recursion in functional programming as an example to point out how different programming paradigms and the problems learners may face when learning programming. To address this issue, we developed a methodology to help novices learn from the definition of recursion. We designed a system to let learners focus on the definition and benefit from cooperative learning with role division.

#### 2. Literature review

In this section, we discuss existing research activities on programming paradigms, cooperative learning, and recursion to clarify the research gap, i.e., the importance and difficulty of learning recursion.

#### 2.1 Programming paradigm

Programming paradigms are how to implement and organize program code, which often determines whether programmers use a language (Floyd, 1979). There have been many programming paradigms developed for resolving different application scenarios. For example, structured programming was suggested to improve the structure of program code (Dahl, Dijkstra, & Hoare, 1972). Procedural programming was discussed to divide program code into numbers of procedures or subroutines. Functional programming was developed to make code declarative and reduce side effects (Hughes, 1989). Object-oriented programming was proposed to model a system with message sending protocol (Goldberg & Robson, 1983). Aspect-oriented programming was developed to modularize crosscutting concerns (Kiczales et al., 1997). Reactive programming follows the idea of data-flow programming to simplify further the usage of events in programs (Zhuang, 2019). Programming paradigms play an essential role in practicing the separation of concerns, which lets programmers focus on implementing and modifying the code of a particular concern (Dijkstra, 1974). The support of the programming paradigm is the most crucial feature in programming language design.

#### 2.1.1 Structured programming and procedural programming

Structured programming is an early milestone in the history of programming languages. It is to encourage programmers to represent their algorithms with the three types of decomposition: concatenation (sequence), selection (branch), and repetition (loop) (Dahl, Dijkstra, & Hoare,

1972). Structured programming also encourages programmers to consider block and scope and represent them in the form of flowcharts (Chapin, 1970; Nassi & Shneiderman, 1973).

Procedural programming suggests considering a program as a series of procedure calls, which can be regarded as a further step toward modularization upon imperative programming and structured programming. It emphasizes the usage of modular procedures to express computational steps in the algorithm to implement. Nowadays, structured programming and procedural programming are supported in most imperative programming languages, such as C and C++.

#### 2.1.2 Functional programming

Functional programming has its origin in lambda calculus and was introduced in early programming languages for artificial intelligence, such as Lisp. It naturally supports recursion without transferring it into repetition and can easily represent recursive ideas in computer science. Although the mainstream languages in the industry nowadays are imperative, functional programming elements like recursion and lambda are partially introduced, i.e., programmers can write imperative programs with "functional flavor."

On the other hand, pure functional programming languages are continuously used in academia and now going popular. Several functional programming languages, such as Haskell and Scala, have a large community and are used in industry.

#### 2.2 Cooperative learning and functional roles

Cooperative learning has shown its effectiveness in education (Roger & Johnson, 1994; Slavin, 1980), including programming learning (Chu & Hwang, 2010; Hwang et al., 2012). Those research activities have reported cooperative learning can increase student achievement. With different kinds of cooperative learning approaches, students' learning performance may be improved (Sharan, 1980). In cooperative learning, positive interdependence can be considered a critical element (Roger & Johnson, 1994), which means students in a cooperative lesson have to ensure not only themselves but also group members can learn the assigned material well. In other words, for group success, all group members need to devote efforts and make their contributions to the joint effort. In such group training and learning, compared with classifying all group members as a single category, giving group members different functional roles is more effective for group growth and production (Benne & Sheats, 1948; Benne & Sheats, 2020).

#### 2.3 Recursion

Recursion is not only an important thinking method but also a representative approach used in computer science. Many important problems and algorithms in computer science are based on recursion, for example, the traversal of trees and graphs. Although the concept of recursion is quite natural, it is always a big challenge for novice programmers. One of the reasons might be that recursion has its root in functional languages like Lisp. At the same time, it was introduced in imperative languages like C. Dijkstra also observed that students who are familiar with Fortran tend to have difficulty in learning recursion since Fortran does not support recursion. Kahney concluded the problem of learning recursion comes from an incorrect understanding of recursion (Kahney, 1983). As Rinderknecht mentioned, many recursion teachings focus on control flow rather than definition (Rinderknecht, 2014).

Although the difficulty of learning recursion is not a new issue, there were rarely research activities on improving it from the viewpoint of definition. Recent research, RecurTutor, emphasized the necessity of practice and showed convincing experimental results (Hamouda et al., 2018). However, according to the findings in the above research results, how to afford a better understanding of recursion from the viewpoint of definition should be more critical. This observation motivated us to design a different learning method for understanding recursive programming.

#### 3. Methodology

To help novices learn the concepts in different programming paradigms, such as recursion, we propose applying cooperative learning with role division to programming learning. We developed a methodology along with a concrete system to let learners focus on the definition of recursion without thinking about the elements in other programming paradigms.

#### 3.1 The definition of recursion

The recursive solution to a problem is to consider and resolve a smaller instance of the same problem (Roberts, 1986). In programming, recursion is performed by letting functions call themselves with smaller data sizes, i.e., reducing the problem size to resolve. Once the functions are called with the base cases that cannot be reduced anymore, the problem can be easily resolved. Therefore, we can define a recursion solution to a given problem as the following three elements: base case, reduction, and rewrite.

Here we use the example of calculating factorial numbers in mathematics to explain the three elements. As a frequently used number in mathematics, the factorial of a nonnegative integer n is denoted by n!, which is the product of all positive integers less than or equal to n. For example, factorial three can be calculated as follows:

3! = 3 \* 2 \* 1 And actually, it is equal to 3 \* (3-1)!: 3! = 3 \* 2!

For the problem of calculating the factorial number of a given n, we can consider the three elements for the recursive solution based on its definition. According to the definition of factorial numbers, the simplest case without any calculation is the case of zero, and the factorial number of zero is one:

0! = 1

Hence, we can set the base cases of our recursive solution to them, try to reduce the given n to one or zero, and rewrite the rules to combine the calculation as shown in Table 1. Suppose we use a function named fact(n) to represent the calculation of factorial n. The three steps of thinking the recursive solution will be considering the following three questions in order:

- 1. How to handle the base case(s) for the given problem?
- 2. Can the problem be reduced to the same problem(s) with a smaller problem size?
- 3. How to rewrite the problem in the form of the same problem(s) with a smaller problem size?

Element	Meaning	Example
Base case	find out the extreme cases of the problem	0! = 1
	that can be simply calculated or got	
Reduction	move one step forward by reducing the	(3 - 1)!
	problem size	
Rewrite	rewrite the given problem statement with a	3! = 3 * 2!
	combination of smaller ones	

Table 1. The three elements in the recursion definition

#### 3.2 Division of roles

In order to let learners focus their attention on the recursion definition itself, the task of thinking of a recursive solution is allotted to three roles: base case finder, problem size reducer, and problem statement rewriter, as shown in Figure 1. By dividing the task of thinking of a recursive solution into three parts, each learner can concentrate on only the assigned part without considering the whole problem once. The role division can relieve the difficulty and help learners start by thinking about a part of the problem.



Figure 1. The role division is based on the three elements in the recursion definition.

#### 3.3 Learning design

The learning design of our methodology is shown in Figure 2. Students in a classroom are first grouped to learn a given problem, where three students per group. After the teaching activities given by the lecturer, students start to think about a part of the problem based on the roles assigned to them, i.e., base case finder, problem size reducer, and problem statement rewriter. When students finish their work, our system will compose their thinking result to compose the solution to the given problem. After completing the given problem, students need to exchange their roles for the next problem to learn how to think about different parts of the problem. Such a design is to help learners think of the solution according to the definition without thinking about control flow.



Figure 2. The learning design of our methodology.

#### 3.4 System implementation

Based on our proposed learning design, we developed a learning system software to support learners. The system is implemented in C# programming language with TCP/IP socket connection. Each group member can launch a client on the computer individually and switch to the mode for the assigned role. Node is and WPF (Windows Presentation Foundation) frameworks are used for the server and client user interface. The object language generated by the system is Python, which will be shown on the client user interface to help learners verify their thinking results.

#### 4. Conclusions and future work

We took recursion as an example to discuss the problem novices may face when they learn various programming paradigms. A methodology using cooperative learning with role division was developed along with a concrete system to help learners focus on the definition of recursion. We have conducted an experiment based on our methodology and system interface compared with learning in an integrated development environment. We are analyzing the experimental results to give a detailed report in a journal paper. In the future, we plan to extend our system to support more complicated recursive problems such as backtracking usage. Currently, our system can only cover typical recursive problems due to the implementation of our system. Showing the effectiveness of our methodology in learning other programming paradigms is also included in our future work.

#### Acknowledgements

This work was supported by the National Science and Technology Council, Taiwan [Grant No.: MOST 110-2511-H-008-003-MY2].

#### References

- Benne, K. D., & Sheats, P. (1948). Functional roles of group members. *Journal of social issues*, 4(2), 41-49. <u>https://doi.org/10.1111/j.1540-4560.1948.tb01783.x</u>
- Benne, K. D., & Sheats, P. (2020). Functional roles of group members. In *Shared Experiences in Human Communication* (pp. 155-163). Routledge.
- Chapin, N. (1970). Flowcharting with the ANSI standard: A tutorial. *Acm computing surveys (csur)*, 2(2), 119-146.
- Chu, H.-C., & Hwang, G.-J. (2010). Development of a project-based cooperative learning environment for computer programming courses. *International Journal of Innovation and Learning*, 8(3), 256-266. <u>https://doi.org/10.1504/IJIL.2010.035029</u>
- Dahl, O.-J., Dijkstra, E. W., & Hoare, C. A. R. (1972). *Structured programming*. Academic Press Ltd. Dijkstra, E. W. (1974). *E.W. Dijkstra Archive: On the role of scientific thought (EWD447)*. Retrieved
- Dec 23 from https://www.cs.utexas.edu/users/EWD/transcriptions/EWD04xx/EWD447.html
- Dijkstra, E. W. (1999). É.W. Dijkstra Archive: Computing Science: Achievements and Challenges (EWD1284). Retrieved Dec 30 from
- https://www.cs.utexas.edu/users/EWD/transcriptions/EWD12xx/EWD1284.html Floyd, R. W. (1979). The paradigms of programming. *Commun. ACM*, 22(8), 455–460. https://doi.org/10.1145/359138.359140
- Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1993). Design patterns: Abstraction and reuse of object-oriented design. European Conference on Object-Oriented Programming,
- Goldberg, A., & Robson, D. (1983). *Smalltalk-80: the language and its implementation*. Addison-Wesley Longman Publishing Co., Inc.
- Hamouda, S., Edwards, S. H., Elmongui, H. G., Ernst, J. V., & Shaffer, C. A. (2018). RecurTutor: An interactive tutorial for learning recursion. *ACM Transactions on Computing Education (TOCE)*, *19*(1), 1-25.
- Hughes, J. (1989). Why Functional Programming Matters. *The Computer Journal*, 32(2), 98-107. https://doi.org/10.1093/comjnl/32.2.98
- Hwang, W.-Y., Shadiev, R., Wang, C.-Y., & Huang, Z.-H. (2012). A pilot study of cooperative programming learning behavior and its relationship with students' learning performance. *Computers & Education*, *58*(4), 1267-1281. https://doi.org/10.1016/j.compedu.2011.12.009
- Kahney, H. (1983). What do novice programmers know about recursion. Proceedings of the SIGCHI conference on Human Factors in Computing Systems,
- Kiczales, G., Lamping, J., Mendhekar, A., Maeda, C., Lopes, C., Loingtier, J.-M., & Irwin, J. (1997). Aspect-oriented programming. European conference on object-oriented programming,
- Nassi, I., & Shneiderman, B. (1973). Flowchart techniques for structured programming. *SIGPLAN Not.*, 8(8), 12–26. <u>https://doi.org/10.1145/953349.953350</u>
- Rinderknecht, C. (2014). A survey on teaching and learning recursive programming. *Informatics in Education-An International Journal*, *13*(1), 87-120.
- Roberts, E. (1986). Thinking recursively. J. Wiley.
- Roger, T., & Johnson, D. W. (1994). An overview of cooperative learning. *Creativity and collaborative learning*, 1-21.
- Sharan, S. (1980). Cooperative Learning in Small Groups: Recent Methods and Effects on Achievement, Attitudes, and Ethnic Relations. *Review of Educational Research*, 50(2), 241-271. <u>https://doi.org/10.3102/00346543050002241</u>
- Slavin, R. E. (1980). Cooperative Learning. *Review of Educational Research*, 50(2), 315-342. https://doi.org/10.3102/00346543050002315
- Zhuang, Y. (2019). A lightweight push-pull mechanism for implicitly using signals in imperative programming. *Journal of Computer Languages*, *54*, 100903. <u>https://doi.org/https://doi.org/10.1016/j.cola.2019.100903</u>

## Designing a Scaffolding Inquiry-based instruction to Promote Non-engineering students in STEM Learning

#### Chia-Jung CHANG<sup>\*</sup>

Department of Information Communication, Yuan Ze University, Taiwan \* ccj@saturn.yzu.edu.tw

**Abstract:** The purpose of this study is to design a scaffolding inquiry-based instruction to facilitate non-engineering students in STEM learning and investigates its impact on students' computer programming, science learning and problem-solving performance. Employing a pretest-posttest experimental design, the investigation involved 19 university students who participated in a 3-week scaffolding inquiry-based program. The results showed that the instruction have positive effects on non-engineering students' logical thinking, debug, and control in computer programming and concept, high-level cognition, practice, apply and communication in science learning. Although the instruction improved students problem-solving performance, some students failed to solve complicated problems.

Keywords: scaffolding, programming, non-engineering, STEM

#### 1. Introduction

Numerous research and educators have paid much attention on STEM education for nurturing students' various competencies. Within STEM activities, students are frequently required to design products or propose solutions to solve problems by integrating different disciplines. The process facilities students' conceptual understanding and problem analysis by the experience of applying and integrating knowledge and skills from disciplines (Dasgupta, Magana, & Vieira, 2019). Hence, STEM has been extensively implemented in varied educational contexts.

Previous studies have indicated that STEM education contributes to the enhancement of problem-solving and computer programming competencies. For example, Lin, Wang, and Wu (2019) designed a modeling-based physics programming instruction for STEM learning, indicating that the modelling process not only enhanced students' programming abilities and physics concepts, but also helped them program solutions to solve problems. Additionally, other studies indicated the experiencing in STEM activities could fosters students' problemsolving competency (Chen et al., 2020).

Recently, a growing body of research has shifted its focus towards non-engineering students within STEM learning. (Hu, Yeh, & Chen, 2020; Lin, Yu, Shih, & Wu, 2021). For instance, Hu et al. (2020) investigated the effect of hands-on activity for non-engineering students on STEM learning. They found that the activity enhanced students' learning performance and attitude by experiencing the problem-solving process. Another similar study by Lin et al. (2021) explored the AI literacy of non-engineering students after participating in STEM-based AI hands-on activity. The results revealed that STEM hands-on activity can improve non-engineering students' AI literacy. However, it is relatively challenging for non-engineering students who lacked learning experience in STEM-related subjects in the past. Previous studies indicated that non-engineering students have a lower level of engineering skills and computer literacy than engineering students (Mansoor & Ahmad, 2019). Moreover, non-engineering students often fail to perform STEM activities, which may produce a negative learning attitude toward STEM learning (Lo et al., 2017). In other words, STEM learning does

not always guarantee effective learning for non-engineering students (Lin et al., 2021). Hence, how to help non-engineering students with STEM learning is an important educational issue.

Although STEM problem-solving activities benefit students, some studies found potential learning problems during students try to solve problems. Tan et al. (2023) indicated that most students encounter problems related to the application of disciplinary knowledge and practice when taking part in STEM problem-solving activities. More specifically, students often applied low-level scientific concepts to solve problems. The same problem is also found in the study by Chang and Tao (2021), indicating that even with inquiry-based instruction, students still failed to use high-level conceptions of learning science in STEM activity. Therefore, they suggested that an intended connection to reduce the gap between science and practice is helpful for students in generating solutions and enhancing STEM competencies.

To help non-engineering students engage in STEM learning, the study developed a scaffolding inquiry-based instruction in STEM learning and investigated the impact of the instruction on non-engineering students' performance in computer programming, science learning, and problem-solving. More specifically, research questions are presented as follows.

- 1. Does the scaffolding inquiry-based instruction facilitate non-engineering students' computer programming self-efficacy?
- 2. Does the scaffolding inquiry-based instruction improve non-engineering students' science learning self-efficacy?
- 3. What are non-engineering students' problem-solving performance in the scaffolding inquiry-based STEM?

#### 2. Method

#### 2.1 Participants

A pilot study using a pretest-posttest experimental design was conducted to explore the effect of the scaffolding inquiry-based instruction on STEM learning. The participants of this study were 19 university non-engineering students who were designs backgrounds (16 females and 3 males). They have no experience in making a game and taking part in STEM education. Hence, these students are suitable as subjects.

#### 2.2 The scaffolding inquiry-based learning activity

The study designed a game-making learning activity based on an inquiry-based process to guide students on the procedural of game development using programming and provided scaffolding questions to help students establish a connection between physics and computer programming. The activity was created on CoSci (<u>https://cosci.tw</u>), developed by Prof. Liu, and used an inquiry process as an instructional framework to engage students in STEM learning. The activity consisted of two parts, including a training in programming and a problem-solving task. The former aimed to teach students the operation of the game design interface, basic concepts of programming, and physics concepts by a sequential inquiry process, including playing, exploration, explanation, practice, and reflection phases.

Except for the playing phases, every phase provided scaffolding questions to guide students in understanding the required tasks. For instance, during the exploration phase, students were required to observe the data sheet provided by the platform and to respond to questions regarding variables that needed to be considered for aircraft/bomb movement. At the explanation phase, the students were required to explain the game object movement using physics concepts. Moreover, students needed to explain the relationships among variables from the previous phase. These questions aimed to help students identify the role of each variable in the game and to explore the relationship among the variables by explaining physics concepts. It was expected that the support of the scaffolding question facilitated the integration and application of disciplinary knowledge. The last phase, the activity provided reflective questions to prompt students evaluate their learning performance and monitor the errors they made while designing a mini game.

Since students lacked computer programming, physics, mathematics and problemsolving skills, an aircraft shooting a tank mini game as a demonstration to help students understand the concept of programming and free fall motion concept. More specifically, the demonstration guided students in establishing the relationships among the variables and physics concepts, to control the movement of the game objects, to create interaction with keyboard input, and handling collision detection. Additionally, there are practice and reflection sessions that provided learning opportunities to reinforce the application and integration of physics and programming. Finally, students were required to complete a problem-solving game task. Differentiating from previous demonstration game, the task was a bombing mission where a player controlled a tank to dodges enemy aircraft attacks and simultaneously shot an enemy aircraft. In the task, there are four problems, including rendering game objects, objects controls, object movement, collision detection.

#### 2.3 Procedure

Before the activity, two questionnaires, including computer programming self-efficacy scale (CPSES) and science learning self-efficacy (SLSE), were used for 30 minutes as a pretest. The scaffolding inquiry-based learning activity lasted for nine 50-min sessions over three weeks. The training was implemented in six sessions and the task took three sessions. After the activity, the same two questionnaires were employed as a posttest.

#### 2.4 Data collection and analysis

The computer programming self-efficacy scale developed by Tsai, Wang, and Hsu (2018) was used to measure students' perception of programming learning before and after the activity. Additionally, the study used the science learning self-efficacy scale (Lin & Tsai, 2013) to investigate their perceptions of science learning. The two questionnaires were slightly revised to align with the learning activity context. The overall Cronbach's alpha of the two revised questionnaires were .93 and .94, respectively, indicating that the questionnaires were sufficiently reliable. The responses to the questionnaires were analyzed by a paired t-test to understand the changes before and after the activity. Besides, students' performance in the task was evaluated. The task consisted of four problems, each worth 4 points, for a total of 16 points.

#### 3. Results

Table 1 showed the result of students' computer programming self-efficacy scale using a paired t-test. The results indicated that compared with before the activity, students perceived a higher level of computer programming skills, including the dimensions of logical thinking (t=3.16, p=.005<0.01), debug (t=-3.13, p=0.006<.01), and control (t=-3.50, p<.01). However, the algorithm was no statistically significant difference between before and after the activity. In other words, students perceived a similar level of algorithm before and after the activity. The results suggested that the activity may be helpful for non-engineering students in condition statements, revised programming problems and control of programming interface.

 Table 1. The result of students' computer programming self-efficacy before and after the scaffolding problem-based activity

Dimensions	Pre	etest	Posttest		t-value	p-value
	Mean	SD	Mean	SD	_	
Logical thinking	3.66	0.94	4.72	0.95	-3.16**	0.005
Algorithm	2.98	0.69	3.44	1.29	-1.65	0.117

Debug	3.46	0.75	4.35	1.09	-3.13**	0.006
Control	3.47	1.11	4.61	1.15	-3.50**	0.003

\*<.05, \*\*<.01, \*\*\*<.001

The result of students' science learning self-efficacy using a paired t-test was shown in Table 2. Significant differences were found in all the dimensions of science learning selfefficacy. More specifically, compared before the activity, students believed that they obtained a higher level of understanding of science concept (t=-3.48, p<.01), and could explain a scientific phenomenon using science theory (t=-3.21, p<.01). In addition, they knew how to set up variables and collect data during executing game (t=-3.98, t<.01) and were able to apply science in the game design (t=-7.36, p<.00). Moreover, they were more capable of expressing their ideas properly and clearly using science (t=-8.53, p<.00). These results suggested that the activity have a positive effect on science learning.

Table 2. The re	sult of students'	science	learning	self-efficacy	before	and after	the s	scaffolding
problem-based	activity							

Pretest		Posttest		t-value	p-value
Mean	SD	Mean	SD	_	-
2.93	0.83	3.33	0.67	-3.48**	0.003
2.87	0.87	3.65	0.56	-3.21**	0.005
3.18	0.93	4.05	0.66	-3.98**	0.001
3.06	0.68	3.74	0.46	-7.36***	0.000
2.67	0.77	4.29	0.54	-8.53***	0.000
	Pre Mean 2.93 2.87 3.18 3.06 2.67	Pretest           Mean         SD           2.93         0.83           2.87         0.87           3.18         0.93           3.06         0.68           2.67         0.77	PretestPostMeanSDMean2.930.833.332.870.873.653.180.934.053.060.683.742.670.774.29	PretestPosttestMeanSDMeanSD2.930.833.330.672.870.873.650.563.180.934.050.663.060.683.740.462.670.774.290.54	PretestPosttestt-valueMeanSDMeanSD2.930.833.330.67-3.48**2.870.873.650.56-3.21**3.180.934.050.66-3.98**3.060.683.740.46-7.36***2.670.774.290.54-8.53***

\*<.05, \*\*<.01, \*\*\*<.001



Figure 1. The results of students' problem-solving performance

Figure 1 showed the outcomes of problem-solving performance that students completed in the task. More than 80% of students could successfully render game objects in appropriated positions on the game scene and controlled objects using input devices. Approximately 75% of the students could correctly display object movement, while the rest of the students only displayed a partial correct movement and failed to move any objects (15%). Regarding collision detection, no students fully completed the four types of collision detection. Approximately 40% of the students partially achieved the task, while the rest only achieved one or two types of collision detections.

#### 4. Conclusion

The study designed scaffolding inquiry-based learning instruction for facilitating nonengineering students in STEM learning. The results showed that the instruction contributed not only to enhancing students' computer programming, but also improving their science learning self-efficacy. However, it is worth noting that most students had difficulties in integrating programming and physics. The findings of the study serve as references for researchers and instructors in integrating scaffolding inquiry learning into STEM learning. However, there were some limitations in this study. Due to the small sample size in the study, the results may not be generalized. Another limitation was the obtained results based on the questionnaire and programming outcome. Moreover, the activity is a short-term program that lasted for three weeks. In future studies, it is suggested to verify this study with a large sample size and a long-term program, and to evaluate students' programming process to identify the potential learning difficulties.

#### Acknowledgements

This research was partially funded by the National Science and Technology Council, Taiwan under contract numbers MOST 110-2511-H-155-002- and MOE Teaching Practice Research Program (PEE1123014).

#### References

- Chang, C.-J., & Tao, S.-Y. (2021). The Impact of Inquiry-based Integrated STEM on Student's Perception of Learning Science and Computer Programming. International Conference on Computers in Education.
- Chen, J. C., Huang, Y., Lin, K. Y., Chang, Y. S., Lin, H. C., Lin, C. Y., & Hsiao, H. S. (2020). Developing a hands-on activity using virtual reality to help students learn by doing. *Journal of Computer Assisted Learning*, *36*(1), 46-60.
- Delahunty, T., Seery, N., & Lynch, R. (2020). Exploring problem conceptualization and performance in STEM problem solving contexts. *Instructional Science*, *48*(4), 395-425.
- Ge, X., & Land, S. M. (2003). Scaffolding students' problem-solving processes in an ill-structured task using question prompts and peer interactions. *Educational technology research and development*, *51*(1), 21-38.
- Hu, C. C., Yeh, H. C., & Chen, N. S. (2020). Enhancing STEM competence by making electronic musical pencil for non-engineering students. *Computers & Education*, *150*, 103840.
- Lin, C. H., Yu, C. C., Shih, P. K., & Wu, L. Y. (2021). STEM based artificial intelligence learning in general education for non-engineering undergraduate students. *Educational Technology & Society*, *24*(3), 224-237.
- Lin, T. J., & Tsai, C. C. (2013). A MULTI-DIMENSIONAL INSTRUMENT FOR EVALUATING TAIWANESE HIGH SCHOOL STUDENTS'SCIENCE LEARNING SELF-EFFICACY IN RELATION TO THEIR APPROACHES TO LEARNING SCIENCE. International Journal of Science and Mathematics Education, 11, 1275-1301.
- Lin, Y. T., Wang, M. T., & Wu, C. C. (2019). Design and implementation of interdisciplinary STEM instruction: Teaching programming by computational physics. *The Asia-Pacific Education Researcher, 28*, 77-91.
- Lo, K. W., Lau, C. K., Chan, S. C., & Ngai, G. (2017, October). When non-engineering students work on an international service-learning engineering project—A case study. In 2017 IEEE Global Humanitarian Technology Conference (GHTC) (pp. 1-7). IEEE.
- Mansoor, T., & Ahmad, M. S. (2019). Learning for Youth Leadership: the Case of Hong Kong, Daniel TL Shek. Grace Ngai & Stephen CF Chan, Service (Eds.), Springer: Singapore, 2019, ISBN: 978– 981–13-0447-7.
- Tan, A. L., Ong, Y. S., Ng, Y. S., & Tan, J. H. J. (2023). STEM problem solving: Inquiry, concepts, and reasoning. *Science & Education*, *32*(2), 381-397.
- Tsai, M. J., Wang, C. Y., & Hsu, P. F. (2019). Developing the computer programming self-efficacy scale for computer literacy education. *Journal of Educational Computing Research*, *56*(8), 1345-1360.

## Explainable AI in the Real World: Challenges and Opportunities

#### Dora HORVAT<sup>a</sup>, Ivica BOTICKI<sup>a\*</sup>, Peter SEOW<sup>b</sup> & Antun DROBNJAK<sup>a</sup>

<sup>a</sup>Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia <sup>b</sup>National Institute of Education, Nanyang Technological University, Singapore \*ivica.boticki@fer.hr

**Abstract:** This paper presents the results of a systematic review of the research papers on the use of explainable AI in the real world. The present body of research indicates there is a huge drive from the academic society in pushing and exploring explainable AI across disciplines from a research perspective, and there is inherent need to design prototypes with increased complexity to tackle the numerous scientific and methodological issues in the process. The main conclusions of the review are that there exist serious methodological issues with the use of XAI in complex systems which reside on vast or layered information systems spanning across multiple organizational units with important data sometimes missing, potentially limiting the validity of the XAI approach used in practice. For XAI to work in the real-world context of education, the approaches to presenting explanations to the stakeholders such as teachers and students should be understandable by them to take appropriate actions or decisions. This would highlight the need to study of human-computer interaction between AI and users that would lead to better transparency, trust and personalization.

Keywords: Explainable AI, industry applications, education implications

#### 1. Introduction

The main goal of the 1955 proposal for the Dartmouth Summer Research Project on Artificial Intelligence (AI), a workshop labeled as the birthplace of artificial intelligence was "...to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it." (McCarthy et al., 2006). Artificial intelligence was then defined as "making a machine behave in ways that would be called intelligent if a human were so behaving", which despite the plethora of definitions remains relevant today. Even though AI definitions today emphasize the broad reach of artificial intelligence, the formulation given by its creators still stands at its core.

Since its beginnings, AI has continuously been developing at a fast pace. In 1964, scientists at MIT developed ELIZA, one of the first language processing computing systems which was able to imitate a Rogerian therapist. In 1997, IBM's Deep Blue computer beat the world champion Gari Kasparov in chess, thus going down in history as the first computer to win a game against a world champion under tournament time conditions. A decade later, a yearly tournament for chess engines began, where engines compete against each other. Today, digital AI-based assistants such as Siri or Cortana are widely implemented and used in everyday life in language processing tasks.

In recent years in particular, the field of artificial intelligence has made major progress in almost all its main sub-areas, including computer vision, speech recognition, natural language processing, expert systems and decision making (Michael L. Littman, 2021). Al is in some form ingrained into all major fields of human work – finance, security, healthcare and medicine, criminal justice, transportation, marketing, telecommunications. With that, Al is no longer observed exclusively in the context of accuracy and model optimization; it is the

role of humans that becomes a crucial factor. As decisions made by intelligent systems are affecting human lives in areas such as medicine or law, the need for understanding how these decisions are furnished by AI methods becomes imperative. The notion of explainable artificial intelligence (XAI) in most cases implies more than just understanding the model. Depending on the area of application and the intended users, XAI is expected to achieve goals other than 'just explanations'. For example, when AI is used for decision-making, it is crucial to ensure the model is fair and unbiased; for systems that work with sensitive and personal data, privacy must be preserved; and an AI system must be trustworthy – the user has to be confident that a model will act as intended.

This paper critically examines the potential of XAI in the real world uses, mapping out the studies which move further from the proof-of-concept efforts in order to illustrate the potential for the application of XAI in the real world. By doing that, the study is a pioneering effort in the field.

#### 2. Background

Although explainability is the core of responsible and trustworthy AI, across multiple disciplines and areas (Guidotti et al., 2018), there is still no agreement in the literature as to what explainability actually is. When talking about explanations in AI, there is no one-sizefits-all solution – the complexity of the explanations and the amount of detail provided is entirely dependent on who the intended users are (Barredo Arrieta et al., 2020). Furthermore, since explainability is ultimately a human-agent interaction problem (Miller, 2019), the solution to explainable AI lies not just in 'more AI', but in considering multiple aspects of human understanding, drawing insights from psychology and social sciences (Barredo Arrieta et al., 2020). Although the uniform definition of explainability remains elusive, the term black-box problem is a defining aspect of XAI. The black-box problem is a well known phenomenon in the field of artificial intelligence and represents the main limitation of effectiveness and usage of machine and deep learning models. As pointed out in the DARPA's Explainable Artificial Intelligence (XAI) Program, there is an inherent tension between machine learning performance (predictive accuracy) and explainability; with often the highest performing methods (e.g., deep learning) being the least explainable, and the most explainable (e.g., decision trees) being less accurate (Gunning and Aha, 2019).

In XAI-related literature, terms explainability and interpretability are often used interchangeably (Tjoa and Guan, 2021). Much like with explainability, there is no general agreement on the definition of interpretability, as well as no clear distinction between the two. However, there exist domains in which they do not convey the same concepts (Ehsan et al., 2018). In general, interpretability refers to the ability to explain or present a model in terms understandable to humans. It denotes passive characteristics of a model – how easy it is for humans to make sense of and identify relations between model's inputs and outputs (Barredo Arrieta et al., 2020) (Došilović, Brčić and Hlupić, 2018). Even though explainability is related to interpretability, explainability remains is associated with internal logic behind an intelligent system. In contrast to interpretability, explainability denotes an active characteristic of a model – how deep human's understanding of a model's inner workings is (Linardatos, Papastefanopoulos and Kotsiantis, 2021).

When differentiating between the two main approaches to generating explanations, the literature makes a clear distinction between post-hoc explainability and intrinsic explainability. Post-hoc explainability refers to methods which are applied to models that are not interpretable by design to improve their interpretability (Barredo Arrieta et al., 2020). They do not directly explain the inner workings of the model, instead they offer 'approximate' explanations such as textual, visual, or example-based explanations. Intrinsic explainability refers to models which are interpretable on their own due to their simple structure, such as linear regression or tree-based models. However, these models, which often stand to as an

interpretable alternative to complex black-box models such as neural networks, are not unconditionally interpretable. Taking high-dimensionality or heavy pre-processing and feature engineering into account, they are often not more intrinsically interpretable than black-box models (Lipton, 2018).

#### 3. Methodology

Since explainable AI is an emerging concept in terms of applicable methods and techniques in machine and deep learning, the focus of this review was to get an insight into the state of explainable AI, i.e. – are explainable models and frameworks being implemented and validated and if so, to what degree. Since explainability is not an easily measured component of AI, studies evaluating explainability of the system on real users were an important point of interest. In addition, studies validating models on real scenarios were also taken into consideration. For the medical and healthcare field, clinicians or experts had to be involved in some way, either via design or through validation.

The search for relevant studies on real-life applications of XAI was conducted on Web of Science electronic database in the period between December 2021 and March 2022. The search query used is a general one in order to gather studies for a wide overview of explainable artificial intelligence in a range of domains, which is the main aim of this review. At the time of the search, the query returned 1777 results from Web of Science database Publication year after 2017 was selected as an additional filter in order to gain focus on the current state of applications of XAI.

The 1777 records from Web of Science were obtained for title and abstract screening using the inclusion and exclusion criteria agreed-upon during the inter-rater process. Application of inclusion and exclusion criteria resulted in the set of 163 studies on which the full-text screening was conducted. The full-text screening was more qualitative in contrast to the title and abstract screening, which had precisely defined inclusion and exclusion criteria. During the full-text screening, focus was on studies that implemented a usable explainable system and evaluated explainability of the system on human subjects through user studies. Following that, 144 papers were excluded, leaving 19 papers in the final list (Figure 1).



Figure 1. Screening the paper for final analysis

#### 4. Inclusion and exclusion criteria

Final inclusion and exclusion criteria are formed as shown in Table 1. Agreed upon inclusion and exclusion criteria between the two raters after the 2nd inter-rater.

Table 1. Agreed upon inclusion	and exclusion criteria	between the two	raters after the 2nd
	inter-rater		

Inclusion criteria	Exclusion criteria
<ul> <li>Include papers proposing XAI models or frameworks and:</li> </ul>	• Exclude reviews and overviews of XAI field and methods
<ul> <li>Demonstrating application and validating it on a real scenario or a simulated real scenario (use- cases) or</li> </ul>	<ul> <li>Exclude papers conducting surveys about explainability and trustworthiness of AI and XAI</li> </ul>
<ul> <li>Performing user-studies or user- evaluations</li> </ul>	<ul> <li>Exclude papers discussing design patterns and principles</li> </ul>
<ul> <li>Include papers about the use of XAI in healthcare, medical and biomedical field if:</li> </ul>	• Exclude papers discussing ethical issues, bias, transparency and trust in the context of AI and XAI
<ul> <li>Clinicians/experts are involved in the study design</li> </ul>	
<ul> <li>User-studies with clinicians/experts are performed</li> </ul>	

#### 5. Results

A total of 14 studies were included in this review (Table 2). The biggest proportion of studies belongs to the healthcare domain, as is the case in a systematic review of explainable AI application domains (Islam et al., 2022). Besides six studies from the healthcare domain, three studies describe the use of XAI in the domain of human-computer interaction in the form of collaborative games or human performance improvement; one study belongs to the telecommunications domain; one to the industry domain in the form of supply chain planning support; and one to the energy domain through recommendations for energy efficiency. Most of the studies are from 2020 to 2021, and they conduct user studies with 5 to 60 participants to evaluate their systems and approaches (except for (Irarrázaval et al., 2021)).

Gradient-boosted decision trees along with the Shapley values for generating explanations are used in two studies, (Chromik, 2021) and (Melançon et al., 2021). In (Deperlioglu et al., 2022), (Xu et al., 2021) and (Wang and An, 2021) neural networks are used as the machine learning model while CAM (along with Grad-CAM and DeconvNet) method is used for providing visual explanations. Studies (Xie et al., 2020) and (Sardianos et al., 2021) do not specify the inner workings of their systems in terms of machine learning and explainable models; the focus is on iterative design of the system by following user requirements and user feedback. In (Xie et al., 2020) a Clinical Decision Support System is constructed. The system has multiple features for generating explanations, such as producing contrastive examples, outputting probabilities and the ability to show the most significant observations.

In (Sardianos et al., 2021) a recommendation system for energy efficiency with explainable and persuasive recommendations is designed. In (Khodabandehloo, Riboni and Alimohammadi, 2021) decision tree learning algorithm is used as an intrinsically explained model: the decision tree is then parsed to produce natural language explanations. In (Samuel, Abdullah and Raj, 2021) granular computing is used to interpret SVM's classification and construct syllogisms which are then transformed into natural language explanations. In (Gao et al., 2020) a spatial-temporal causal And-Or graph (STC-AoG) is used as representation of a robot's knowledge and inferred user's mental state. By parsing this graph and applying the proposed explanation generation framework, the robot can generate explanations with an aim to correct sub-optimal human behavior in human-robot collaboration tasks. In (Das and Chernova, 2020) a Rationale Generating Algorithm is proposed; it produces rationales (natural language expressions) which aim to aid the user's decision-making process and consequently increase the user's understanding of the performed task. In (Ehsan et al., 2018) an encoder-decoder network is used to translate between state and action information and natural language rationalizations, which serve as rationalizations for describing agent behavior. In (Sabol et al., 2020) semantically explainable fuzzy classifier CFCMC explains the decision by giving a semantic explanation on the possibilities of misclassification, and visual explanations by showing the training sample most responsible for a given prediction as well as training samples from other, conflicting classes. And finally, in (Irarrázaval et al., 2021) data is first clustered into groups using an unsupervised learning approach, after which a CART algorithm is used to construct a decision tree from which a set of rules is extracted.

Author and year	Domain	ML model	Explainable method	User study participants
(Xie et al., 2020)	Healthcare	-	-	6
(Xu et al., 2021)	Healthcare	ANN	Grad-CAM, Guided Grad-CAM	9
(Deperlioglu et al., 2022)	Healthcare	CNN	CAM	15
(Samuel, Abdullah and Raj, 2021)	Healthcare	SVM	Granular Computing	5
(Khodabandehloo, Riboni and Alimohammadi, 2021)	Healthcare	Decision Tree	-	8
(Sabol et al., 2020)	Healthcare	Fuzzy Model	Cumulative Fuzzy Class Membership Criterion (CFCMC)	14
(Irarrázaval et al., 2021)	Telecomm unications	Clustering, CART	-	-
(Wang and An, 2021)	Education*	CNN	CAM, DeconvNet	30
(Gao et al., 2020)	Human- computer interaction	-	Rationalization	29
(Ehsan et al., 2018)	Human- computer interaction	Encoder- decoder NN	Rationalization	53

Table 2. The final set of selected studies after the application of the literature review steps

(Das and Chernova, 2020)	Human- computer interaction	-	Rationalization	60
(Chromik, 2021)	General*	XGBoost	SHAP	16
(Melançon et al., 2021)	Industry	XGBoost	SHAP	*
(Sardianos et al.,	Energy	-	Explainable	8
2021)			Recommendation	
(Westerski et al.,	Industry	ADM	Rationalization	-
2021)				
(ten Broeke et al.,	Healthcare	-	Explainable	15
2021)			Decisions	
(Chakraborti et al.,	IT	-	Explainable	-
2019)			Recommendation	
(Gonzalo-Cristóbal	Education	Monte Carlo +	Explainable	12
et al., 2021)		ANN	Recommendation	
(Mirchi et al., 2020)	Healthcare	SVM	Adaptive	50
			Explanation	

#### 6. Discussions

There is no doubt that Explainable AI presents one of the most challenging topics in contemporary research. With the rise of artificial intelligence, the importance of explaining its inner workings is ever so important across disciplines, with the healthcare domain being most prominent, especially in the context of medical diagnosis and decision-support tools (Islam et al., 2022; Tjoa & Guan, 2021). The prospects and opportunities for medical XAI are at first glance many: the improvement of medical diagnosis, help with allocation of resources, reduction of bias, further AI development and increased adoption. Nevertheless, practical implications of the use of explainable AI across disciplines remain and are therefore explored as part of the study presented in this review paper.

Designing explainable AI often comes with difficult design decisions lying at the intersection of AI technology and its practical applications. The reviewed studies indicate that the ambiguity of the term explainable presents a great challenge. The concept of explainability is abstract and must be observed in the context of the user and its environment. Consequently, explanations should not be regarded just as just a product, but a multidisciplinary process (Khosravi et al., 2022). To achieve this, the human must be put in the focus of the design of explainable systems. Just as in decision-support systems, the human-centric approach and consideration of context should be the focus when developing XAI platforms. Such an approach can be noted in (Mirchi et al., 2020), where the experts' opinion altered the importance of metrics (features) that was determined by the AI model, to ensure that the framework was in concordance with the current guidelines of neurosurgical education. In the design of the explanations, the authors followed several cognitive and learning models and theories, while also ensuring that the given feedback mimics real-life experiences through textual, audio and video-based instructions. This type of feedback is claimed to improve selfguided learning and develop responsibility, which is highlighted as highly beneficial (Winne, 2021). The claims about benefits and relevance of the framework are however not verified, as the conducted user study only validates the technical workings of the model.

Once the explainable AI designs and solutions are in place, they are trialed and tested in real-life environments, but the testing is done in a laboratory fashion. The whole experiential setup is arranged with maximum support given to the party applying XAI solutions, often neglecting the idiosyncrasies of the real world and oftentimes not addressing the realistic performance of XAI. The vast majority of screened papers from the medical field only report

results in the form of technical metrics such as accuracy and precision, whereas the explainable part of the machine learning system often appears to be implemented in order to formally justify the XAI label, without much thought or validation on whether these explanations are useful in real-life settings. This is also noted in (Liao & Varshney, 2021), where the "disconnect between technical XAI approaches and supporting user's end goals in usage context" is identified as one of the pitfalls of developing explainable models. The disconnect can be observed in two related aspects; one aspect refers to the lack of cross-disciplinary research and studies on users' needs and preferences. The other aspect refers to the lack of performance and relevance evaluation of XAI models in real-life settings. (Liao & Varshney, 2021) identify the absence of studies providing evidence that incorporation of explainable components in AI systems and solutions improves realistic user performance in judgment and decision making, while (Adadi & Berrada, 2018) report the lack of XAI models evaluation no assessment of their relevance to the user. Some studies warn that giving up predictive power in favor of transparency and explainability should be carefully considered and properly justified upon (Lipton, 2018).

In the education context, XAI aims to address concerns related to fairness, accountability, transparency and ethics in educational interventions supported by AI algorithms (Khosravi et al., 2022). XAI can benefit teachers by gaining a better understanding of how AI systems work and make decisions, which can help them to better integrate AI tools into their teaching practice. For students, XAI can personalize their learning experiences by providing explanations that are tailored to their individual needs and characteristics. Explainable AI can support student's self-regulation by providing transparency and interpretability of the predictions and recommendations, which can help students better understand their performance and take appropriate actions to improve it (Afzall et al., 2021). For XAI to work in the real-world context of education, the approaches to presenting explanations to the stakeholders such as teachers and students should be understandable by them to take appropriate actions or decisions. This would highlight the need to study of human-computer interaction between AI and users that would lead to better transparency, trust and personalization.

Serious methodological issues surface with the use of XAI in complex systems which reside on vast or layered information systems spanning across multiple organizational units, where data sharing is limited. Such inherent issues reflect on the quality of integration of XAI solutions heavily and raise the question of whether the totality of information contributing to the actual workings of the processes is well modeled and described. What is more, in some cases important data from certain section of a system might be totally missing, potentially limiting the validity of the XAI approach used in practice. (Melancon, 2021), in their cooperation with Michelin, points out the lack of data as a restricting challenge, due to the absence of a standard practice of archiving all data in detail. In the medical domain, despite continued improvements of electronic health records, data quality and availability still present an issue (Gerlings, 2022).

#### 7. Conclusions

The complexity of XAI solutions presents a hurdle to their seamless user adoption. This does not come as a surprise, since there is a huge drive from the academic society in pushing and exploring explainable AI across disciplines from a research perspective, and there is inherent need to design prototypes with increased complexity to tackle the numerous scientific and methodological issues in the process. (Westerski, 2021) noted that, in their case, the introduction of a framework which is significantly different from users' current habits had a negative impact on adoption of the system. To minimize those negative effects, users should be involved in the process of design and validation, with keeping in mind that the same models have very different results depending on organization profile.

Last but not the least, the issues of data collection processes and data privacy are tightly related to the implementation of XAI systems in real-life scenarios, as such systems heavily relying on a variety of data sources. Although the manipulation and processing of such data presents a challenge itself, data privacy remains an insurmountable issue, since explaining data inevitably leads to discovering certain sensitive bits and information. Ethical and trust issues regarding model transparency and the overall black-box problem of AI emerge across domains. As an example, the explanations are crucial for pedagogical effectiveness of a digital system, as well as gaining students' and teachers' trust in given decisions (Conati et al., 2018).

#### References

- Afzaal, M., Nouri, J., Zia, A., Papapetrou, P., Fors, U., Wu, Y., X, Li, & Weegar, R. (2021). Explainable Ai For Data-driven Feedback and Intelligent Action Recommendations To Support Students Selfregulation. Frontiers in Artificial Intelligence, (4). https://doi.org/10.3389/frai.2021.723447
- Barredo Arrieta, A. et al. (2020) "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," Information Fusion, 58. doi:10.1016/j.inffus.2019.12.012.
- Došilović, F.K., Brčić, M. and Hlupić, N. (2018) "Explainable artificial intelligence: A survey," in 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 210–215. doi:10.23919/MIPRO.2018.8400040.
- Ehsan, U. et al. (2018) "Rationalization: A Neural Machine Translation Approach to Generating Natural Language Explanations," in AIES 2018 - Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society. doi:10.1145/3278721.3278736.
- Guidotti, R. et al. (2018) "A survey of methods for explaining black box models," ACM Computing Surveys, 51(5). doi:10.1145/3236009.
- Gunning, D. and Aha, D. (2019) "DARPA's Explainable Artificial Intelligence (XAI) Program," AI Magazine, 40(2), pp. 44–58. doi:10.1609/aimag.v40i2.2850.
- Linardatos, P., Papastefanopoulos, V. and Kotsiantis, S. (2021) "Explainable ai: A review of machine learning interpretability methods," Entropy. doi:10.3390/e23010018.
- Lipton, Z.C. (2018) "The mythos of model interpretability," Communications of the ACM, 61(10). doi:10.1145/3233231.
- Melo, E., Silva, I., Costa, D. G., Viegas, C. M. D., & Barros, T. M. (2022). On the Use of eXplainable Artificial Intelligence to Evaluate School Dropout. Education Sciences, 12(12), 845. MDPI AG. Retrieved from http://dx.doi.org/10.3390/educsci12120845.
- McCarthy, J. et al. (2006) "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955," Al Magazine, 27(4), p. 12. doi:10.1609/aimag.v27i4.1904.
- Littman, M. L., Ajunwa, I., Berger, G., Boutilier, C., Currie, M., Doshi-Velez, F., ... & Walsh, T. (2022). Gathering strength, gathering storms: The one hundred year study on artificial intelligence (AI100) 2021 study panel report. arXiv preprint arXiv:2210.15767. Miller, T. (2019) "Explanation in artificial intelligence: Insights from the social sciences," Artificial Intelligence. doi: 10.1016/j.artint.2018.07.007.
- Shubhendu, S. and Vijay, J. (2013) Applicability of Artificial Intelligence in Different Fields of Life. Available at: www.ijser.in.
- Shukla Shubhendu, S. and Vijay, J. (2013) "Applicability of Artificial Intelligence in Different Fields of Life," International Journal of Scientific Engineering and Research (IJSER), 1(1).
- Tjoa, E. and Guan, C. (2021) "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI," IEEE Transactions on Neural Networks and Learning Systems, 32(11). doi:10.1109/TNNLS.2020.3027314.
- Altman, D.G. (1990) Practical statistics for medical research. CRC press.
- Chromik, M. (2021) "Making SHAP Rap: Bridging Local and Global Insights Through Interaction and Narratives," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). doi:10.1007/978-3-030-85616-8\_37.
- Das, D. and Chernova, S. (2020) "Leveraging rationales to improve human task performance," in International Conference on Intelligent User Interfaces, Proceedings IUI. doi:10.1145/3377325.3377512.
- Deperlioglu, O. et al. (2022) "Explainable framework for Glaucoma diagnosis by image processing and convolutional neural network synergy: Analysis with doctor evaluation," Future Generation Computer Systems, 129. doi:10.1016/j.future.2021.11.018.

- Ehsan, U. et al. (2018) "Rationalization: A Neural Machine Translation Approach to Generating Natural Language Explanations," in AIES 2018 - Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society. doi:10.1145/3278721.3278736.
- Gao, X. et al. (2020) "Joint Mind Modeling for Explanation Generation in Complex Human-Robot Collaborative Tasks," in 29th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2020. doi:10.1109/RO-MAN47096.2020.9223595.
- Irarrázaval, M.E. et al. (2021) "Telecom traffic pumping analytics via explainable data science," Decision Support Systems, 150. doi:10.1016/j.dss.2021.113559.
- Islam, M.R. et al. (2022) "A Systematic Review of Explainable Artificial Intelligence in Terms of Different Application Domains and Tasks," Applied Sciences (Switzerland), 12(3). doi:10.3390/app12031353.
- Khodabandehloo, E., Riboni, D. and Alimohammadi, A. (2021) "HealthXAI: Collaborative and explainable AI for supporting early diagnosis of cognitive decline," Future Generation Computer Systems, 116. doi:10.1016/j.future.2020.10.030.
- Landis, J.R. and Koch, G.G. (1977) "The Measurement of Observer Agreement for Categorical Data," Biometrics, 33(1). doi:10.2307/2529310.
- Melançon, G.G. et al. (2021) "A machine learning-based system for predicting service-level failures in supply chains," Interfaces, 51(3). doi:10.1287/INTE.2020.1055.
- Sabol, P. et al. (2020) "Explainable classifier for improving the accountability in decision-making for colorectal cancer diagnosis from histopathological images," Journal of Biomedical Informatics, 109. doi:10.1016/j.jbi.2020.103523.
- Samuel, S.S., Abdullah, N.N.B. and Raj, A. (2021) "Interpretation of SVM to Build an Explainable Al via Granular Computing," in Studies in Computational Intelligence. doi:10.1007/978-3-030-64949-4\_5.
- Sardianos, C. et al. (2021) "The emergence of explainability of intelligent systems: Delivering explainable and personalized recommendations for energy efficiency," International Journal of Intelligent Systems, 36(2). doi:10.1002/int.22314.
- Wang, C. and An, P. (2021) "A Mobile Tool that Helps Nonexperts Make Sense of Pretrained CNN by Interacting with Their Daily Surroundings," in Extended Abstracts of MobileHCI 2021 - ACM International Conference on Mobile Human-Computer Interaction: Mobile Apart, Mobile Together. doi:10.1145/3447527.3474873.
- Xie, Y. et al. (2020) "CheXplain: Enabling Physicians to Explore and Understand Data-Driven, Al-Enabled Medical Imaging Analysis," in Conference on Human Factors in Computing Systems -Proceedings. doi:10.1145/3313831.3376807.

## Empowering Students Computational Thinking through Robotics-enabled STEM Education

Daner SUN<sup>a\*</sup>, Peiyao TIAN<sup>b</sup>, Kam Yuen LAW<sup>c</sup>, Wai Han CHEUK<sup>d</sup>

 <sup>a,b</sup>Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR, China
 <sup>c,d</sup> Innovation and Technology Team, The Education University of Hong Kong Jockey Club Primary School, Hong Kong SAR, China
 \*dsun@eduhk.hk

**Abstract:** The present study aims to conceptualize and develop a comprehensive series of STEM lessons enriched by robotics, meticulously guided by the 5E inquiry model. By skillfully weaving together robotics and STEM education, while adhering to a well-defined pedagogical framework, this research endeavor seeks to bridge the existing gap in the effective amalgamation of CT education within the diverse realms of STEM education. A pilot study conducted in a primary school demonstrated the positive results in CT skills test.

Keywords: Computational thinking, Robotics, STEM, 5E inquiry model

#### 1. Introduction

With the collaborative efforts of teachers, researchers, and educators, the field of STEM education has witnessed a remarkable surge in research activities. Students now have the privilege of engaging in diverse forms of STEM education, notably those enriched with technology-driven components (Tekdal, 2021). In an era characterized by rapid technological advancements and the proliferation of computing tools—such as programming and coding platforms—the integration of these tools into STEM activities has become an inevitable progression (Shute et al., 2017; Chongo et al., 2020). This infusion of computing tools amplifies the advantages of STEM education for students, prominently fostering the development of Computational Thinking (CT) skills—a quintessential 21st-century competency (Barr & Stephenson, 2011). Computational Thinking embodies a systematic approach to unraveling problems and designing systems. It draws upon foundational concepts such as logic, abstraction, pattern recognition, and algorithmic design to skillfully deconstruct and conquer intricate challenges (Wing, 2006; National Research Council, 2010).

Traditionally entrenched in the domain of computer science and computing-related pursuits, Computational Thinking now finds itself pervading a spectrum of STEM contexts (Lee & Malyn-Smith, 2020). Numerous instances underscore the integration of CT skills within STEM education, with coding tools like mBot, Scratch, and app Inventor lending support to this trend. Nevertheless, the seamless amalgamation of CT education into STEM contexts, harmonized by well-crafted pedagogical principles, remains somewhat constrained.

In response, this study endeavors to conceive and develop a comprehensive series of robotics-facilitated STEM lessons, meticulously guided by the 5E inquiry model. By innovatively interweaving robotics and STEM education, while adhering to a well-defined pedagogical structure, this study aims to bridge the existing gap in the effective integration of CT education within STEM domains.

#### 2. Literature Review

Computational Thinking (CT), initially introduced by Papert (1980) and subsequently elaborated upon by Wing (2006), has emerged as a foundational concept. It was initially recognized as a cornerstone within the realm of computer science, involving "problemsolving, system design, and comprehension of human behavior, drawing upon fundamental computer science concepts" (Wing, 2006). The construct of CT was refined by Campbell and Heller (2019) and Yin et al. (2020) to encompass decomposition, abstraction, pattern recognition, and algorithmic thinking. Grover and Pea (2013) further delineated CT skills to comprise elements such as abstraction, pattern generalization, algorithms, logic, problem decomposition, debugging, productivity and performance constraints, parallel thinking, and systematic information processing. Amid the contemporary landscape, CT has solidified its place as a pivotal competency in the 21st century. Consequently, over the past decade, a multitude of theoretical and empirical studies have been conducted to unravel its nuances (Sands et al., 2018). Among these explorations, investigations into CT education within STEM contexts have consistently demonstrated its efficacy in nurturing CT development. bolstering STEM knowledge acquisition, and fostering higher-order cognitive skills, including creativity and problem-solving prowess (Ogegbo & Ramnarain, 2021).

As a captivating pedagogical approach in STEM education, robotics has garnered significant attention due to its inherently interdisciplinary nature. Prior research has underscored the potential of infusing CT education through robotics design to foster robotics literacy, encompassing concepts like simple machines, sequencing, order, and control (Cejka et al., 2006). These facets naturally align with engaging contexts such as environmental exploration, innovative creation, and pragmatic problem-solving. The rapid strides taken in the realms of artificial intelligence and robotics technologies have facilitated the integration of diverse programming tools within AI-driven projects. Yet, the intricate nexus between intelligent robotics technology and the cultivation of computational thinking within the framework of STEM education remains a domain ripe for exploration. In light of this, the current study explored the impacts of robotics-enabled STEM lessons on students' computational thinking and their motivation to engage with these innovative lessons.

#### 3. Research Questions

Based on the above research background, this study aims to design a robot-enabled STEM program to develop students' computational thinking skills and enhance their learning motivations, Specifically, the research questions are as follows:

1) To what extent does robotics-enabled STEM education impact students' computational thinking knowledge?

2) To what extent does robotics-enabled STEM education influence students' learning motivations?

#### 4. Methodology

#### 4.1 Participants

To ascertain the sample's representativeness and accessibility, the study employed the purposive sampling method. This approach involved selecting all Grade 4 students from a local primary school in Hong Kong, renowned for its emphasis on ICT education. Prior to commencing the intervention, requisite permissions were secured from both the students' parents and the students themselves, ensuring voluntary participation. Subsequently, a total of 56 students, aged approximately 11 years on average, were chosen from four distinct classes.

#### 4.2 Instruments

#### 4.1.1 Computational Thinking skill Test

Drawing upon the works of Shute et al. (2017) and Curzon et al. (2019), an assessment of students' computational thinking skills was executed through a carefully designed test. This evaluative tool encompassed 20 questions, each necessitating students to articulate their problem-solving approach. The comprehensive problem-solving process comprised five distinct stages of computational thinking: identification and representation of strategies for problem resolution, decomposition of problems into manageable segments and identification of critical information, recognition of patterns, trends, and regular inferences, and a final review of the adequacy of the problem-solving solution. The scoring system allocated up to 5 points to each question, reflecting the progression through these stages (Table 1). Consequently, the maximum achievable score was set at 100 points. To ensure the validity of the assessment, the test questions underwent meticulous screening and correction by both frontline teachers and subject matter experts prior to its formal administration.

Component	Definition	Score
Algorithmic	Identifying and representing routines to solve a problem or	1
thinking	task, like ordered, step-by-step instructions	
Decomposition	Breaking down one problem, algorithm or process into smaller	2
	parts such that the partial results can be later integrated to	
	more easily solve or understand the whole problem	
Abstraction	Identifying essential elements of a problem or process; this	3
	involves simplifying and hiding detail	
Pattern	Inferring and identifying patterns, trends or regularities in a	4
recognition	certain problem or process	
Evaluation/	Reviewing the adequacy of solutions or elements to a problem	5
debugging		

Table 1. Dimension Classification of the Student Motivation Questionnaire

#### 4.1.2 Learning Motivation Questionnaire

The students' motivation was assessed using a questionnaire consisting of 13 questions, categorized according to the framework developed by Glynn et al. (2011) (as presented in Table 2). This questionnaire comprises a range of motivation-related aspects, including 3 items related to intrinsic motivation, such as "I find the prospect of working with robots intriguing," 3 items reflecting extrinsic motivation, such as "Acquiring knowledge about robotics will benefit my learning and future growth," 4 items concerning self-determination, such as "I am committed to investing substantial effort into learning robotics," and 3 items gauging self-efficacy, such as "I hold the belief that I performed well in the robotics activity." Respondents were requested to provide their responses on a 5-point Likert scale, spanning from "strongly agree" to "strongly disagree." In ensuring the reliability and validity of the questionnaire, a preliminary test of the questions was administered. This validation process yielded favorable results, with a Cronbach's alpha coefficient of 0.781 and a Kaiser-Meyer-Olkin (KMO) measure of 0.870, indicating strong internal consistency and reliability.

Dimension	Item	
Intrinsic motivation	I think it is interesting to work with robots.	

	I think knowledge of computer engineering, mathematics, physics and mechanical sciences is necessary to master robotics. Winning events and competitions is very important to me.			
Extrinsic motivation	Understanding robotics technology will help my learning and future development.			
	Participating in robotics activities will improve my academic performance.			
	I think I will be able to use what I have learned from the robotics activities in other courses.			
Self-determination	My future career dream is to become a scientist			
	I intend to put a lot of effort into learning robotics.			
	I will gather information from different sources, such as			
	math and physics, to use in robotics activities.			
	If I have a problem, I will continue to try and solve it by			
	reading the material without anyone's help.			
Self-efficacy	I am able to express and explain my ideas to my team			
	I believe that I have done a good job in robotics activities.			
	I participate actively in team activities.			

#### 4.2 Robotics-enabled STEM Curriculum

Collaborative efforts between educators and researchers culminate in the design and execution of Computational Thinking (CT)-focused STEM lessons, synergistically bolstered by the utilization of VEX Robotics toolkits. Specifically, the VEX Robotics toolkit, recognized for its educational value, offers an array of block-based coding tools tailored to students from kindergarten to grade 12. This inclusive approach translates into an "interactive, programmable robot that bridges the realms of Computer Science and Computational Thinking, transcending digital screens to become a tactile experience for Pre-Primary students." Guided by the established 5E inquiry model, the lessons seamlessly align with existing VEX activities available at https://education.vex.com/stemlabs/go/activities. The incorporation of the 5E inquiry model ensures an engaging and comprehensive learning process encompassing engagement, exploration, explanation, elaboration, and evaluation. Therefore, this study drew on the existing activities of VEX GO and modified them according to their characteristics and the cognitive characteristics of grade 4 primary students to form the final robot-supported STEM curriculum. The lesson plan can be found in Table 3, where students constructed an Astronaut Vault. The course lasted for 5 classroom hours. By intertwining CT principles with robotics education through the VEX Robotics toolkit and grounding the pedagogical approach in the 5E inquiry model, the collaborative endeavor seeks to foster an immersive and impactful STEM learning journey.

5E Inquiry phase	Student Activities				
Engagement	Students are introduced the lunar rover through watching videos and answer the related questions about lunar rover.				
Exploration	Student are asked to work with their classmates to build an astronaut vault which could speed quickly toward an object and stop before hitting it. They are discussing the design ideas and build the model				
Explanation	Students are invited to present their design and do demonstrations, with further explaining how it works.				

Table 3. Exemplar of lesson plan

Elaboration	Student are further elaborate their design based on teacher
	feedback and other group's comments.
Evaluation	The students answer the questions about CT knowledge and
	skills in the activities.

#### 4.3 Data Collection and Analysis

The dataset for this study comprises three distinct categories: the pre-test and post-test results measuring students' computational thinking skills, the pre-test and post-test outcomes gauging students' motivation through a questionnaire, and unstructured interviews conducted with the students. To ensure data collection consistency, the students' computational thinking skills and motivation to learn questionnaires was administered by their classroom teachers both prior to and following the intervention. Furthermore, unstructured interviews were carried out with three selected students under the guidance of a facilitator, who then recorded the collected data.

The ensuing data analysis encompasses both quantitative and qualitative methodologies. Initially, the students' computational thinking skills test results and their academic motivation outcomes underwent descriptive statistical analysis in SPSS. This encompassed the computation of total mean scores, standard deviations, and paired sample t-tests to ascertain the disparities between students' pre-intervention and post-intervention scores. Additionally, variations in each facet of academic motivation was visually represented using line graphs generated in EXCEL. Lastly, the interview transcripts from the students were analyzed to glean profound insights into their computational thinking skills and motivation to learn.

#### 5. Results

#### 5.1 Computational Thinking skills

An examination of descriptive statistics pertaining to students' total scores in computational thinking skills, both in the pre-test and post-test (refer to Table 4), revealed a notable trend: the students' post-test scores surpassed their pre-test scores. Evidently, this signifies a discernible enhancement in the students' overall computational thinking skills following the intervention.

Test	М	SD	SE
Pretest Score	67.717	6.298	.865
Posttest Score	81.422	6.387	.894

Table 4. Scores of the pre-test and post-test

Subsequently, a paired samples t-test was employed to analyze the pre-test and post-test results of the students' computational thinking skills assessment. The outcomes indicated a substantial disparity between the students' pre-test and post-test scores (p < .05), affirming a noteworthy improvement in their computational thinking skills subsequent to the intervention.

#### 5.2 Learning Motivation

Descriptive statistics were applied to the students' pre-test and post-test motivation scores, with the findings outlined in Table 5. The results unveil varying degrees of advancement in the four dimensions: intrinsic motivation, extrinsic motivation, self-determination, and self-efficacy, following the intervention. Among these dimensions, post-intervention self-efficacy registered the highest mean score, closely followed by intrinsic motivation. This pattern underscores how the intervention notably bolstered students' confidence in ICT-related

learning and fostered a heightened willingness to propel their learning through personal initiative.

Test	Dimension	М	SD	Skewness	SE	Kurtosis	SE
Pretest	IM	3.71	1.09	-1.124	0.319	0.748	0.628
	EM	3.63	1.18	-0.850	0.319	0.042	0.628
	SD	3.14	1.1	-0.067	0.319	-0.557	0.628
	SE	3.81	1.11	-1.045	0.319	0.731	0.628
Posttest	IM	4.01	1.06	-1.369	0.322	1.600	0.634
	EM	3.83	1.14	-1.149	0.322	0.814	0.634
	SD	3.43	1.16	-0.388	0.322	-0.439	0.634
	SE	4.04	1.08	-1.271	0.322	1.458	0.634

Table 5. Students' learning motivations of the pre-test and post-test

Note: IM: Intrinsic Motivation; EM: Extrinsic Motivation; SD: Self-Determination; SE: Self-efficacy

Paired-sample t-tests were subsequently conducted for each individual dimension as well as the overall mean, aiming to ascertain any distinctions between the pretest and posttest conditions. The findings are presented in Table 6. When analyzing the overall mean, the posttest total score significantly exceeded that of the pretest (p < .05). This substantial discrepancy points towards a marked elevation in students' overall motivation to learn subsequent to the intervention. Further examination of each dimension showcases noteworthy trends. Intrinsic motivation, self-determination, and self-efficacy displayed notable increments post-intervention, exhibiting statistically significant improvements. However, the variation in extrinsic motivation—pertaining to external incentives like monetary rewards or other external inducements—was relatively modest between the pretest and post-test stages. This suggests that post-intervention, students' motivation wasn't largely driven by external factors. Instead, the intervention primarily centered on bolstering their learning confidence and cultivating a genuine interest in the subject matter itself.

Pair	t	р
IM	-1.926	0.039*
EM	-1.244	0.219
SD	-1.575	0.041*
SE	-1.573	0.021*
Total	-1.816	0.045*

Table 6. Paired t test of students' pre and posttest of learning motivations

Note: IM: Intrinsic Motivation; EM: Extrinsic Motivation; SD: Self-Determination; SE: Self-efficacy

#### 5.3 Interviews

At the end of the intervention, three randomly selected students were interviewed to further confirm the effectiveness of the robot-assisted STEM programme. Insights from the interview data indicated that students consistently agreed on the benefits of integrating robotics into the STEM curriculum. The integration of robotics into the STEM curriculum was praised by students for infusing an inquiry-based and problem-solving approach, which promoted computational thinking skills.

Students confirmed the effectiveness of these classes in promoting problem solving skills by asking "Question 1: Do you think robot-assisted STEM classes have helped you with your problem-solving skills?" Students confirmed the efficacy of these courses in promoting STEM knowledge and competence.

Student A: It was okay, I learned general patterns of problem solving in the classroom activities.

Student B: I clearly grasped the content of the course, which will help me solve problems in the future.

Student C: It was helpful in developing my computational thinking skills because I need to solve a lot of problems on my own.

For "Question 2: After listening to the whole lesson, did you find the STEM programme more interesting?" They became interested in the STEM curriculum through robot-assisted STEM learning. In addition, as the intervention culminated, the students became more and more emotional-enthusiastic about the upcoming STEM course and renewed their resolve to take the related academics seriously.

In response to "Question 3: How well do you feel you are listening in STEM classes now compared to before?" Overwhelmingly, respondents reported that the classes hold their attention and provide a way for them to participate more effectively in the classroom. Students agreed that they performed better and were more attentive in STEM classrooms that leveraged robotics.

Student A: He likes the way the class is taught compared to before and listens attentively in every class.

Student B: I am interested in robotics, so I am more motivated than before.

Student C: Sometimes I get sleepy in class, but now I try to stay awake and concentrate on the STEM lessons.

Overall, this instructional approach served as a catalyst for student enthusiasm and motivation in the STEM field of study. It not only enriched students' current learning experiences, but also fostered their continued interest in future STEM projects.

#### 6. Discussions

This study entailed the design and implementation of a robot-enabled STEM program, structured upon the 5E model, with the primary objective of enhancing students' computational thinking skills and fostering motivation to learn. Significant advancements in students' computational thinking skills were observed post-program intervention, signifying a tangible improvement as a result of the intervention. This observation aligns with prior research indicating that STEM programs incorporating computational thinking elements contribute positively to the computational thinking skills of pre-service teachers (Ciftci & Topçu, 2023), as well as similar findings in the context of robotics-enhanced STEM summer camp activities and game design for children from diverse regions (Chiang et al., 2022; Leonard et al., 2016; Shang, 2023). The current study reinforces these findings, affirming that a robot-enabled STEM program effectively enhances computational thinking among primary school students in Hong Kong. This pedagogical intervention, brimming with inquiry and problem-solving processes, propels students to collaboratively tackle real-life issues. This immersive engagement fosters a deeper appreciation for the intricacies of problemsolving, thereby nurturing their computational thinking prowess (Shute et al., 2017; Barak & Assal, 2018).

Turning to the aspect of student motivation, the preponderance of prior research indicates that robot-enabled STEM programs distinctly elevate students' self-efficacy, learning attitudes, and enthusiasm for learning (Gomoll et al., 2016; Sisman et al., 2021; Üçgül et al., 2022). This study's findings reaffirm these conclusions, while also delving deeper to explore the program's influence on self-determination and extrinsic motivation. The results underscore that, upon intervention completion, students displayed substantial enhancements across intrinsic motivation, self-determination, and self-efficacy. Nonetheless, concerning extrinsic motivation, the mean scores witnessed growth, yet the variance from the pre-intervention phase did not attain statistical significance. This could be attributed to the program's intrinsic nature, which aims to cultivate enjoyment and interest by immersing students in inquiry and problem-solving activities, consequently boosting their self-efficacy in addressing challenges (Barak & Assal, 2018). As a result, future iterations of the course might consider incorporating immediate classroom assessments to amplify external motivation in alignment with their learning objectives.

#### 7. Conclusions

This study devised a robot-enabled STEM program rooted in the 5E instructional model. Through a comparative analysis of students' computational thinking skills and motivation prior to and following the intervention, significant insights emerged. The study revealed that the integration of robotics within STEM education holds the potential to foster considerable enhancements in students' computational thinking skills. Moreover, the infusion of robotics into STEM curricula yielded positive outcomes, ameliorating students' intrinsic motivation, self-determination, and self-efficacy.

This realization underscores the prospect for educators to leverage similar STEM programs in the future. By orchestrating comparable initiatives, teachers can ignite students' fascination for STEM learning, nurture adept problem-solving abilities, consequently augmenting their computational thinking proficiencies. In doing so, a harmonious advancement of students' STEM literacy can be achieved.

Nevertheless, this study is not without its limitations. Although it validated the efficacy of robotic-enabled STEM courses in bolstering intrinsic motivation, self-determination, and self-efficacy, the discernible impact on students' external motivation was not statistically significant between pre- and post-tests. This aspect presents an avenue for future refinement. A potential approach involves the integration of immediate classroom assessments, strategically aimed at amplifying external motivation.

In conclusion, this research underscores the transformative potential of robot-enabled STEM initiatives. By harnessing the power of robotics within the framework of STEM education, educators can foster a well-rounded development of students' computational thinking skills, intrinsic motivation, and self-determination, thereby paving the way for a comprehensive advancement of students' STEM literacy.

#### References

- Barak, M., & Assal, M. (2018). Robotics and STEM learning: students' achievements in assignments according to the P3 Task Taxonomy—practice, problem solving, and projects. *International Journal of Technology and Design Education*, 28, 121-144.
- Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K-12: What is involved and what is the role of the computer science education community?. *Acm Inroads*, *2*(1), 48-54.
- Campbell, L. O., & Heller, S. (2019). Building computational thinking: Design and making in teacher education. In *Recruiting, preparing, and retaining stem teachers for a global generation* (pp. 163-189). Brill.
- Cejka, E., Rogers, C., & Portsmore, M. (2006). Kindergarten robotics: Using robotics to motivate math, science, and engineering literacy in elementary school. *International Journal of Engineering Education*, *22*(4), 711.
- Chiang, F. K., Zhang, Y., Zhu, D., Shang, X., & Jiang, Z. (2022). The influence of online STEM education camps on students' self-efficacy, computational thinking, and task value. *Journal of Science Education and Technology*, *31*(4), 461-472.
- Chongo, S., Osman, K., & Nayan, N. A. (2020). Level of Computational Thinking Skills among Secondary Science Student: Variation across Gender and Mathematics Achievement. *Science Education International*, *31*(2), 159-163.
- Çiftçi, A., & Topçu, M. S. (2023). Improving early childhood pre-service teachers' computational thinking skills through the unplugged computational thinking integrated STEM approach. *Thinking Skills and Creativity*, 101337.
- Curzon P., Bell T., Waite J., Dorling M. (2019). "Computational thinking," In Fincher S., Robbins A. (Eds.), Cambridge Handbook of Computer Science Education Research (pp. 513–546). Cambridge University Press.
- Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, *48*(10), 1159-1176.

- Gomoll, A., Hmelo-Silver, C. E., Šabanović, S., & Francisco, M. (2016). Dragons, ladybugs, and softballs: Girls' STEM engagement with human-centered robotics. *Journal of Science Education and Technology*, *25*, 899-914.
- Grover, S., & Pea, R. (2013). Computational thinking in K–12: A review of the state of the field. *Educational researcher*, *42*(1), 38-43.
- Lee, I., & Malyn-Smith, J. (2020). Computational thinking integration patterns along the framework defining computational thinking from a disciplinary perspective. *Journal of Science Education and Technology*, *29*, 9-18.
- Leonard, J., Buss, A., Gamboa, R., Mitchell, M., Fashola, O. S., Hubert, T., & Almughyirah, S. (2016). Using robotics and game design to enhance children's self-efficacy, STEM attitudes, and computational thinking skills. *Journal of Science Education and Technology*, 25, 860-876.
- National Research Council. (2010). *Report of a workshop on the scope and nature of computational thinking*. National Academies Press.
- Ogegbo, A. A., & Ramnarain, U. (2022). A systematic review of computational thinking in science classrooms. *Studies in Science Education*, *58*(2), 203-230.
- Papert, S. (1980). Personal computing and its impact on education. *The computer in the school: Tutor, tool, tutee*, 197-202.
- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33.
- Sands, P., Yadav, A., & Good, J. (2018). Computational thinking in K-12: In-service teacher perceptions of computational thinking. *Computational Thinking in the STEM Disciplines: Foundations and Research Highlights*, 151-164.
- Shang, X., Jiang, Z., Chiang, F. K., Zhang, Y., & Zhu, D. (2023). Effects of robotics STEM camps on rural elementary students' self-efficacy and computational thinking. *Educational Technology Research and Development*, 1-26.
- Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review*, 22, 142-158.
- Sisman, B., Kucuk, S., & Yaman, Y. (2021). The effects of robotics training on children's spatial ability and attitude toward STEM. *International Journal of Social Robotics*, *13*, 379-389.
- Tekdal, M. (2021). Trends and development in research on computational thinking. *Education and Information Technologies*, *26*(5), 6499-6529.
- Üçgül, M., & Altıok, S. (2022). You are an astroneer: The effects of robotics camps on secondary school students' perceptions and attitudes towards STEM. *International Journal of Technology and Design Education*, 1-21.
- Wing, J. M. (2006). Computational thinking. Communications of the ACM, 49(3), 33-35.
- Yin, Y., Hadad, R., Tang, X., & Lin, Q. (2020). Improving and assessing computational thinking in maker activities: The integration with physics and engineering learning. *Journal of Science Education and Technology*, 29, 189-214.

# W13

ICCE Workshop on "Towards the Practice of Seamless Interest-Driven Creators (SIDC) Theory with Technological Supports"

### Role of Seamless Learning in Enhancing Interest-Driven Creator Theory

#### Su Luan WONG<sup>a\*</sup>, Mas Nida MD KHAMBARI<sup>b</sup> & Sai Hong TANG<sup>c</sup>

<sup>a</sup>Faculty of Educational Studies, Universiti Putra Malaysia, Malaysia
 <sup>b</sup> Faculty of Educational Studies, Universiti Putra Malaysia, Malaysia
 <sup>c</sup>Faculty of Engineering, Universiti Putra Malaysia, Malaysia
 \*suluan@upm.edu.my

**Abstract:** This paper discusses the concept of seamless learning and its alignment with Interest-Driven Creator (IDC) Theory in response to Tak-Wai Chan's keynote at the 1<sup>st</sup> MetaACES 2023. He introduced the Seamless Interest-Driven Co-Creator (SIDC) learning framework — an integration of seamless learning and IDC learning. This paper also explores the educational implications of SIDC learning framework.

**Keywords:** seamless learning, interest-driven creator theory, SIDC learning framework

#### 1. Introduction

In 2006, Chan and his team introduced the concept of a "seamless learning space" in their groundbreaking work titled "One-to-one technology-enhanced learning: An opportunity for global research collaboration." In this work, Chan et al. (2006) emphasised the shifting trends in student learning beyond traditional classrooms. They attributed this shift to the increased affordability of personal digital devices among schoolchildren and college students. Essentially, the widespread availability of affordable digital devices gave these students the opportunity to engage in learning outside of school, anytime and anywhere, with anyone. This innovative development raised students' expectations regarding the use of personal digital devices for learning within their school environment. Chan and his colleagues firmly believed that the line separating learning inside and outside classrooms could become increasingly blurred, due to advancements in future learning technologies. In their view, it was entirely feasible for students to seamlessly continue their learning experiences across various environments.

The notion of seamless learning space is about designing learning environments that leverage digital technologies, placing emphasis on "where to learn", "when to learn", "who to learn with", and "what kind of tools for learning" (Chan et al., 2006). At the same time, this seamless learning space must integrate the aspect of "how to learn" within the future educational learning landscape. Within this context, the tenets of Interest-Driven Creator (IDC) learning aligns seamlessly with the concept of a cohesive learning space. IDC theory emphasises the importance of tailoring learning to students' interests, promoting creative activities driven by those interests, and fostering a habitual culture of learning (Chan, et al., 2018). In other words, IDC learning champions how students engage in learning activities based on their individual interests, making it a natural fit for the seamless learning environment.

#### 2. Seamless Learning and IDC Learning

Integrating a seamless learning space and IDC Theory spawns a new concept — Seamless Interest-Driven Co-Creator (SIDC) Learning. Tak-Wai Chan presented the SIDC Learning framework in his keynote at the 1<sup>st</sup> MetaACES 2023 (14 June 2023, Hong Kong). The framework emphasises four fundamental orientations: learning, harmony and well-being (Harwell), equity, and environment. He envisaged that educators can optimise the proposed framework to design practical and unique student learning experiences in the seamless learning space. By considering the aforesaid orientations, educators can create a holistic environment that ignites and supports students' interests, promotes their well-being, provides equitable access to resources, and ensures a safe learning environment.

The learning element in the framework focuses on developing students' interest in a seamless learning world. IDC learning recognises that students become more engaged and motivated to create knowledge when they are genuinely interested in a topic. When students practice learning routines that are meaningful to them, they are conditioned to focus and engage in the learning process. When these learning routines are repeated incessantly, students form a learning habit that becomes second nature to them. It is possible to extend this learning habit into a learning ritual. Learning ritual is more than learning habit because it revolves practices that are meaningful to the learners. When learners practice a set of learning ritual, their mind and body become conditioned to focus and engage in the learning process. This structure can enhance concentration and overall learning performance. Seamless learning complements this by providing diverse contexts and opportunities for students to pursue their interests. Educators can create a seamless learning space that is non-iudgmental — a learning environment where students are free and comfortable to express their ideas, try new things and take risks without detrimental learning consequences to themselves. Educators may inject some elements of confusion and wonder in the SIDC learning environment. Exposing students to learning activities that would create a little confusion but are interesting enough to hold the learner's attention, can create mental space for students to respond to new or unfamiliar learning situations (Wong et al., 2023). Designing learning activities with well-regulated confusion will likely hold students' attention thus leading to deeper learning.

In addition, the "seamless" element in IDC learning encourages educators to curate diverse learning resources such as textbooks, online articles, videos, podcast and many others to support learner's needs. Educators will have more flexibility to tailor learning experiences to each student learning needs and preferences. By curating diverse learning resources, educators can provide multiple pathways for their students to learn according to their learning styles and interests. Sourcing for rich and varied learning contents allows students to have wider exposure to differing viewpoints and cultures thus fostering a more comprehensive understanding of the subject matter. When educators design multimodal learning activities that incorporate diverse resources, students are able to stay focused and become more engaged learners (Wong & Md Khambari, 2022). In other words, multimodality promotes a better learning experience for learners across different learning contexts (formal education, informal learning, and real-world experiences) as it incorporates the use of multiple sensory modalities such as visual, auditory and tactile methods.

Learning assessment can also be conducted more fluidly — continuous alternative assessments and feedback can be integrated into the home and school learning environments. It is an excellent idea to incorporate diverse learning assessment methods to provide timely feedback and support learners' progress regardless of being in the home or school learning environment. The SIDC learning environment provides opportunities for students to engage in self-learning or collaborative learning experiences, explore real-world applications of their knowledge, and acquire academic and 21st-century competencies.

In his keynote, Tak-Wai Chan also introduced the concept of "Harwell" as another critical aspect of the SIDC framework. It is a portmanteau of two words — Harmony and well-being. SIDC learning equates to happy learners. Learning happiness pertains to experiencing joy, both during and following the learning journey. When students are engaged in interest-driven activities, they experience a sense of fulfilment and satisfaction. This corresponds with Maslow's hierarchy of needs, as students' harmonious interactions with themselves, others, and the environment contribute to their overall well-being. Learners need to be in harmony with themselves, the people around them, and the environment. Harmony can be considered an affective learning outcome — it is achieved when the student's learning needs are fulfilled, gaining a sense of satisfaction and experiencing inner peace. Inevitably, this leads students to feel a sense of belongingness, being happy and eventually becoming lifelong learners. This outcome resonates well with IDC theory which promotes the following assumptions for designing learning activities to optimise students' learning experience:

- 1. Learning is tailored to the learner's interests;
- 2. Learning activities are conceived as creative endeavours driven by these interests;
- 3. The integration of interest-driven creative activities into students' everyday schedules fosters a habitual culture of learning.

Educators must play a vital role in ensuring Harwell is supported in the SIDC learning environment. Students need to have a clear understanding of whom to connect with — to nurture emotional, mental and physical well-being. They must be able to harmonise these three facets to flourish as happy learners. It is the educators' responsibility to create a constructive synergy between themselves and their students and also families to create a harmonious learning environment.

SIDC learning's focus on accessibility promotes equity in education. The seamless world is extremely dynamic, often enabled by affordable digital technologies. This allows students from different socioeconomic backgrounds to access learning resources without discrimination. The SIDC learning experience can be heightened as the cost of owning digital technologies has significantly reduced in recent years. Governments worldwide are giving serious attention to disadvantaged students by equipping them with quality digital devices for learning. No students are left behind when learning switches from the school learning environment to the home learning environment. In other words, students' interest-driven learning process is not disrupted when such a switch occurs. This aspect of SIDC learning is crucial for addressing educational disparities and ensuring all students have equal opportunities to pursue their interests. By reducing barriers to access, seamless learning fosters inclusiveness and empowers students to explore diverse areas of knowledge.

Environmental safety, both physical and digital, is an essential consideration in both seamless learning and IDC learning. The SIDC framework acknowledges the need for environmental safety and hygiene, particularly in light of emerging challenges such as infectious diseases. The COVID-19 pandemic is an excellent example of this worry. In his keynote, Tak-Wai Chan warned us that the world is on the brink of peril. He stressed that the pervasive use of digital technology has caused many detrimental impacts on society. The ubiquity of digital technology at an affordable price has also exacerbated the situation. The phenomenon of digital addiction is a significant cause of concern. Digital addiction encompasses the excessive and compulsive use of digital devices, mobile applications and online contents, often causing detrimental effects on the users' well-being. In the context of learning, students are drawn towards the attraction of social media, video games, streaming platforms and other digital distractions that interferes with their studies. The instant gratification and continuous connectivity afforded by ubiquitous digital technologies contribute to the reinforcement of addictive behaviours.

The SIDC framework also highlights the importance of digital literacy and online safety. As students engage in seamless learning experiences across various environments, educating them about responsible digital citizenship, online norms, and critical evaluation of online information is crucial.

#### 3. Implications for Education

The integration of seamless learning and IDC Theory has some implications for education. Educators can cultivate students' curiosity and motivation by designing learning experiences that leverage students' interests, leading to deeper engagement and better learning outcomes. Seamless learning provides a platform for students to explore their passions beyond the boundaries of traditional classrooms, fostering creativity, innovation, and self-directed learning.

Furthermore, the SIDC framework promotes a learner-centric approach to education. It recognises that students have diverse interests, learning styles, and needs. By embracing seamless learning, educators can personalise learning experiences to cater to individual student's interests and preferences. This customisation enhances student agency. Additionally, it empowers educators to present the learning contents that pique learners' interest, maximising their understanding and capability to connect with the curriculum and the real world settings. In essence, students are more likely to view their studies as more meaningful, leading to a more positive learning experience.

SIDC Learning also encourages collaboration and community engagement. Through seamless learning spaces, students can connect with peers, experts, and communities of practice. These interactions provide opportunities for cooperative learning, sharing ideas, and receiving feedback, thereby enriching the learning experience. Additionally, community involvement and real-world connections enhance the authenticity and relevance of students' learning, bridging the gap between classroom knowledge and its application worldwide.

However, implementing SIDC learning requires thoughtful design and responsible use of technology. Educators must ensure seamless connections between different learning contexts support meaningful learning experiences rather than distractions or addictive behaviors. Attention should be given to designing educational games and digital resources that align with learning goals and promote engagement, rather than perpetuating superficial learning experiences. The onus of designing interesting learning activities lies in the hands of the educators who are willing to embrace SIDC learning. Exposing student teachers to the SIDC framework at the teacher preparatory level would serve as a good starting point to shape their instructional practices to be resonant with the philosophy of SIDC learning. They in turn will likely apply the philosophy in their instructional practices when they teach in the actual classrooms as full-fledged teachers.

#### 4. Conclusions

When learning is driven by interest, better learning outcomes can be achieved. These learning outcomes can be fortified when educators play a vital role in supporting students' well-being and creating a harmonious learning environment while ensuring accessibility and inclusiveness in education. However, it is crucial for the stake holders to address the responsible use of technology, considering environmental safety and digital literacy, to avoid distractions and addictive behaviours. Overall, the SIDC learning framework offers opportunities for students to explore their interest, collaborate with peers and experts, and

bridge the gap between classroom knowledge and real-world application. By embracing this framework, educators can empower students to take ownership of their learning journey and promote the love for lifelong learning.

#### References

- Chan, T. W., Roschelle, J., Hsi, S., Kinshuk, Sharples, M., Brown, T., Patton, C., Cherniavsky, J., Pea, R., Norris, C., Soloway, E., Balachefe, N., Scardamalia, M., Dillenbourg, P., Looi, C.-K., Milrad, M., & Hoppe, U. (2006). One-to-one technology-enhanced learning: An opportunity for global research collaboration. *Research and Practice in Technology-Enhanced Learning*, *1*(1), 3-29.
- Chan, T.-W., Looi, C.-K., Chang, B., Chen, W., Wong, L.-H., Wong, S. L., Yu, F.-Y., Mason, J., Liu, C.-C., Shih, J.-L., Wu, Y.T., Kong, S.-C., Wu,L., Chien,T.-C., Liao, C.C.Y., Cheng, H., Chen, Z.-H., Chou, C.-Y. (2019). IDC theory: Creation and the creation loop. *Research and Practice in Technology Enhanced Learning*, 14 (26).
- Chen, W., Chan, T. W., Wong, L. H., Looi, C. K., Liao, C. C. Y., Cheng, H. N. H., Wong, S. L., Mason, J., So, H.-J., Murthy, S., Gu, X., & Pi, Z. (2020). IDC theory: Habit and the Habit Loop. *Research and Practice in Technology Enhanced Learning*, 15 (10).
- Wong, S.L. & Md. Khambari, M. N. (2021). Multimodal Learning during the COVID19 Pandemic: Exploring students' preferences. In Iyer, S. et al. (Eds.) Proceedings of the 30th International Conference on Computers in Education (Volumn II). Kuala Lumpur: Asia-Pacific Society for Computers in Education. pp.173-177.
- Wong S.L., Md. Khambari, M. N., Wong, L. H. & Tang, S. H. (2023). Exploring the development of student teachers' interest in educational technology through Interest Driven Creator Theory. *Research and Practice in Technology Enhanced Learning.* 18(39), 1-21.
- Wong, L.-H., Chan, T.-W., Chen, W., Looi, C.-K., Chen, Z.-H., Liao, C. C. Y., King, R. B., & Wong, S. L. (2020). IDC theory: Interest and the Interest Loop. Research and Practice in Technology Enhanced Learning, 15(3). https://doi.org/10.1186/s41039-020-0123-2
### MEGA World – A Platform of Multiplayer Educational Game for All

#### Maiga CHANG<sup>\*</sup>

School of Computing and Information System, Athabasca University, Canada \*maiga.chang@gmail.com

**Abstract:** This paper presents MEGA World, a platform that allows teachers to create their own virtual worlds for their course so their students can learn required knowledge and skills as well as be stealth assessed in a non-pressure environment. Students will walk through the learning and assessment activities by wondering and exploring the virtual world via interest driven. In this paper, various quest types will be introduced, and their use cases will be explained with the content from four graduate level and undergraduate level courses in Athabasca University. At the end, users' experiences and their preliminary perceptions toward the use of MEGA World will be summarized – users spent on average 70 to 81 minutes on the gameplay activities even they were introduced MEGA World and its activities in the last second week of the term/semester; moreover, they had very positive opinions on the use of it to replace the supervised exam.

**Keywords:** Multiplayer, Interest-Driven, Educational Game, Platform, Stealth Assessment, Role-Playing

#### 1. Introduction

Students in the 21st century have more different ways to learn than the students before (Prensky & Berry, 2001). They live in a media-rich environment where they can learn different knowledge from innovative and interactive ways such as games (Chang & Kinshuk, 2010). Computer games have an engaging and immersive nature, have inspired academics to adopt and deploy educational games and game-based learning to courses (Tobias & Fletcher, 2007). Educational games are interactive and can get students motivated in learning; therefore, they can enhance the cognitive process with gradual increase in difficulty while improving students' learning performance and their understanding of the subject (Tobias, S. et al., 2011).

Barr (2018) finds the attitude the undergraduate students have toward the use of commercial video games to develop skills and competencies. Their result shows that the students have positive perception in favor of having video games to assist their learning and also believe they can develop their communication skill as well. Similarly, Wouters and colleagues (2013) find that the adoption of digital game in the learning activities can improve students' learning motivation and make them be more active while working on problem-oriented learning activities.

Although there are different game genres, Gros (2007) lists seven game genres that are agreed broadly: (1) action game; (2) adventure game; (3) fighting game; (4) role-playing game; (5) simulation game (6) sports game; and (7) strategy game. Among the above mentioned seven game genres, Podjačevs and Skorobogatova (2017) indicate that the use of role-playing in learning is one of the most efficient methods to learn a foreign language as well as to develop mental personality. The role-playing game allows students to play a role of fictional character to explore the virtual world – which is aligned with developing skills of decision-making, cooperation, problem-solving, and social interactions (Chang & Lin, 2014).

MEGA World (Multiplayer Educational Game for All) is a platform developed by Chang and Kinshuk (2010). The platform allows teachers to create virtual worlds, quests for learning and assessment, non-player characters (NPCs), quest and reward items/tools for their students to learning via role-playing in the game. The rest of the paper is organized as follows: Section 2 introduces MEGA World and Section 3 explains the various quest types with real world course content. A preliminary user perception toward the use of MEGA World in a formal university course is summarized in Section 4. At the end, Section 5 briefly makes a summary and discusses the limitations.

#### 2. MEGA World

MEGA World hosts multiplayer game worlds (i.e., virtual worlds) in which students can interact with non-player characters (NPCs), items, and other students. MEGA World as an educational role-playing game platform was proposed and developed by Chang and Kinshuk (2010). MEGA World supports any language and can access any existing external resources, e.g., multimedia, materials, online meetings, etc. Teachers can create their virtual worlds (see Figure 1) as well as create learning and assessment activities (i.e., quests in the game) for students. Students can learn specific knowledge and reach the learning goal of ActionScript programming language (Kuo, et al., 2010), research methodology (Li, et al., 2018), and English (Xu, et al., 2016) by taking and solving those quests while playing.



Figure 1. MEGA World v3.0 in Dark Theme (access: https://megaworld.game-server.ca).

MEGA World v3.0 currently has **twelve** (12) quest types include calculation, conversation (Chang, et al., 2019), fill-in-the-blank, greeting (check-in), item collection and delivery (check-in with item), multiple choice, short answer, single choice, sorting, speaking, treasure hunting (coordinates), and true/false quest type. Teachers are also provided management portal (see Figure 2) in which editors like map and quest editors are available so they can freely expand their game worlds and design quests in different levels for different subjects/topics according to their own teaching plan.

A virtual world (map) can host students who enroll in different courses. Teachers can design one or more professions to associate to a course they teach and any quest the teachers design and create should be linked to a profession if it is related to the course/subject matter. With this design, students must become a designated profession first (see Figure 3) before they can start the learning process and they won't be able to see



advanced level of the knowledge/skill as the teachers can control the learning progress via

*Figure 2.* Management Portal for Teachers.



Figure 3. A profession is associated with a course or a learning subject.

#### 3. Example Quests in Real Course

profession levels and/or hierarchy.

In this section, several quest types include Sorting Quest, Speaking Quest, Treasuring Hunting Quest, and Conversation Quest, are introduced with the course content in the real world.

As Figure 4 shows below the sorting quest type could be used by teachers to ask their students to match and connect the key concepts and their definitions. Of course, the

sorting quest type can also allow teachers to ask their students about the correct steps and actions of a procedure or process. The teacher of Java Programming language course asks students to match the definitions of the three types of Java: Java Mobile Edition, Java Enterprise Edition and Java Standard Edition.



Figure 4. Sorting Quest

Besides the open-end short answer quest type, MEGA world also supports "speaking" based short answer quest type. It is like the short answer quest, but it asks students to enter their answer via speaking into a microphone instead of typing their answer. With the speaking-based quests, teachers can force students to practice their speaking skill under a non-pressure environment. Figure 5 shows a similar speaking quest that International English Language Testing System (IELTS) has. While speaking to the mic, students can see how the computer recognizes their words correctly or incorrectly – so they can make proper changes and practice more to ensure people could read their words.

= Menu	🌣 100 🖉 195 🎝 2983 \$ 94	
4	Another Ted Quest 🛞	Q
Location 0, 0 Plaza 3	Ted's Tester Level 0	٩
Details 2	Ted can ask you in speaking test. You should prepare ideas for these topics but not memorise whole answers. You answers should be natural when you speak in the test. Prepare your daily life, your recent memories, your childhood memories, your opinions, popular things in your country etc. Remember –part 1 is all about you and your. (see https://ieltsliz.com/ielts- speaking-part-1-topics/) Click the microphone button and speak the answer What is your job? Where do you work? Why did you choose that job?	
	<u> </u>	
Bread Bread	Submit Answer	Another Ted Quest Aug 17, 2023, 11:30 a.m.
	← Back Mana 195 Health 127 Carry Power 15	
<b>*</b>		

Figure 5. Speaking Quest

Another worth to mention quest type in MEGA World is the treasure hunting (coordinates) quest. This quest type allows teachers to create a scenario that students may see in the real world while conducting a task. For example, for conducting literature review

students may use their devices (e.g., computer, laptop, or even mobile phone) to access the academic database via their university library. Figure 6 shows a quest, "finding the literature", that gives the student a "Computer" as tool and asks the student to "go to library at (4, 1)" on the map "digging" the mentioned paper out and bringing back to the NPC "Ted's Tester". When the student reaches the library at (4, 1), he or she uses the given tool to "Dig" and finds the article. This kind of quest solving experience makes students not only get familiar with the operation and flow of doing a task, but also emphasizes the **important tools** students will need for completing a task.



Figure 6. Treasure Hunting Quest

MEGA World also bridged with Speaking-based Conversation Quest (SCQ) System at <u>https://conversation.megaworld.game-server.ca</u> (Chang, Chen, Wu, & Yu, 2019; Chen, Chang, Wu, & Yu, 2018; Li, Chang, & Wu, 2020). Teachers can create and manage (see Figure 7) the conversations that they want their students to learn and practice via speaking to the mic. A conversation can be on any topic and difficulty level in any language. The speech recognition is supported by the Web Speech API built-in majority of the browsers include Chrome, Edge and Safari (<u>https://developer.mozilla.org/en-US/docs/Web/API/SpeechRecognition#browser\_compatibility</u>).

4 D C A 6	https://conversation.megaworld.game-server.ca/conversationLis	st.php					Ŀ	2 0	Ξ
MEG		Demo Video	NPC List	Conversation List	NPC Create	Add Conversation	Sign In		Î
	Conversation List								
	Торіс	Title		Level					
	Food	Ordering food and beverages	Subway	3		/ III (O			
	Greetings	How are you doing?		1		/ III (O			
	Greetings	How are you doing?		2		/ III (O			
	Hobbies	What did you do during the las	t holiday?	2		/ III (O			
	Hobbies	What do you do in your free ti	me?	1		/ III (O			
	Hobbies	What will you do during the ne	ext holiday?	3		/ III (O			
	Introducing yourself	What is your name?		1		/ III (O			
	Introducing yourself	What's your name?		2		N == (0			
	Introducing yourself	What's your name?		3		/ III (O			
	Movie	What movie would you like to	see?	1		/ III (O			
	Music	Have you ever been in a Coldp	lay concert?	3		/ III (O			
	Music	Were you a Michael Jackson fa	n?	2		/ III (O			
	Music	What kind of music do you like	?	1		/ III (O			

Figure 7. Conversations can be created by teachers.

Since teachers would like to have their students learn proper responses to people in speaking a language, they can create and arrange the conversation branches alone or with other teachers together. A branch can be reused, and a conversation can continue by jumping from one place to another place in the conversation tree as Figure 8 shows below.



Figure 8. A Conversation Tree.

Teachers then can create a conversation quest in MEGA World and link the quest in MEGA World with the conversation tree in the SCQ System with the conversation identifier (i.e., an id number). When students work on solving the quest, they will be redirected to the SCQ System and speak to the Non-Player Character (NPC). When they complete the conversation, the NPC will send a mark back to MEGA World so the conversation quest can be concluded.



Figure 9. Conversation Quest

#### 4. Preliminary User Perception

Four undergraduate and graduate level courses at Athabacsa University have virtual worlds (i.e., maps), NPCs, and quests created on MEGA World; they are Java Programming, Managing Innovation and Change in Information System, Research Methodology, and History of Psychology. This section summarizes users' perceptions toward the use of MEGA World in Java Programming learning subject. The Java Programming course has four maps, twenty-four NPCs, and 108 quests distributed in 10 levels.

In the Java Programming course, 11 students were asked to use MEGA World as a self-assessment tool and then voluntarily report their perceptions include whether or not they have seen similar game before, do they think MEGA World can be a self-assessment tool, do they think if MEGA World can replace the supervised exam, and do they want to play more in other courses in the future. Among nine users who provided their opinions, at the end seven users' feedback is considered and summarized in this paper due to two users were not really playing in the course's virtual world but other virtual worlds instead.

In average, the seven users played 70 to 81 minutes and most of them solved 10 to 25 quests while two of them solved only 2 to 6 quests. Six users rated the quests are easy, not difficult, beginner level, or easier side while one user rated 7 out of 10 for the difficulty level of the quests. Most of the users, five of them, had never seen similar games before while one had seen, and one had not seen such game in Massively Multiplayer Online (MMO) game form.

Regarding their perceptions on whether MEGA World can be a self-assessment tool, most of them believe so – only one user disagreed and explained that he or she preferred quizzes with selectable answers and automated code testing instead. While six out of seven users believed MEGA World can be used as a self-assessment tool, they all suggested having more quests in higher difficulty levels. This is understandable as the virtual worlds' first 28 quests belong to the first two levels of the course and none of the users proceeded beyond that probably because this activity is voluntarily, and they want to spend more time on it after 70 to 81 minutes or more.

For the possibility of replacing the traditional supervised exams with MEGA World, six of them except the one who preferred quizzes and automated code testing had very positive

opinions including "while I prefer project-based assessment but consider it is and supervised exams are equivalent," "I get very stressed and overwhelmed on exams and I would think this platform would lighten that aspect for me," "this is a good idea to replace traditional exams actually," "assuming the exam question are replaced by the question or activities in the quest, then it could replace supervised exams," "I would heavily support this." Besides the positive perceptions, users also suggest to still having some supervision so MEGA World is not considered an open book exam that students can research answers while playing it.

#### 5. Conclusion

This paper introduces MEGA World, a multiplayer education game platform that allows teachers to create their own virtual world maps, non-player characters (NPCs) and role-playing quests for their courses. MEGA World currently has 12 quest types that cover most of the question types used in a quiz and exam. Besides the common question types like true/false, single choice, multiple choice, and fill-in-the-blank, MEGA World also supports sorting, speaking, conversation, treasure hunting, and open-end short answer quest types.

The research team has created virtual worlds, NPCs, and quests for four courses at Athabasca University, including Java Programming, Managing Innovation and Change in Information System, Research Methodology, and History of Psychology. Taking Java Programming course as example, the course has 24 NPCs dispatched on 4 maps offering students 108 quests distributed in 10 levels.

The research reports users' preliminary perceptions toward the use of MEGA World in the Java Programming course as a self-assessment tool and the possibility of replacing supervised exams with MEGA World. Six out of seven users believed MEGA World is good for use as self-assessment tool and had very positive opinions on the use of it to replace the supervised exam. One user expressed his/her feeling about exams – stressful and overwhelming – and thought MEGA World can lighten this aspect very much. Finally, one user considered MEGA World to be equivalent to supervised exams while he or she prefers project-based assessment.

There is one recommendation according to the results the research team would like to remind practitioners, including teachers and educators who may want to adopt multiplayer educational game activities into their courses. As the results show that on average the users spent 70 to 81 minutes on the gameplay in MEGA World, they have done about 10 to 25 quests out of the 108 quests the virtual world has for the course content. The 10 to 25 quests belong to the first two difficulty levels while the 108 quests created for 10 difficulty levels. Under such circumstances, users explicitly express their desire to see more difficult quests and have positive willingness and attitude toward the use of MEGA World to replace the supervised exams.

Since the course has MEGA World and its gameplay activities arranged at the last second week of the course, one week before they are supposed to write the supervised final exam. They cannot put more effort and time in terms of doing gameplay activities and completing more quests. The research team would recommend practitioners to start the gameplay activities earlier and even at the beginning of the course. In such a case, users can have time to play more – the more they play, the more difficult quests they can see and practice to solve. Their gameplay progress could match the syllabus and weekly teaching/learning plan and even be considered as either preview or review of the course content that they have been taught in the class.

#### Acknowledgements

Thank all the NSERC, Mitacs Globalink, SIL research interns/assistants and various university students (see <u>https://megaworld.game-server.ca/#team</u>) who help to improve MEGA World in terms of (sub-)systems development, features implementation, graphic assets design and creation as well as help to create virtual worlds, NPCs, quests for courses.

#### References

- Barr, M. (2018). Student attitudes to games-based skills development: Learning from video games in higher education. *Computers in Human Behavior*, *80*, 283-294. https://doi.org/10.1016/j.chb.2017.11.030
- Chang, M., & Kinshuk. (2010). Web-based Multiplayer Online Role Playing Game (MORPG) for Assessing Students' Java Programming Knowledge and Skills.
- Chang, M., Chen, C.-T., Wu, K.-H., & Yu, P.-S. (2019). Conversation Quest in MEGA World (Multiplayer Educational Game for All). In *the Proceedings of International Conference on Smart Learning Environments (ICSLE 2019)* (pp. 77-82), Denton, TX, USA, March 18-20, 2019.
- Chang, S.-M., & Lin, S. S. J. (2014). Team knowledge with motivation in a successful MMORPG game team: A case study. *Computers & Education*, 73, 129-140.
- Chen, C.-T., Chang, M., Wu, K.-H., & Yu, P.-S. (2018). Web-based Conversation Quest for Enhancing English Speaking Skills. In *the Workshop Proceedings of the 26th International Conference on Computers in Education*, Manila, Philippines, November 26-30, 2018. Access: https://apsce.net/conference 4.php
- Gros, B. (2007). Digital games in education: The design of games-based learning environments. *Journal of Research on Technology in Education*, 40(1), 23-38.
- Kuo, R., Chang, M., Kinshuk, & Liu, E. Z.-F. (2010). Applying Multiplayer Online Game in Actionscript Programming Courses for Students Doing Self-Assessment. In *the Proceedings of Workshop on New Paradigms in Learning: Robotics, Playful Learning, and Digital Arts (pp. 351-355)*. Putrajaya, Malaysia.
- Li, K.-C., Chang, M., & Wu, K.-H. (2020). Developing a Task-based Dialogue System for Language Learning. *Education Sciences*, *10*(11), Article #306. Access: https://doi.org/10.3390/educsci10110306
- Li, Z., Zou, D., Xie, H., Wang, F. L., & Chang, M. (2018). Enhancing Information Literacy in Hong Kong Higher Education through Game-based Learning. In *the Proceedings of 22nd Global Chinese Conference on Computers in Education (GCCCE 2018)* (pp. 595-598), Guangzhou, China, May 25-29, 2018
- Podjačevs, P., & Skorobogatova, O. (2017). Computer Role-Playing Games as One of the Most Efficient Methods of Teaching and Learning in Higher Education. *International Journal on Information Technologies and Security*, 9, 65-78.
- Prensky, M., & Berry, B. D. (2001). Do they really think differently? On the Horizon, 9(6), 1-9.
- Tobias, S., & Fletcher, J. D. (2007). What research has to say about designing computer games for learning. *Educational Technology*, 47(5), 20-29. Access: <u>https://www.jstor.org/stable/i40186175</u>
- Tobias, S., Fletcher, J. D., Dai, D. Y., & Wind, A. P. (2011). Review of research on computer games. In S. Tobias & J. D. Fletcher (Eds.), *Computer Games and Learning* (pp. 127-222). Charlotte, NC: Information Age Publishing.
- Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249.
- Xu, B., Chang, M., Chen, G., Chen, Z.-H., & Kinshuk. (2016). Preliminary Study on the Influence of Role Playing Quests in Educational Game. In *the Proceedings of 20th Global Chinese Conference in Computer Education (GCCCE 2016)* (344-347), Hong-Kong, May 23-27, 2016.

# Web-based VR Education Contents Supporting VR-goggles and User Study

Yoshihiro OKADA<sup>a\*</sup>, Kosuke KANEKO<sup>b</sup>, Wei SHI<sup>c</sup>

<sup>a</sup>Innovation Center for Educational Resources, Kyushu University, Japan <sup>b</sup>Academic Research and Industrial Collaboration Management Office, Kyushu University, Japan <sup>c</sup>Research Institute for Information Technology, Kyushu University, Japan \*okada.yoshihiro.520@m.kyushu-u.ac.jp

**Abstract:** This paper treats the user study for web-based VR education contents supporting VR-goggles. Educational materials using multimedia have ability to enhance the education efficiency rather than text books. Similarly, educational materials using new technologies like VR have ability to attract much interests of students. Based on Interest-Driven Creator Theory (IDC Theory), this means educational materials using VR may have high educational efficiency. For experiment subjects and exercise subjects, the most efficient learning environments should be the same as the real ones as possible. Educational materials using VR can provide virtually students with high immersive learning environments. Therefore, educational VR materials supporting VR-goggles may have higher education efficiency rather than the others. This may justify Seamless Interest-Driven Creator Theory (SIDC Theory). So, this paper introduces several web-based VR education contents, and also show simple user study results especially for the comparison between their web-browser mode and VR goggle mode.

Keywords: IDC Theory, SIDC Theory, Educational materials, VR, User study

#### 1. Introduction

In this paper, we introduce several VR education contents supporting VR-goggles and show the user study results for the comparison between their web browser mode and VR goggle mode. This research is one of the activities of our center called ICER (Innovation Center for Educational Resources)(ICER, Aug. 2023) belonging to Kyushu University Library of Kyushu University, Japan because the mission of ICER is to provide students with educational materials using recent ICT and to support teachers for their creating such educational materials. In addition, our university established new institution called Data-Driven Innovation Initiative in April, last year. It consists of five divisions. We also belong to one of the divisions called NOE(Division of Next generation Open Education promotion)(NOE, Aug. 2023). Its mission is to promote Education DX(Digital Transformation), currently, we focus on the development of XR(VR/AR/MR) contents. Our ICER and NOE have various equipment and devices such as VR goggles, 3D scanners, and 360VR cameras for developing e-learning materials including XR contents like VR tours. Using these devices effectively, we want to develop a lot of e-learning materials for students of our university. To do so, we have already proposed the development framework for web-based VR education contents that supports lidar data(PCD: Point Cloud Data) and 360VR images/videos(Okada, et. al, Feb. 2023, Okada, et. al. Jul. 2023). In addition, we have proposed web viewers supporting VR-goggles for 3D model data, PCD and 360VR images/videos (Okada & Oki, Mar. 2023).

Educational materials using multimedia have ability to enhance the education efficiency rather than text books. Similarly, educational materials using new technologies like VR have ability to attract much interests of students. This can be derived from Interest-Driven Creator Theory (IDC Theory). For experiment subjects and exercise subjects, the most efficient

learning environments should be the same as the real ones as possible. Educational materials using VR can provide virtually students with high immersive learning environments. Therefore, educational VR materials supporting VR-goggles may have higher education efficiency rather than the others. This may justify Seamless Interest-Driven Creator Theory (SIDC Theory). So, this paper introduces several web-based VR education contents, and also show simple user study results especially for the comparison between their web-browser mode and VR goggle mode.

The remainder of this paper is organized as follows: next Section 2 describes related work. We explain the proposed framework and its functionalities using the VR tour contents in Section 3. In Section 4, we briefly introduce the proposed web viewers. Section 5 shows the user study results and discuss about them. Finally, we conclude the paper in Section 6.

#### 2. Related Work

Usually, we have to use any toolkit systems when creating web-based interactive 3D educational materials. Our laboratory proposed IntelligentBox(Okada & Tanaka, 1995) that is a development system for interactive 3D graphics desktop applications and uses it for our research purposes. There have been many desktop applications actually developed using IntelligentBox so far. In addition, we proposed web-version of IntelligentBox(Okada, 2013) and developed web-based applications using it. However, those cannot support VR goggles such as Meta Quest 2. Unity is the most popular game engine in the world (Unity, April 2023), that enables creating web contents supporting VR goggles such as Meta Quest 2. It is a very powerful toolkit for creating desktop and web-based 3D graphics applications including VR applications. However, the use of Unity requires programming knowledge and skills of the operations for it. Therefore, it is not easy for standard end-users like teachers to use Unity. Therefore, we proposed the development framework for web-based VR contents like VR tours.

There are many commercial services for creating interactive web contents using 360VR camera images. The service of RICOH (THETA, April 2023) does support 360VR images but not 360VR videos nor PCD. The service of Matterport (Matterport, April 2023) does support 360VR images and PCD but not 360VR videos. Contrarily, our proposed development framework consists of three systems those are for walkthrough contents of 360VR images, for navigation contents of 360VR videos and for walkthrough contents of PCD. These are differences between our development framework and the others. For the convenience of standard users like teachers, we also proposed web viewer applications for 3D model data, PCD and 360VR images/videos.

As the educational efficiency of e-learning materials using VR, we have not investigated related works enough yet. As a preliminary research, we show user study results for web-based VR education contents supporting VR-goggles in this paper.

#### 3. Web-based VR contents supporting VR-goggles

Last year, we made VR tour contents of the university library building for our open campus event to appeal that the building is one of the biggest university libraries in Japan. In the followings, we explain two types of VR tour contents of the university library using 360VR images and 360VR videos.

#### 3.1 Walkthrough content of 360VR images

Figure 1 includes three screen shots of the walkthrough content of 360VR images, 3rd floor of the library building. The map of 3rd floor appears in the left upper part of each screen shot. On the map, there are many orange dots totally over 100 each indicates the place at that each 360VR image was taken. By clicking on the orange dot, its corresponding 360VR image will be displayed as the next 360VR scene. Similarly, by clicking on one of the thin grey cylinders in a 360VR scene, the 360VR image taken at the same location will be displayed as the next

360VR scene. In this way, you can walk through on the floor. This VR tour content supports multimedia like standard image files and movie files as shown in the middle part(standard movie) and right part (standard image) of Figure 1.



*Figure 1.* Four screen shots of the walkthrough content of 360VR images, 3rd floor of the library building.

#### 3.2 Navigation content of 360VR videos

Figure 2 includes three screen shots of the navigation content of 360VR video, 3rd floor of the library building. The lower middle part of each of the three screen shots is a control panel for the 360VR video, i.e., play, pause, backward, forward, etc. The upper left part of each figure is the map of 3rd floor. There is a cyan color closed polyline that means the moving path actually the person wearing a portable 360VR camera on the top of his/her head moved through the floor when taking this 360VR video. By clicking on the polyline, the 360VR video moves to the corresponding 360VR scene. Similar to the system for the above walkthrough content, this content also supports multimedia like standard image files as shown in the middle figure and the right figure. The 360VR scene automatically changes according to the current playing point of the 360VR video. Therefore, we call this type of contents are navigation ones.



*Figure 2. Three* screen shots of the navigation content of 360VR video, 3rd floor of the library building, and 3rd floor map.

#### 4. Web Viewers for 3D model, PCD and 360VR images/video

Figure 3 shows four screen shots of the web viewers. Its left upper part is the 3D model viewer of a brain model. Its right upper part is the PCD viewer of CT exercise room. Its left lower and right lower parts are the 360VR image viewer and the 360VR video viewer of CT exercise room. You can see [Enter VR] button on the right lower part in each screen shot. This button is for changing web-browser mode into VR-goggle mode.





*Figure 3.* Four screen shots of the web viewers of 3D model(left upper), PCD(right upper), 360VR image(left lower) and 360VR video(right lower).

#### 5. User Study Results

We carried out simple user evaluation by the students of Graduate School of ISEE(Information Science and Electrical Engineering) of our Kyushu University. The number of students is 26. The user evaluation contents are VR tours (Walkthrough and Navigation) of University Library already shown in Figure 1 and Figure 2, and 3D model viewer of a Skeleton model. These contents in VR-goggle mode are shown in Figure 4, Figure 5 and Figure 6.

Table 1 includes three questions and Figure 7 shows their results. The left of the figure is for the walk-through content of 360VR images, the middle is for the navigation content of 360VR video, and the right is for the 3D model viewer of a skeleton model.

About question 1: Operability and responsiveness, the answer results are not bad because the mean values of the three contents are 3.2, 3.5 and 3.8. Although the usage of VR-goggle controllers was not explained, most students can operate by their selves.

#### Table 1: Questions

- Q1: Operability and responsiveness (1: Bad 5: Good)
- Q2: Immersive (1: Weak 5: Strong)
- Q3: 3D sickness (1: Strong 5: Weak)



*Figure 4.* Screen shot of the walkthrough content of 360VR images, 3rd floor of the library building, in VR-goggle mode.



*Figure 5.* Screen shot of the navigation content of 360VR video, 3rd floor of the library building, in VR-goggle mode.



*Figure 6.* Screen shot of the 3D model viewer, a skeleton model, in VR-goggle mode.



*Figure 7.* Answer results against the questions of Table 1, left is for the walk-through content of 360VR images, middle is for the navigation content of 360VR video, and right is for the 3D model viewer of a skeleton model.

About question 2, Immersive: the answer results are very good because the mean values of the three contents are 3.8, 4.1 and 3.7. Because the experience of a VR-goggle is the first time for most students, many students felt wonderful experience by the high immersion of VR-goggle.

About question 3, 3D sickness: the answer results are not bad because the mean values of the three contents are 3.4, 2.8 and 3.8. The navigation content is the moving action of a 360VR video, so most students felt 3D sickness rather than the other two contents.

We also received some free description comments such as "This was my first time using VR goggles, and the immersive experience greatly changed due to the image quality and operability. I was intrigued by the high level of technology that makes me think that the VR industry will become more popular in the future" and "I was surprised because it was more immersive than I expected. I wanted some VR goggles". Totally, most of the comments are positive and a few are negative. Therefore, we should develop more VR education contents to provide students.

#### 6. Conclusion

The development of educational VR contents is not easy and needs much time. So, in this paper, we introduced our proposed development framework for web-based VR education contents and our proposed web viewers for 3D model data, PCD and 360VR images/videos. We should use new technologies for educational materials because we think that they can attract much interests of students and have possibility to enhance educational efficiency derived from IDC Theory. In this paper, we also showed simple user study results about webbased VR education contents like VR tours. From some comments of the students, it was clarified that education contents supporting VR-goggles are more interesting than those not supporting VR-goggles.

As future work, we will try to ask teachers to use the proposed development framework for actually creating more web-based VR education contents to clarify the usefulness of the framework. Furthermore, we will create several educational VR contents with the framework and ask students to learn using the contents. After that, we will consult the students for more deeply evaluating educational efficiency of the contents.

#### Acknowledgements

This research was partially supported by JSPS KAKENHI Grant Number JP22H03705 and VISION EXPO Project of Kyushu University.

#### References

ICER (August 2023): https://www.icer.kyushu-u.ac.jp/en

- NOE (August 2023): https://dx.kyushu-u.ac.jp/ cms dir/uploads/2022/12/unit oe.pdf
- Okada, Y., Kaneko, K., Shi, W. (February 2023). Development Framework Using 360VR Cameras and Lidar Scanners for Web-Based XR Educational Materials Supporting VR Goggles. In: Barolli, L. (eds) Advances in Internet, Data & Web Technologies. EIDWT 2023. Lecture Notes on Data
  - Engineering and Communications Technologies, vol 161. Springer, pp. 401-412.
- Okada, Y., Kaneko, K., Shi, W. (July 2023). Development Framework for Web-based VR Tours and Its Examples. 27th International Conference Information Visualisation (IV), pp. 420-425.
- Okada, Y. and Oki, K.(March 2023). Web viewers for educational VR contents. IADIS International Conferences e-Society and Mobile Learning 2023 Proceedings, pp. 359-366.
- Okada, Y. and Tanaka, Y. (1995), IntelligentBox: a constructive visual software development system for interactive 3D graphic applications, Proc. of Computer Animation '95, IEEE CS Press, pp. 114-125.

Okada, Y. (2013), Web Version of IntelligentBox (WebIB) and Its Integration with Webble World, Vol. 372 of the series, Communications in Computer and Information Science, Springer, pp. 11-20. Unity (April 2023) : https://unity.com/

- THETA 360.biz (April 2023): https://www.theta360.biz/
- Matterport (April 2023): https://matterport.com/

Three.js (April 2023) : https://threejs.org/

# Development of TETPR: Technology-Enhanced Total Physical Response for Elementary Students to Learn English Vocabulary in Indonesia

#### **Riska SAPUTRA**

Graduate Institute of Network Learning Technology, National Central University, Taiwan riskasaputra@g.ncu.edu.tw

**Abstract:** This study explores the potential of Technology Enhanced Total Physical Response (TETPR) as an engaging approach to teaching English vocabulary to Indonesian students to support their English reading. TETPR integrates technology to enhance language acquisition, using the web, multimedia, AI, and remote. The study involves elementary students in Indonesia and employs a pretest-posttest control-group design. Data analysis will utilise an independent sample t-test to compare vocabulary acquisition between the experimental group taught with TETPR and the control group taught through traditional methods. The hypothesis predicts a significant difference in favour of the TETPR group. If successful, this study will contribute to IDC Theory by emphasising the role of interest-driven approaches on TETPR in improving students' English vocabulary acquisition.

Keywords: IDC, TELL, EFL, Total Physical Responses, Elementary Students

#### 1. Introduction

English is widely recognised as a global language (John & Yunus, 2021). Internationalisation is happening in all nations worldwide. (Tseng, 2012). More than 1.5 billion people are reportedly learning English (Dagvadorj, 2020). A student needs an extensive vocabulary to learn four English skills because vocabulary is essential for reading, speaking, writing, and listening (Mutalib, 2014). To introduce students to English book reading activities, Book Reading Centric (BRC) English begins with learning phonics, building basic vocabulary, and recognising common sentence patterns (Looi et al., 2023). However, Indonesian readers encounter a significant challenge: a deficiency of language aptitude when engaging with English books (Hidayati, 2018). Limited vocabulary can impede comprehension and make it difficult for readers to fully understand the text (Nanda & Azmy, 2020). According to Nation (2001), learning vocabulary in another language requires a systematic approach that considers the learners' needs and the best use of class time. TPR (total physical response) is an excellent technique for teaching vocabulary, particularly pre-reading or hearing words (Dixon, 2016). Total Physical Response (TPR) is a method for teaching language that emphasises the coordination of speech and action through physical activity (Astutik et al., 2019). TPR is particularly effective for teaching vocabulary because it engages students in physical movements and actions that assist them in comprehending and remembering new words (Nurhajati & Sulistyani, 2022). One of the benefits of using TPR for vocabulary learning is that it makes the learning process enjoyable and entertaining for students, resulting in greater motivation and engagement (Nuraeni, 2019). TPR can also be used in e-learning environments, where students can follow online instructions and perform physical movements (Dewi & Fatmawati, 2022). TPR can be effectively implemented when teaching young or adult learners (Nuraeni, 2019a; Nuraeni, 2019b).

This study aims to explore the potential of using Technology Enhanced Total Physical Response (TETPR) as a structured and engaging approach to teaching Indonesian elementary students to learn English vocabulary in the context of supporting their reading of English books.

#### 2. Literature Review

#### 2.1 IDC Theory on English Study

The interest-driven creator (IDC) theory is a design theory that aims to inform the design of future education in Asia (Chen et al., 2020). The theory posits that when students' learning is driven by interest, they can be engaged in knowledge creation (Looi et al., 2023). The Total Physical Response (TPR) is a fun and enjoyable method for students (Nuraeni, 2019b). Total Physical Response uses movement to teach a new language pleasantly and efficiently (Cook, 2022). TPR helps students to remember English words and expressions (Nuraeni, 2019a; Nuraeni, 2019b). Through multimodal learning, TPR activities using physical actions and commands can improve memory retention (Dewi & Fatmawati, 2022). TPR's fun factor can also improve the classroom atmosphere (Ningrum, 2021). This can boost students' English learning interest (Zur & Selfieni, 2022). TPR method uses psychology and kinesiology to help students' learn and remember language. Physical and verbal movement improves young students' vocabulary (Mariyam & Musfiroh, 2019). Students learn vocabulary better by linking it with physical motions (Ningrum, 2021). This method teaches language through physical activity by emphasising speech-action synchronisation (Dewi & Fatmawati, 2022).

In the context of English study, the IDC theory has been applied to improve English learning outcomes for undergraduate students in China (Wang & Chen, 2023).

A study designed a compulsory IDC-based English course in a blended setting for Chinese college students (Wang & Chen, 2023). The study aimed to enhance the listening and speaking proficiency of the students (Wang & Chen, 2023). In the IDC School in Taiwan, the BRC English approach is rooted in IDC Theory, prioritising immersive English book reading. Reader's Theatre activities or English drama performances connect reading with speaking and listening. Students create digital English picture books, aligning with IDC Theory's Creation Loop. Besides English classes, some subjects like arts and social studies are taught in English (Looi et al., 2023). Sulikatin (2023) has implemented a mini-lesson TPR activity supported by an online picture dictionary for learning English Listening and Vocabulary (LSV) at IDC Experimental School in Taiwan.

#### 2.2 Technology-Enhanced Total Physical Response

TPR is particularly effective for teaching vocabulary, as it helps students remember words and expressions (Nurhaiati & Sulistvani, 2022; Khakim & Anwar, 2020). It can be applied in large and small classes and is suitable for learners of all ages, including young learners and adults (Nuraeni, 2019). A technology-enhanced Total Physical Response (TPR) system in the classroom addresses the requirement for practical language learning approaches that engage students and encourage active engagement. TPR teaches vocabulary and grammar through haptic experiences (Swain & Lapkin, 2000). By using technology in TPR, educators improve learning and offer value. First, technology-enhanced TPR can support language acquisition using multimedia. Audio, video, and interactive features help students understand and motivate authentic language and cultural information (Nisa et al., 2022). Second, technology enables personalised and adaptive learning. Students can practice TPR at their own pace and receive rapid feedback using instructional software or applications (Jack & Higgins, 2019). This effective customised strategy lets students focus on their language learning needs and proceed at their own pace (Nisa et al., 2022). Technology-enhanced TPR fosters community and gives pupils social language practice (Richards & Rodgers, 2001). Several studies conducted TPR with technology, such as Wang et al. (2019), implemented Collaborative kinesthetic EFL learning with a collaborative total physical response. Kuo et al. (2013) evaluated the effects of the Embodiment-based TPR approach on student English vocabulary learning achievement, retention, and acceptance. Huang and Wang (2021) implemented an Artificial Intelligence learning approach through total physical response embodiment, teaching French vocabulary learning retention. Sulikatin (2023) has implemented a mini-lesson TPR activity supported by an online picture dictionary for learning English Listening and Vocabulary (LSV) at IDC Experimental School in Taiwan. The result showed a positive impact.

#### **3. Learning Environment**

#### 3.1 Technology Enhanced TPR for Learning English Vocabulary

Total Physical Responses is a structured learning strategy that facilitates language acquisition through commands, speech, and action (Astutik et al., 2019). This technology-enhanced TPR for learning English vocabulary for elementary students differs from previous studies. As shown in Table 1, this study has innovated from the technology, user, learning activity and language.

Comparison	This study	Kuo, <i>et al</i> . (2013)	Wang, et al. (2019)	Huang & Wang (2021)	Sulikatin (2023)
TPR with Technology	Ø	0	Ø	0	0
Picture Animation	Ø	8	⊗	8	0
TPR with Technology and Teacher	Ø	8	8	<b>I</b>	♦
Audio of Word from Computer/Web	0	8	8	0	0
Microsoft Kinect	8	<b>Ø</b>	Ø	8	8
Using Video for Learning	8	<b>I</b>	8	8	8
Collaborative Learning	Ø	8	<b>Ø</b>	8	8
Learn English	0	<b>I</b>	Ø	8	0
Audio Generated with AI	Ø	8	8	8	0
Participant (Elementary Student)	Ø	8	8	8	0
Friendly User Interface	<b>Ø</b>	8	8	8	8
Using Air Remote	<b>Ø</b>	8	8	8	8

Table 1. The Comparison of Technology Enhanced TPR



Figure 1. Learning Environment

As shown in Figure 1, this study uses technology to teach TPR for learning English vocabulary in the classroom. During a learning session, the teacher can utilise an air remote mouse to control the website. The website was modified to help Total Physical Response—AI speech generated by the text, and picture animation was developed as an online picture dictionary. Al was employed with children's sounds for primary or young learners. Al turned the text to audio. The text is the command inserted from the book into the web. Canva made the picture animation. Find two images to animate in Canva. Set the time in Canva. Export as a GIF file. Figure 1 displays the website's system architecture with three primary components: audio, text, and animation. The whiteboard must be linked to the internet to use this system. Google Site helps the teacher easily select and connect the learning content from an online picture dictionary. The projector shows the learning content on a whiteboard so the student can see the picture animation and text commands from the online picture dictionary. Teachers can easily control the web from afar with an air mouse.

#### 3.2 Learning Design

TETPR in the English class at the elementary school every week they have one time for learning English; it has 40 minutes for one hour of learning. The learning material will use a book named "Teaching English through Action: A Dynamic Way of Teaching the Basics of English to Students of all ages" by Berty Segal Cook. In this study, students will learn from units 1 to unit 2 (lesson 1 to lesson 17), in which the students learn about body parts, classroom items, colours, and numbers. Cook (2022) divided it into four steps to teach TPR in the lesson: New Word, Review Command, Command, and Novel Command. The experiment class (experimental group) will begin the lesson by showing new vocabulary or new words on the picture dictionary website, then show the review command to review last week's lesson, the command for the meeting today, and the novel command to review all learned lessons.

At every meeting, the teacher will teach one lesson. There are several steps to teach English vocabulary through this TETPR: 1) The teacher stands at the front of the classroom; 2) The teacher displays the content of the new word page; 3) The teacher instructs students to watch the picture animation, listen, and repeat; 4) The teacher directs students to perform the action word from the new word page; 5) The teacher shows the content of the review command/command/novel command pages; 6) The teacher asks students to listen and repeat the commands; 7) The teacher requests students to act out the commands at the front of the class. However, in the traditional class (control group), students learn English vocabulary by reading and memorising.

#### 4. Method

#### 4.1 Participants

The participants are 48 fourth-grade elementary students at SDS Aisyiyah Bengkalis, Riau, Indonesia. The school gave research permission to the student's parents for the research ethical. Student assignment data will be collected and recorded with photos and videos during the learning process.



Figure 2. Experimental design

#### 4.2 Instruments and Experimental Procedure

This research used pre-tests and post-tests from the Teaching English through Action book assessment. The researcher will administer 25 questions taken from the book. The correct response to each of the 25 questions on the pre-test is worth 4 points. The cumulative score is, therefore, 100. The test will evaluate the vocabulary skills of the students. In both the pre-test and post-tests, students will be written using pencil and paper.

This study will use quantitative methods to test TETPR's learning efficiency, which is the pretest-posttest control-group design will be used. It is an excellent experimental design, effectively controlling for rival hypotheses threatening the experiment's internal validity (Johnson & Christensen, 2014). The investigation will be implemented for three months (from July to October 2023).

#### 4.4 Data Analysis & Hypothesis

The study will employ an independent sample t-test. The Independent Samples t-test compares the means of two separate groups to answer the hypothesis. Before conducting further analysis, the researcher will assess whether the data is a normal distribution. The researcher intends to employ SPSS 27 to perform data analysis.

Null Hypothesis (H0): There is no significant difference in vocabulary acquisition between the group of students taught using TETPR (experimental group) and the group of students taught using traditional methods (control group).

Alternative Hypothesis (Ha): There is a significant difference in vocabulary acquisition between the group of students taught using TETPR (experimental group) and the group of students taught using traditional methods (control group).

#### 5. Conclusion

Suppose the study's findings can enrich the Interest-Driven Creation (IDC) Theory by emphasising the significance of interest-based reading choices, intrinsic motivation, active engagement, and the development of lifelong reading habits, especially if students demonstrate the positive impact of interest-driven approaches like Total Enhanced Total Physical Response (TETPR) on vocabulary acquisition and comprehension when reading English books. These findings align with IDC Theory's fundamental principles and promote an approach to English book reading centred on the learner.

#### Acknowledgements

The authors thank the National Science Technology Council of the Republic of China, Taiwan, for financial support (112-2410-H-008-020-MY3) to the Research Center of Science and Technology for Learning at the National Central University, Taiwan.

#### References

- Astutik, Y., Aulina, C. N. & Megawati, F. (2019). Total Physical Response (TPR): How is it used to teach EFL young learners?. *International Journal of Learning, Teaching and Educational Research, 18*(1), 92-103. <u>https://doi.org/10.26803/ijlter.18.1.7</u>
- Chen, W., Chan, T., Wong, L., Looi, C., Liao, C. C., Cheng, H. N., ... & Pi, Z. (2020). IDC theory: habit and the habit loop. Research and Practice in Technology Enhanced Learning, 15(1). <u>https://doi.org/10.1186/s41039-020-00127-7</u>
- Cook, B. S. (2020). Teaching English through Action: A dynamic way of teaching the basics of English to students of all ages (2nd ed.). North Holland: Arcos Publishers.
- Dagvadorj, N. (2020). A review of the effectiveness of cooperative learning in enhancing English language learner's fluency in Mongolia. International Journal of Trend in Research and Development, 7(4). http://www.ijtrd.com/papers/IJTRD22247.pdf
- Dewi, N.L. & Fatmawati, A. (2022). The Implementation of Total Physical Response (TPR) Method to Teach Vocabulary in E-Learning. *Jurnal Pendidikan Bahasa Inggris Proficiency*, *4*(1). http://doi.org/10.32503/proficiency.v4i1.2246

- Dixon, S. (2016). 100 TESOL Activities for Teachers Practical ESL/EFL Activities for the Communicative Classroom. Oregon, USA: Waygoose Press.
- Hidayati, K. (2018). Teaching writing to EFL learners: an investigation of challenges confronted by Indonesian teachers. Langkawi: Journal of the Association for Arabic and English, 4(1), 21. <u>https://doi.org/10.31332/lkw.v4i1.772</u>
- Huang, T.-H. & Wang, L.-Z (2021) Artificial intelligence learning approach through Total Physical Response embodiment teaching on French vocabulary learning retention. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2021.2008980
- Jack, C. & Higgins, S. (2019). Embedding educational technologies in early years education. *Research in Learning Technology*, 27(0). <u>https://doi.org/10.25304/rlt.v27.2033</u>
- John, E. & Yunus, M. M. (2021). A systematic review of social media integration to teach speaking. *Sustainability*, 13(16). <u>https://doi.org/10.3390/su13169047</u>
- Johnson, R. and L. Christensen (2014). *Educational Research Quantitative, Qualitative, and Mixed Approaches* (Fifth Ed.). California, USA: SAGE.
- Khakim, L. & Anwar, C. (2020). Improving students' vocabulary mastery through Total Physical Response learning method. Proceedings of the 2nd Social and Humaniora Research Symposium (SoRes, 2019). <u>https://doi.org/10.2991/assehr.k.200225.110</u>
- Kuo, F. R., Hsu, C-C., Fang, W-C. & Chen, N-S. (2013). The effects of Embodiment-based TPR approach on student English vocabulary learning achievement, retention and acceptance. Journal of King Saud University – Computer and Information Sciences, 26. <u>https://doi.org/10.1016/j.jksuci.2013.10.003</u>
- Looi, C.-K., Wong, S. L., Kong, S.-C., Chan, T.-W., Shih, J.-L., Chang, B., Wu, Y.-T., Liu, C.-C., Yeh, C. Y. C., Chen, Z.-H., Chien, T.-C., Chou, C.-Y., Hung, H.-C., Cheng, H. & Liao, C. C. Y. (2023). Interest-Driven Creator Theory: Case study of embodiment in an experimental school in Taiwan. *Research and Practice in Technology Enhanced Learning*, 18(23). <u>https://doi.org/10.58459/rptel.2023.18023</u>
- Mariyam, S. N., & Musfiroh, T. (2019). Total physical response (TPR) method improves English vocabulary acquisition of 5-6 years old children. *Tadris: Jurnal Keguruan dan Ilmu Tarbiyah, 4*(2), 257-264. https://doi.org/10.24042/tadris.v4i2.4071
- Mutalib, A. H. B. A., Kadir, R. B. A., Robani, R. B., & Majid, F. A. (2014). Vocabulary Learning Strategies among Malaysian TEVT Students in German-Malaysian Institute (GMI). *Procedia - Social and Behavioral Sciences, 123*, 361-368. <u>https://doi.org/10.1016/j.sbspro.2014.01.1434</u>
- Nanda, D. W. and Azmy, K. (2020). Poor reading comprehension issue in EFL classroom among Indonesian secondary school students: scrutinising the causes, impacts and possible solutions. *Englisia: Journal of Language, Education, and Humanities, 8*(1), 12. <u>https://doi.org/10.22373/ej.v8i1.6771</u>
- Nation, I. (2001). Learning Vocabulary in Another Language (Cambridge Applied Linguistics). Cambridge: Cambridge University Press. <u>https://doi.org/10.1017/cbo9781139524759</u>
- Ningrum, M.P. & Ma'rifah, U. (2021). Teaching English Vocabulary through TPR Method in Virtual Classroom: Students' Perception at Fifth Grade Thai Students. *Journal of English Teaching Literature and Applied Linguistics*, *5*(1). http://doi.org/10.30587/jetlal.v5i1.2158
- Nisa, K., Ramadhan, S. & Thahar, H. (2022). Student perceptions of technology needs in scientific writing. *Proceedings* of the 5th International Conference on Language, Literature, and Education (ICLLE-5 2022), 287-295. https://doi.org/10.2991/978-2-494069-85-5 31
- Nuraeni, C. (2019a). Using Total Physical Response (TPR) Method on Young Learners English Language Teaching. Metathesis Journal of English Language Literature and Teaching, (1). http://doi.org/10.31002/metathesis.v3i1.1223
- Nuraeni, C. (2019b). Promoting Total Physical Response (TPR) Method on Early Childhood English Language Teaching. Jurnal Penelitian Humaniora, 20(2). http://doi.org/10.23917/humaniora.v20i2.7144
- Nurhajati, D. & Sulistyani, S. (2022). Teaching Vocabulary through Total Physical Response to Elementary Students. *Journal of Development Research, 6*(2), 158-163. <u>https://doi.org/10.28926/jdr.v6i2.231</u>
- Richards, J., & Rodgers, T. (2001). Approaches and Methods in Language Teaching (2nd ed). Cambridge: Cambridge University Press. <u>https://doi.org/10.1017/cbo9780511667305</u>
- Sulikatin, S. (2023). Design and Implementation of Total Physical Response Mini-Lessons Supported by Online Picture Dictionary [Master Thesis]. National Central University.
- Swain, M. & Lapkin, S. (2000). Task-based second language learning: the uses of the first language. *Language Teaching Research, 4*(3), 251-274. <u>https://doi.org/10.1177/136216880000400304</u>
- Tseng, C. J. & Chao, C. T. (2012). Teaching culture in Taiwan's EFL classroom. *Procedia- Social and Behavioral Sciences*, 47. <u>https://doi.org/10.1016/j.sbspro.2012.06.910</u>
- Wang, F., Hwang, W-Y., Li, Y-H., Chen, P-T. & Manabe, K. (2019). Collaborative kinesthetic EFL learning with collaborative total physical response, *Computer Assisted Language Learning*, <u>https://doi.org/10.1080/09588221.2018.1540432</u>
- Wang, Y. & Chen, P. (2023). Interest-driven creator English course for developing Chinese private college students' listening and speaking proficiency based via blended setting. Asian-Pacific Journal of Second and Foreign Language Education. <u>https://doi.org/10.21203/rs.3.rs-2755448/v1</u>
- Zur, S. & Selfieni, S. (2022). Students' Perception of the Implementation on Total Physical Response (TPR) in English Teaching. *International Journal of Transdisciplinary Knowledge*, 3(2). <u>http://doi.org/10.31332/ijtk.v3i2.30</u>

# W14

# Quantification of Learning in Virtual Learning Environments

## Dynamic Facial Expression Recognition through Partial Label Learning and Federated Learning

Mohammad Alif DAFFA<sup>a\*</sup>, Manas GUPTA<sup>b</sup>, Hao CHEN<sup>b</sup> & Cheryl Sze Yin WONG<sup>b</sup>

<sup>a</sup>Singapore University of Technology and Design (SUTD), Singapore <sup>b</sup>Institute for Infocomm Research (I2R), Agency for Science Technology and Research (A\*STAR), Singapore {manas\_gupta, chen\_hao, cheryl\_wong}@i2r.a-star.edu.sg

- **Abstract:** In this paper, we present a model development pipeline for dynamic Facial Expression Recognition (FER) aimed at quantifying learning in virtual classrooms. The proposed pipeline involves the use of partial labels for training dynamic FER models, followed by the use of a self-supervised federated learning approach in further enhancing the model's performance on new subjects, addressing both continual learning needs and privacy concerns. This work ultimately contributes to advancing learning quantification in virtual classrooms by integrating partial label training and federated learning strategies for dynamic FER.
- Keywords: Federated learning, Privacy preserving deep learning, Facial expression recognition, Partial Label Learning

#### 1. Introduction

The assessment of student engagement consists of varied interactive components that can generally be divided into behavioral, affective and cognitive dimensions (Mandernach, 2015). In online learning environments, student engagement has typically been determined based on facial video recordings (Gupta et al., 2016; Dhall et al., 2020; Shen et al., 2022). Hence, dynamic facial expression recognition (FER) is an area of research that can be leveraged on for student engagement.

Despite the extensive research on both static and dynamic FER (Li & Deng, 2020), there still exists challenges that hinder the use of FER in applications. In general, FER suffers from subjective annotations and inherent similarity between emotion classes (Wang, Weijie & Sebe, Nicu & Lepri, Bruno., 2022). In addition, the variability in emotional expressions across subjects, as well as the lack of large dynamic FER datasets further enhances the difficulty of the task.

However, this raises yet another significant challenge which is the case of privacy. As FER data typically contains participant's faces, many would prefer not to have this data shared with others. A possible solution to this is the use of feature extraction methods to provide anonymity of training data. However, feature-extraction based methods generally achieve worse results when compared to end-to-end based methods (Tsalera et al., 2022). In order to overcome this issue, (Salman & Busso, 2022) developed a privacy preserving personalisation method for dynamic FER using federated learning. Their proposed method used a lightweight model that reduced the computation required for local training on an edge device. Furthermore, the use of federated learning allows for continual adaptation of the model without

comprising user privacy. The federated learning approach (Salman & Busso, 2022) allows for the local unsupervised training of a dynamic FER model, where local models are trained using pseudo labels generated by an image FER, and then used to update the central model via FedAvg. By combining this federated learning methodology with partial label learning, our work presents a generalized pipeline for developing a model which addresses the various challenges inherent to dynamic FER, which we evaluate on the CREMA-D dataset.

#### 2. Related Work

#### 2.1 FER

Based on the form of input data, the FER task can be further categorized as image facial expression recognition and video or dynamic facial expression recognition. In dealing with video FER, several approaches (Kahou, Michalski, Konda, Memisevic, & Pal, 2015; Lee, Choi, Kim, & Song, 2019; Lu et al., 2018) utilized convolutional neural networks architectures such as the VGG (Simonyan & Zisserman, 2014) or ResNet (He, Zhang, Ren, & Sun, 2016) to capture spatial features. These are then often paired with a Recurrent Neural Network (RNN) to handle the temporal features. The Long Short-Term Memory (LSTM) (Hochreiter & Schmidhuber, 1997) model is commonly used for this purpose.

#### 2.2 Federated Learning

Federated learning is a decentralized approach for machine learning where the training process takes place across multiple client devices rather than in a centralized server (Zhang et al. 2021). In this setup, each client holds its own local dataset and computes updates to a local model. Instead of sending raw data to a central server, only the model updates (gradients) are communicated back and aggregated by a central entity known as the aggregator. This in turn preserves privacy as raw data is not shared.

#### 3. Methodology

#### 3.1 Models

Similarly to (Salman & Busso, 2022), the proposed pipeline relies on two models - an image facial expression recognition model (IFER) and a dynamic video facial expression recognition model (VFER), both of which share the same feature extractor. For our feature extractor and IFER, we leverage the pretrained EmoNet model from (Toisoul et al., 2021), due to being trained for image facial expression recognition on the large diversity of faces in the AffectNet dataset, as well as having a small size of 2M parameters, making it ideal for performing inference on edge devices. We extract the final classifier module of the EmoNet as our IFER, taking only the 6 relevant emotions in the CREMA-D dataset, and the rest of the EmoNet model is used as the shared feature extractor between the IFER and VFER. Both the feature extractor and IFER are kept frozen throughout our experiments. On the other hand, the VFER consists of an LSTM model that takes in the sequences of the extracted features from the EmoNet backbone before predicting the overall emotion of a video sequence. While the VFER is used for dynamic FER, the IFER is used to generate pseudo labels necessary for self-supervised federated learning, by averaging the output of the IFER across every frame of the input video, and assigning the most probable class as the pseudo label if its maximum confidence, c passes the threshold of c=0.5. Otherwise, the sample is discarded and not used for training.



Figure 1. Model architectures for VFER and IFER

#### 3.2 Using probability distributions as labels

Previous studies (Salman and Busso, 2022) have utilized the majority rule for allocating the final annotation to each clip. In this work, we improve upon the annotation scheme by using partial labels to reflect the relative probabilities of the annotations in the final label. As per (Wang, Weijie & Sebe, Nicu & Lepri, Bruno., 2022), modeling FER as a Partial Label Learning (PLL) task addresses subjective annotations and inherent similarities among various facial expressions. In our implementation, we convert the distribution of votes into a probability distribution using the softmax transformation. To ensure the emotions with zero votes correspond to a probability of zero, we first map every instance of zero votes to a large negative number before applying the softmax function. We then train our VFER on these probability distributions using the Kullback-Leibler Divergence loss (Kullback & Leibler, 1951).

Formula for transforming vote distributions into target probability distribution:  $P(i) = Softmax(v_i')$  where  $v_i' = v_i$  if  $v_i \neq 0$  else  $v_i' = -10^{-9}$ 

#### 3.3 Federated Learning

During the inference phase, we deploy separate copies of the VFER to multiple test subjects, which we refer to as the local model. In this phase, we first train the local VFER model for a given test subject by using a subset of the given subject's videos. To simulate federated learning, we discard the associated labels and assign the local videos with pseudo labels generated by the IFER instead. The local VFER copies are each then trained on the respective subject's newly labeled videos. Once all the local copies of the VFER models have been fine-tuned, FedAvg is used to aggregate the local copies into a single central model. Following this, the updated central model is evaluated on the unseen test set and the results are reported.



*Figure 2.* Federated learning framework for proposed dynamic facial expression recognition

#### 4. Experimental evaluation

#### 4.1 Dataset split

The dataset chosen is the CREMA-D dataset (Cao et al., 2014), which contains 7,442 audio-visual video clips from 91 actors spread across six emotion classes - happiness, sadness, neutral, anger, fear and disgust. Each clip has 4 to 12 annotations with 95% of the clips having 8 or more annotations. We select labels produced solely based on visual data, aligning with our model's exclusive use of video modality. This reduces noise from audio information unavailable to our model. To prevent data leakage and allow for an effective simulation of federated learning, we group video samples by actors before splitting the data into train, validation and test sets, ensuring that samples from the same actor only appear in one set. We also aim to keep the gender distribution in each set balanced. This results in a final split of 67 actors in our training set, 12 actors in our validation set, and 12 actors in our test set. For federated learning, the test set is split into 2 equal segments, with each segment containing half of the video samples of every actor. One segment will then be used as the adaptation set for federated learning, and the other is used as our test set for evaluating the adapted model.

#### 4.2 Implementation

For initial training of the VFER model, we use the Adam optimizer with a learning rate of 0.0001 for 20 epochs. On each epoch, the model is evaluated on our validation set to detect for overfitting. We then save the VFER weights on the epoch which achieves the lowest validation loss. For tuning the hyperparameters for federated learning, we evaluated the performance of local models on the validation set after each epoch. This led to the observation that locally trained models experience severe catastrophic forgetting (McCloskey & Cohen, 1989), immediately demonstrating deteriorating validation loss from the first epoch. As a result, we found the ideal

hyperparameters for local adaptation to be only a single epoch of training with a low learning rate of 0.00001.

#### 4.3 Results

We present the comparison of results on our proposed implementation of partial label learning on the CREMA-D dataset in Table 1. As shown in Table 1, using partial labels for the training of our initial VFER resulted in an improved model across measures of average precision, recall and f1 score on the testing set. Among the 6 emotions, the model trained on the partial labels achieved a higher average score in classifying Neutral, Happiness, Fear and Anger while underperforming in the classifying of Sadness and Disgust, emotions which occur the least frequently in the CREMA-D dataset.

Emotion	Precision(	%)	Recall(%)	Recall(%)		F1_Score(%)	
	Discrete	Partial labels	Discrete	Partial labels	Discrete	Partial labels	
Neutral	72.3	65.2	74.5	75.2	73.4	69.8	
Happiness	81.6	88.9	94.1	94.1	87.4	91.4	
Sadness	37.5	32.8	38.9	38.9	38.2	35.6	
Fear	58.2	51.3	44.4	55.6	50.4	53.3	
Disgust	61.0	75.0	52.2	43.5	56.3	55.0	
Anger	49.4	62.9	54.7	52.0	51.9	56.9	
Micro-mean	63.0	64.7	63.4	63.6	62.9	63.3	
Macro-mean	60.0	62.7	59.8	59.9	59.6	60.4	

Table 1: VFER trained using discrete vs partial labels

Table 2 presents the effect of federated learning on the model initially trained on partial labels. Our results demonstrate that federated learning further enhances the performance of the initial VFER trained on partial labels with improvements in average precision, recall and fl score.

Emotion	Precision(%)		Reca	Recall(%)		F1_Score(%)	
	Before	After	Before	After	Before	After	
Neutral	65.2	67.1	75.2	77.4	69.8	71.9	
Happiness	88.9	89.9	94.1	94.1	91.4	92.0	
Sadness	32.8	35.7	38.9	37.0	35.6	36.4	

Table 2: VFER Performance Before vs After Federated Learning

Fear	51.3	52.0	55.6	54.2	53.3	53.0
Disgust	75.0	75.5	43.5	53.6	55.0	62.7
Anger	62.9	60.0	52.0	52.0	56.9	55.7
Micro-mean	64.7	65.4	63.6	65.2	63.3	64.9
Macro-mean	62.7	63.4	59.9	61.4	60.4	61.9

#### 5. Conclusion

In conclusion, our study introduces a training pipeline for dynamic Facial Expression Recognition (FER). We first verify the effectiveness of using a partial label learning paradigm to address subjective labels and intra-class similarities between facial expressions (Wang, Weijie & Sebe, Nicu & Lepri, Bruno., 2022) in the context of dynamic FER, demonstrating overall improvement in our initial dynamic FER model. Subsequently, to address the need for continual learning while preserving user privacy, we validate the use of a previously established federated learning technique (Salman & Busso, 2022), demonstrating even further improvements on our previously trained model's performance.

Ultimately, by leveraging the methodologies of both partial label learning and self-supervised federated learning, we establish a formal model development pipeline and demonstrate its effectiveness in addressing the challenges inherent to FER, providing a promising technique for practical applications in the quantification of learning in virtual classrooms.

#### Acknowledgements

We would like to thank the Artificial Intelligence, Analytics & Informatics (AI3) and Institute for Infocomm Research (I2R), A\*STAR for supporting this work.

#### References

- Kahou, S. E., Michalski, V., Konda, K., Memisevic, R., & Pal, C. (2015). Recurrent neural networks for emotion recognition in video. In ICMI (pp. 467–474).
- Lee, M. K., Choi, D. Y., Kim, D. H., & Song, B. C. (2019). Visual scene-aware hybrid neural network architecture for video-based facial expression recognition. In FG (pp. 1–8).
- Lu, C., Zheng, W., Li, C., Tang, C., Liu, S., Yan, S., & Zong, Y. (2018). Multiple spatio-temporal feature learning for videobased emotion recognition in the wild. In ICMI (pp. 646–652).
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In CVPR (pp. 770–778).
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.
- Antunes, R. S., André da Costa, C., Küderle, A., Yari, I. A., & Eskofier, B. (2022). Federated learning for healthcare: Systematic review and architecture proposal. ACM Transactions on Intelligent Systems and Technology (TIST), 13(4), 1-23.
- Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., ... & Roselander, J. (2019). Towards federated learning at scale: System design. Proceedings of machine learning and systems, 1, 374-388.

- Cao, H., Cooper, D. G., Keutmann, M. K., Gur, R. C., Nenkova, A., & Verma, R. (2014). CREMA-D: Crowd-sourced Emotional Multimodal Actors Dataset. IEEE transactions on affective computing, 5(4), 377–390. https://doi.org/10.1109/TAFFC.2014.2336244
- Dhall, A., Sharma, G., Goecke, R., & Gedeon, T. (2020, October). Emotiw 2020: Driver gaze, group emotion, student engagement and physiological signal based challenges. In Proceedings of the 2020 International Conference on Multimodal Interaction (pp. 784-789).
- Gupta, A., D'Cunha, A., Awasthi, K., & Balasubramanian, V. (2016). Daisee: Towards user engagement recognition in the wild. arXiv preprint arXiv:1609.01885.
- Li, S., & Deng, W. (2020). Deep facial expression recognition: A survey. IEEE transactions on affective computing, 13(3), 1195-1215.
- Long, G., Tan, Y., Jiang, J., & Zhang, C. (2020). Federated learning for open banking. In Federated Learning: Privacy and Incentive (pp. 240-254). Cham: Springer International Publishing.
- Mandernach, B. J. (2015). Assessment of student engagement in higher education: a synthesis of literature and assessment tools. Int. J. Learn. Teach. Educ. Res. 12, 1–14. doi: 10.1080/02602938.2021.1986468
- McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communicationefficient learning of deep networks from decentralized data. In Artificial intelligence and statistics (pp. 1273-1282). PMLR.
- McMahan, H. B., Moore, E., Ramage, D., & y Arcas, B. A. (2016). Federated learning of deep networks using model averaging. arXiv preprint arXiv:1602.05629, 2, 2.
- Pandya, S., Srivastava, G., Jhaveri, R., Babu, M. R., Bhattacharya, S., Maddikunta, P. K. R., ... & Gadekallu, T. R. (2023). Federated learning for smart cities: A comprehensive survey. Sustainable Energy Technologies and Assessments, 55, 102987.
- Salman, A., & Busso, C. (2022, November). Privacy preserving personalization for video facial expression recognition using federated learning. In Proceedings of the 2022 International Conference on Multimodal Interaction (pp. 495-503).
- Shen, J., Yang, H., Li, J. et al. Assessing learning engagement based on facial expression recognition in MOOC's scenario. Multimedia Systems 28, 469–478 (2022). https://doi.org/10.1007/s00530-021-00854-x
- Tsalera, E., Papadakis, A., Samarakou, M., & Voyiatzis, I. (2022). Feature Extraction with Handcrafted Methods and Convolutional Neural Networks for Facial Emotion Recognition. Applied Sciences, 12(17), 8455. <u>https://doi.org/10.3390/app12178455</u>
- Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021). A survey on federated learning. Knowledge-Based Systems, 216, 106775.'
- Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. The Annals of Mathematical Statistics, 22(1), 79–86. https://doi.org/10.1214/aoms/1177729694
- McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The Sequential Learning Problem. *Psychology of Learning and Motivation*, 109–165. https://doi.org/10.1016/s0079-7421(08)60536-8
- Toisoul, A., Kossaifi, J., Bulat, A., Tzimiropoulos, G., & Pantic, M. (2021). Estimation of continuous valence and arousal levels from faces in naturalistic conditions. *Nature Machine Intelligence*, 3(1), 42–50. <u>https://doi.org/10.1038/s42256-020-00280-0</u>

# An Empirical Investigation on Google Classroom: Use Behaviour of Malaysian School Teachers

Priscilla MOSES<sup>a\*</sup>, Jacqueline Chung Ling LAU<sup>b</sup>, Phaik Kin CHEAH<sup>c</sup>, Phoebe Soong Yee YAP<sup>d</sup>, Mas Nida Md KHAMBARI<sup>e</sup>, Su Luan WONG<sup>f</sup>

<sup>abcd</sup>Universiti Tunku Abdul Rahman, Malaysia <sup>ef</sup>Universiti Putra Malaysia, Malaysia \*priscilla@utar.edu.my

Abstract: Google Classroom (GC) has become a prominent online learning platform in Malaysia since the government's official announcement and later sees its boost in usage again during the COVID-19 pandemic due to the implementation of emergency remote teaching. The literature has unveiled that the mean score pertaining to technological knowledge in the context of GC usage was the lowest. Therefore, this study aims to investigate the level of GC use and the significant influence of facilitating conditions, habit, and behavioural intention on GC use behaviour among Malaysian teachers. A total of 37 secondary school teachers participated in an online quantitative survey consisting of 22 five-point Likert scale items. Descriptive statistics and standard regression analysis were used to analyse the data. Findings found that the extent of agreement and utilisation of GC during the pandemic were only moderate, requiring increased focus as a precautionary measure for potential future scenarios, such as another pandemic, where remote learning becomes the sole viable option. Results also suggested habit ( $\beta$  = 1.128, p < .0005) as the most significant factor that influences Malaysian teachers' GC use behaviour. Thus, future initiatives in boosting the intentions of using GC among Malaysian teachers need to be considered.

Keywords: Google Classroom, Use Behaviour, Secondary Teachers, Malaysiaii

#### 1. Introduction

Before the outbreak of the pandemic, the Malaysian Ministry of Education (2019) officially announced the use of Google Classroom (GC) as the main online teaching platform for national schools. Subsequently, the outbreak of the pandemic has driven the implementation of emergency remote teaching, where teachers were obligated to fully adopt GC as their main medium of instructional delivery. It is undeniable that the application of e-learning has improved due to the outbreak of the pandemic (Hoque et al., 2020; Mokhtar, 2020). However, in the study by Zakaria et al. (2021), they indicated that educators recognise the benefits of Google Classroom as an effective online learning platform but obstacles arise due to limited knowledge and facilities making the integration of online learning through Google Classroom more difficult as compared to face-to-face teaching. A recent study by Piaralal et al. (2023) also revealed that the mean score for technological knowledge in relation to the use of GC was the lowest. Consequently, they concluded that teachers should prioritize upgrading their technological proficiency to effectively promote the intention to use Google Classroom (Piaralal et al., 2023).

The sudden shift from traditional face-to-face classrooms to the new norm of sole online teaching and learning indisputably created new challenges for educators around the globe, not to mention those from underrepresented countries where information and communication technology (ICT) growth is only at the emerging stage. Thus, this study focuses on three constructs, which are facilitating conditions, habit, and behavioural intention to investigate the GC use behaviour among teachers. These teachers as the actual implementers of GC can reveal crucial information after the execution of the platform for more than two years. The key information on GC use from this study can add great significance to the existing body of knowledge regarding the use of GC in Malaysia.

#### 1.1.1 Main research objective

Addressing the problems stated earlier, the main objective of this paper is to determine the level of agreement of the three constructs, the level of GC use behaviour and how much variance in GC use behaviour can be explained by facilitating conditions (FC), habit (HB), and behavioural intention (BI).

#### 1.1.2 Specific research objectives

This paper thus sets out to achieve the following specific research objectives:

- 1. To determine the level of agreement on facilitating conditions, habit, behavioural intention of using GC among secondary school teachers.
- 2. To determine the level of GC use behaviour among secondary school teachers.
- 3. To investigate the influence of facilitating conditions on GC use behaviour among secondary school teachers.
- 4. To investigate the influence of habit on GC use behaviour among secondary school teachers.
- 5. To investigate the influence of behavioural intention on GC use behaviour among secondary school teachers.

#### 1.1.3 Research questions

This paper aims to answer the following research questions:

- 1. What is the level of agreement on facilitating conditions, habit and behavioural intention of using GC among secondary school teachers?
- 2. What is the level of GC use behaviour among secondary school teachers?
- 3. Is there a significant influence of facilitating conditions on GC use behaviour among secondary school teachers?
- 4. Is there a significant influence of habit on GC use behaviour among secondary school teachers?
- 5. Is there a significant influence of behavioural intention on GC use behaviour among secondary school teachers?

#### 1.1.4 Hypotheses

Aligned with the objectives of the study, the following hypotheses were proposed:

H1: There is a significant influence of facilitating conditions on GC use behaviour among secondary school teachers.

H2: There is a significant influence of habit on GC use behaviour among secondary school teachers.

H3: There is a significant influence of behavioural intention on GC use behaviour among secondary school teachers.

#### 2. Literature Review

Generally, it has been found that past studies on educational technologies and online learning platforms like GC were often focusing on higher educational settings (Basher, 2017; Dash, 2019; Dewi et al., 2022; Kumar & Bervell, 2019; Mafa, 2018; Olumorin et al., 2022; Shaharanee et al., 2016). Besides, most of the past studies also focused on the effectiveness and impacts of these online learning platforms (Basher, 2017; Dash, 2019; Shaharanee, 2016), limited studies were found to focus on the actual use of GC and applications.

#### 2.1 The utilisation of Google Classroom in Malaysia

Before GC was launched by Google Apps for Education in 2014 (Piaralal et al., 2023), the government introduced the Frog Virtual Learning Environment (VLE). Frog VLE is an online based system, offering various features such as uploading documents, videos, and chatroom for teachers to communicate with their students (Rahayu et al., 2019). Today, GC is a part of the Google Workspace for Education (formerly known as G Suite in Education). It is an application that is user-friendly that enhances teachers' work performance by fostering active learning during online teaching and learning (Ahmad et al., 2022). In addition, GC also allows teachers to share teaching resources such as pictures and videos effortlessly, besides distributing assignments, carrying out quizzes, and doing grading works online (Chung, 2022; Rosni et al., 2022). Following the discontinuation of the Frog Virtual Learning Environment, GC which offers a variety of benefits, further presents itself as a superior platform and has become a prominent platform to schools in Malaysia.

#### 2.2 Common challenges encountered during online teaching and learning

Often in the process of online teaching and learning, poor Internet connectivity poses the main challenge to both educators and students (Cheok et al., 2017). Educators and students from rural areas were reported as facing more challenges compared to those in urban areas to conduct online teaching and learning (Nation, 2019). On the other hand, Cheok et al. (2017) and Lepp et al. (2021) mentioned that some teachers viewed lack of time as the major concern in the implementation of e-learning. The teachers claimed that the overload of work in school had restrained their intentions of exploring and learning the system further (Cheok et al., 2017). Teachers have expressed the need to be trained so as to overcome the challenges of using educational technology in online teaching and learning. The issue of inadequate ICT skills among teachers has also been highlighted by Nidup (2022) as one of the prominent barriers to teaching online. In fact, both Cheok et al. (2017) and Rauf and Swanto (2020) have indicated that without an appropriate training environment, access to the necessary technology, and support from relevant authorities, teachers would lack the self-confidence to successfully incorporate educational technologies within their classrooms.

In addition to this, Rauf and Swanto (2020) have highlighted that there are several other aspects that may affect their use of using educational technologies in online classroom teaching and learning. Therefore, this study sets out to investigate the level of GC use and the significant influence of facilitating conditions, habit, and behavioural intention on GC use behaviour among Malaysian teachers.

#### 2.3 Theoretical and Conceptual Frameworks

This study adapted Venkatesh et al. (2012)'s Unified Theory of Acceptance and Use of Technology 2 (UTAUT 2) by narrowing the focus solely on the three constructs (facilitating conditions, habit and behavioural intention) that directly contribute to the use behavior of GC among teachers. Venkatesh et al. (2003) explained facilitating conditions were initially a core construct from the Model of PC Utilisation (MPCU) which was derived mainly from the theory of human behaviour. Facilitating conditions were defined as objective factors in the environment that observers agree which make an act easy to accomplish (Venkatesh et al., 2003). In this study, facilitating conditions is referred to the provision of support such as Internet connection, modern technology, technical assistants, computers, and laptops (Malaysian Ministry of Education, 2013) for Malaysian teachers to utilise GC during the pandemic.

On the other hand, habit was defined as the act of automatically performing behaviours because of the learning that has occurred (Venkatesh et al., 2012). Nikian et al. (2013) have also stated that teachers' repetitive use of technology would be a form of habitual routine in daily teachings. In this study, habit is referred to Malaysian teachers' routine use of GC automatically for online teaching and learning during the pandemic.



Figure 1. Conceptual framework of this study.

Behavioural intention to use was initially defined by Fishbein and Ajzen (1980) as the degree of likelihood that a person makes use of a certain technology or application. It has also been stated by Venkatesh et al. (2012) that behavioural intention plays a role as the key predictor of certain technology use. In this study, behavioural intention is referred to as whether Malaysian teachers intend to use GC in the future.

Finally, use behaviour has been referred to as the actual usage of a certain technology or application by Venkatesh et al. (2012). In this study, use behaviour is referred to as the consistency of using GC for online teaching and learning among Malaysian teachers during the pandemic. Consequently, with all the constructs and variables defined, the research framework of this study is depicted in Figure 1, along with the hypotheses that were proposed to align with the research objectives of the study.

#### 3. Method

This study employed a quantitative approach to answer all the research questions of the study. An online questionnaire was used to gather the data needed in investigating the influence of facilitating conditions, habit, and behavioural intention on the use behaviour of GC among Malaysian teachers. Besides, the level of agreement on facilitating conditions, habit, behavioural intention of using GC among the teachers and the level of GC use behaviour among the teachers during the pandemic were also determined by using descriptive statistics (research questions 1 & 2). Multiple regression was employed as the inferential statistics method to address research questions 3, 4, and 5.

#### 3.1 Research instrument, sampling, and data collection

In this study, an online survey tool known as Qualtrics was used to build and distribute the online questionnaire survey. The questionnaire consisted of two sections: Section A for the collection of respondents' demographic, while Section B comprised 22 items measuring the facilitating conditions, habit, behavioural intention, and use behaviour of GC respectively (https://tinyurl.com/ympr6tsx). Each item was measured on a five-point Likert scale.

As for the data collection, 56 responses were collected initially through Qualtrics. After the cleaning of data by removing partial responses and speeding responses, there was a total of only 37 responses which were valid for further analysis. As suggested by Maxwell (2000) and Statistics Solutions (2021, August 2), the rule of thumb for getting an adequate sample size is following the ratio of subjects to predictors be at least 10:1. Thus, the sample size of 37 subjects to three predictors (facilitating conditions, habit, and behavioural intention) of this study is sufficient to obtain valid and actual findings.

#### 3.2 Reliability and validity of the instrument

The Cronbach's Alpha test was used to access the reliability of the questionnaire used to collect data for this study. As depicted in Table 1, the Cronbach's Alpha values for all the variables ranged from .783 to .974, which were considered ideal as all of them were more than .7 (Taber, 2018), proving that the instrument used in this study had an ideal internal

consistency and reliable. In addition, a panel of two experts in the field of Malaysian educational technology have also reviewed the questionnaire to ensure its content validity.

Variable	Cronbach's Alpha	Cronbach's Alpha based on Standardized Items	Number of Items
Facilitating conditions	.783	.778	5
Habit	.944	.945	5
Behavioural Intention	.971	.972	5
Use Behaviour	.974	.975	7

Table 1. Reliability statistics of the predictor variables and dependent variable

#### 4. Results and Findings

#### 4.1 Demographic information

In this study, there were 11 males (29.73%) and 26 females (70.27%). Figure 2 depicts the age distribution of the respondents, the mean score obtained is 36.30 and the mode is 26 with the standard deviation of 9.70. Moreover, 13 teachers had more than 15 years of teaching experience, followed by 9 teachers with 1 to 5 years of teaching experience, 6 teachers with 11 to 14 years of teaching experience, 5 teachers with 4 to 10 years, and only 4 teachers had less than a year of teaching experience.



Figure 2. Age distribution of the respondents.

#### 4.2 Findings on the level of agreement for each variable

 Table 2. Indication of level by categorization of mean scores

Category of Scores	Level
Below 2.33	Low
Between 2.33 and 3.66	Moderate
Above 3.66	High

Table 3. Level of agreement on FC	, HT, BI and level of Google	Classroom use behaviour
-----------------------------------	------------------------------	-------------------------

Variable	Number of Items	Mean of Items	Level
Facilitating conditions	5	3.492	Moderate Agreement
Habit	5	3.259	Moderate Agreement
Behavioural Intention	5	3.400	Moderate Agreement
Use Behaviour	7	3.459	Moderate Use

The scale to indicate the level of agreement with the items according to their respective means was adapted from Moses et al. (2010). The level of agreement of the variables was quantified on five different levels using a five-point Likert scale, ranging from "Strongly disagree", "Disagree", "Neither agree nor disagree", "Agree" and "Strongly agree". The items in each independent variable ranged from "Strongly disagree" – 1 point to "Strongly agree" – 5 points. The data collected is equally clustered into three levels of agreement namely low agreement, moderate agreement, and high agreement as presented in Table 2. From the findings, all the variables showed a moderate level of agreement (Table 3) as their means were between 2.33 and 3.66. "Facilitating conditions" has the highest mean which is 3.492; the lowest mean score is "habit" with a mean of 3.259.

Meanwhile, the level of GC use behaviour was quantified on five levels ranging from "Never", "Sometimes", "Occasionally", "Most of the time" and "Always". The data gathered showed a moderate level of use as the mean was 3.459, also between 2.33 and 3.66.

#### 4.3 Findings from the standard multiple regression

Standard multiple regression was performed to determine if the predictor variables (facilitating conditions, habit, and behavioural intention) have had an influence over the use behaviour among teachers during the pandemic. The following section reports the multiple regression findings, particularly on the model summary, Analysis of Variance (ANOVA), coefficients, and normal p-p plot as suggested by Hinton et al. (2014) as these are sufficient in reporting multiple regression reports.

As shown in Table 4, the R-value was reported as .868 while the R2 was at .753. When interpreting the findings from the model summary, R2 is often referred to as it explains the extent of variance that occurred among the variables being investigated (Hinton et al., 2014). In this case, as the R2 was reported as .753, it meant that 75.3% of the variance that occurs among variables can be significantly explained. The ANOVA table on the other hand, as depicted in Table 5, has successfully reported the occurrence of the three predictor variables, facilitating conditions, habit, and behavioural intention as existent when using GC. This was proven to be true as the df value was reported at three in the regression and these are significant at p < .0005.

Model	R	R Square	Adjusted I	R Square	Std. Error of the Estimate
1	.868a	.753	.730		.71041
Notes a	Prodictors: ((	Constant) Mean	BL Maan EC	Moon HB	

Table 4. Model Summary

*Notes.* a. Predictors: (Constant), Mean\_BI, Mean\_FC, Mean\_HB b. Dependent Variable: Mean\_UB

Table 5. Analysis of Variance (ANOVA)

Мо	odel	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	50.739	3	16.913	33.512	.000b
	Residual	16.654	33	.505		
	Total	67.393	36			
A.L. (	-					

*Notes.* a. Dependent Variable: Mean\_UB

b. Predictors: (Constant), Mean\_BI, Mean\_FC, Mean\_HB

However, in the coefficients as shown in Table 6, two interesting findings were obtained. When reporting the coefficients, Wagner (2015) has stated that unstandardized coefficients are referred to explain the effects of predictor variables on dependent variables. In the coefficients table at unstandardized coefficients, prominently habit has the highest beta value (1.393), which meant that for every occurrence of habit, the use behaviour increases by 1.393. Contrary to habit, the beta value ( $\beta$ ) for behavioural intention was found to be at negative, as it was reported as -.397, which indicated that for every occurrence of behavioural intention, the use behaviour declines by .397. As for the facilitating conditions, although  $\beta$  is recorded at .090, the p-value is .605 which is not significant.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		β	Std. Error	β	_	
1	(Constant)	044	.511		086	.932
	Mean_FC	.090	.172	.056	.522	.605
	Mean_HB	1.393	.229	1.128	6.085	.000
	Mean_BI	397	.225	343	-1.766	.087

#### Table 6. Coefficients

Notes. a. Dependent Variable: Mean\_UB

Drawing the results from multiple regression, the research hypotheses, H1 and H3 were not supported as both facilitating conditions and behavioural intention have no significant influence on the GC use behaviour among the teachers during the pandemic. Meanwhile, only the second research hypothesis, H2 is supported, and it is proven that habit significantly influenced the use behaviour of GC among the teachers during the pandemic.

Table 7. Summary of results

Hypothesis	Description	P-value	Result
H1	Facilitating conditions → Use Behaviour	.605	Not Supported
H2	Habit $\rightarrow$ Use Behaviour	.000*	Supported
H3	Behavioural Intention → Use Behaviour	.087	Not Supported
Natas to 1 OF			

*Notes.* \**p* < .05

#### 5. Discussion and conclusion

Descriptive statistics reveal that all three levels of agreement on each variable, with facilitating conditions, habit, and behavioural intention respectively, have mean values between 2.33 and 3.66, which is at a moderate level. Given that GC has only been fully utilised since the eruption of the pandemic, such a moderate level of agreement and use behaviour points out that there might be a need in providing training to teachers. This is supported by the statement of Cheok et al. (2017) that teachers have expressed the need to be trained to overcome the challenges of using educational technology for online teaching and learning.

Based on the findings, it was found that facilitating conditions and behavioural intention did not significantly influence the GC use behaviour among secondary school teachers. This can be attributed to several reasons why these two independent variables have an insignificant influence on GC usage. In this study, facilitation conditions refer to the provision of support such as Internet connection, modern technology, technical assistants, computers, and laptops provided to support the use of GC. If the teachers lacked access to reliable internet connections, technical support, or appropriate devices, it could have definitely hindered their capability to effectively use GC for their lessons. In addition to that, inadequate professional development opportunities related to GC could also lead to a lack of facilitating conditions or support among the teachers to successfully implement GC in their teaching and learning. Therefore, the facilitating conditions did not significantly influence the use behaviour of GC could be due to the lack of facilitating conditions that were available to the teachers, as they were working from their respective homes during the pandemic.

As for behavioural intention, it refers to whether Malaysian teachers plan to use GC in their teaching practices in the future. It can also be said as the teachers' willingness to adopt and utilise GC in their lessons but if the teachers are resistant to change or if they did not perceive the benefits of using GC, their behavioural intention to use GC might be low. Some teachers could have concerns about the effectiveness of using GC in achieving teaching objectives and students' learning outcomes. Another reason could be due to the preference that the teachers have for traditional teaching methods compared to online teaching and learning could also negatively influence their behavioural intention to utilise GC. These could
be some of the reasons that influenced the teachers' decision to not fully embrace or use GC for teaching and learning purposes.

On the other hand, this research has shown that the use of GC in classrooms during the pandemic may have potentially become a normalised routine among teachers due to the significance of habit found in the study, which subsequently indicates that it has become a habitual routine among the teachers. Not only that, teachers were compelled to utilise GC as it served as the sole platform or method available to establish connectivity with students to conduct their lessons during the outbreak of the pandemic. Yet, even though the utilisation of GC in classrooms was mandatory, it does not necessarily imply that all teachers are equally proficient or knowledgeable in using the platform (Zakaria et al., 2021; Piaralal et al., 2023).

Venkatesh et al. (2003) have previously emphasized that only when facilitating conditions, habit, and behavioural intention occur, a person would be most likely to use certain technology in the given setting. Narrowing these findings to the Malaysian context, it is thus arguable for teachers who lack facilitating support, who are reluctant to routinized GC as part of their lessons or who show less likelihood of using GC. This could possibly cause the GC use behaviour of teachers consequently to be lesser.

To summarize, among all the predictor variables, the habit was the best predictor of the use behaviour of GC among the Malaysian. In addition, there may be other factors that have influenced the use behaviour of GC as shown by the variance, R2, as the three constructs only explained 75.3% of the GC use behaviour among the teachers, there was 24.7% variance that was not explained. It could be other potential variables which were not included in this study that have affected the GC use behaviour among the teachers.

As part of the Malaysian Ministry of Education's (2013) aspiration to provide quality education to students across the nation, when teachers display disinclination in embracing educational technology within classrooms, it not only hinders their ability to achieve the ministry's goals but also prevents Malaysia from maintaining its educational standards in the global context. Hence, the findings from this study found that the extent of utilisation of GC during the pandemic was not fully embraced which requires increased focus as a precautionary measure for potential future scenarios, such as another pandemic, where remote learning becomes the sole viable option.

# 6. Implications of the study

This study has shown a number of pertinent aspects of the actual usage of GC among Malaysian teachers. One of the most prominent findings is that Malaysian teachers do not have positive intentions in the continuous usage of GC. However, they were using GC mostly due to their habitual routine as a part of their regular teaching practice. This is further supported by Awang et al. (2019) who stated that proper planning and implementation would lessen challenges, thereby alleviating teachers' resistance toward GC. It is believed that with continuous support from teachers, administrators, and the relevant government agencies to enhance the quality of online teaching and learning experiences, the initiatives of achieving the aspiration of Education 4.0 will be accomplished in the near future (Hogue et al., 2020; Nikian et al., 2013). The Malaysian Education Blueprint 2013-2025 has stated that a total of 10,000 public primary and secondary schools in Malaysia will be provided with 4G Internet connectivity and a Virtual Learning Environment (VLE) system to boost the quality of educational ICT innovation (Malaysian Ministry of Education, 2013). Further, the Malaysian Ministry of Education (2013) has also stated that an environment where teachers are constantly trained on using ICT would create a learning environment that has high quality and brings great impact on the overall learning achievements. The findings of this study can be considered for the planning of future initiatives in boosting the uptake intentions of e-learning among teachers in Malaysia.

# 7. Limitations and recommendations

This study, however, may be biased as the data collection was gathered only from the Malaysian secondary school teachers' point of view. In addition, merely employing a quantitative method may be inadequate in delivering generalizable and comprehensive data that truthfully describes the scenario being examined. Future studies should employ a more rigorous research method, include qualitative methods to enrich the quality of research by generating rich, interpretive findings from the teachers on the use of educational technology during the pandemic, and also recruit a larger sample size to improve the overall reliability of the findings.

# Acknowledgements

The research was funded by the Ministry of Higher Education, Fundamental Research Grant Scheme (FRGS), FRGS/1/2021/SSI0/UTAR/02/2. We express our deepest gratitude towards the government for funding this research project.

# References

- Ahmad, R. R., Hassan, Z., & Wahab, N. H. A. (2022). Kemudahan penggunaan teknologi Google bilik darjah dalam kalangan guru sekolah menengah di Malaysia. *Sains Humanika*, *14*(2), 69–77. <u>https://doi.org/10.11113/sh.v14n2.1936</u>
- Awang, H., Aji, Z. M., Osman, W. R. S., Nasir, A. A., Deli, M. M., & Hamat, W. Y. W. (2019). Virtual Learning Environment (VLE) implementation strategy: An analysis of practicality for Google Classroom implementation in Malaysian schools. *Journal of Educational Research and Indigenous Studies*, 2(1). <u>https://www.researchgate.net/profile/Aidayani-Abdul-Nasir/publication/335834482\_Virtual\_Learning\_Environment\_VLE\_Implementation\_Strategy\_An\_Analysis\_of\_Practicality\_for\_Google\_Classroom\_Implementation\_in\_Malaysian\_Schools/links/5d\_7f43ba4585155f1e4f6a29/Virtual-Learning-Environment-VLE-Implementation-Strategy\_An\_Analysis-of-Practicality-for-Google-Classroom-Implementation-in-Malaysian-Schools.pdf
  </u>
- Basher, S. A. O. (2017). The impact of Google Classroom application on the teaching efficiency of pre-teachers. *International Journal Social Sciences and Education*, 7(2), 45-54. http://ijsse.com/sites/default/files/issues/2017/v7i2/Paper-4.pdf
- Cheok, M. L., Wong, S. L., Ahmad, F. A., & Rosnaini, M. (2017). Teachers' perceptions of e-learning in Malaysian secondary schools. *Malaysian Online Journal of Educational Technology*, 5(2), 20-33. <u>https://files.eric.ed.gov/fulltext/EJ1142390.pdf</u>
- Chung, M. H. C. (2022). Google Classroom sebagai salah satu platform dalam penyampaian pengajaran dan pembelajaran: Satu kajian tinjauan. *Jurnal Kurikulum & Pengajaran Asia Pasifik*, *10*(1), 1–9. <u>https://ejournal.um.edu.my/index.php/JUKU/article/view/35298</u>
- Dash, S. (2019). Google Classroom as a learning management system to teach biochemistry in a medical school. *Biochemistry and molecular biology education, 47*(4), 404-407. https://doi.org/10.1002/bmb.21246
- Dewi, C. A., Muhali, M., Kurniasih, Y., Lukitasari, D., & Sakban, A. (2022). The impact of Google Classroom to increase students' information literacy. *International Journal of Evaluation and Research in Education*, *11*(2), 1005–1014. <u>https://doi.org/10.11591/ijere.v11i2.22237</u>
- Fishbein, M., & Ajzen, I. (1980). *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, NJ: Prentice Hall.
- Hinton, P., McMurray, I., & Brownlow, C. (2014). SPSS Explained. 711 Third Avenue, NY: Routledge.
- Hoque, M., Mohamed, Y., Salaeh, A., & Kadir, K. A. (2020). Students' attitudes towards educational technology. *International Journal of Advanced Research in Engineering and Technology*, 11(10), 267-274. <u>https://oarep.usim.edu.my/jspui/handle/123456789/6010</u>
- Kumar, J. A., & Bervell, B. (2019). Google Classroom for mobile learning in higher education: Modelling the initial perceptions of students. *Education and Information Technologies*, 24(2), 1793-1817. <u>https://doi.org/10.1007/s10639-018-09858-z</u>
- Lepp, L., Aaviku, T., Leijen, Ä., Pedaste, M., & Saks, K. (2021). Teaching during COVID-19: The decisions made in teaching. *Education Sciences*, *11*(2), 1–21. https://doi.org/10.3390/educsci11020047
- Mafa, K. R. (2018). Capabilities of Google Classroom as a teaching and learning tool in higher education. *International Journal of Science Technology & Engineering, 5*(5), 30-34. <u>https://www.ijste.org/articles/IJSTEV5I5006.pdf</u>

- Malaysian Ministry of Education. (2013). *Malaysia education blueprint 2013-2025 (Preschool to post-secondary education)*. <u>https://www.moe.gov.my/menumedia/media-cetak/penerbitan/dasar/1207-malaysia-education-blueprint-2013-2025/file</u>
- Malaysian Ministry of Education. (2019, July 1). *Kementerian Pendidikan Malaysia: Google Classroom.* Ministry of Education.

https://www.moe.gov.my/en/pemberitahuan/announcement/google-classroom-gc

- Maxwell, S. E. (2000). Sample size and multiple regression analysis. *Psychological methods, 5*(4), 434.
- Mokhtar, H. S. (2020, March 31). Teachers explore online teaching methods during MCO. *New Straits Times*. <u>https://www.nst.com.my/education/2020/03/579992/teachers-explore-online-teaching-methods-during-mco</u>
- Moses, P., Wong, S. L., Abu Bakar, K., & Mahmud, R. (2010). A preliminary study: Level of laptop competence among secondary school teachers. In Z. Abas, I. Jung & J. Luca (Eds.), *Proceedings of Global Learn Asia Pacific 2010* (pp. 195-204). Association for the Advancement of Computing in Education (AACE). <u>https://www.learntechlib.org/primary/p/34173/</u>
- Nation. (2019, June 27). Contract between Education Ministry and 1BestariNet ends June 30. *The Star*. <u>https://www.thestar.com.my/news/nation/2019/06/27/1bestarinet-contract-ends-june-30</u>
- Nidup, Y. (2022). Teachers ICT Skills and application of ICT in the middle and higher secondary schools in teachers' ICT skills and application of ICT in the middle and higher secondary schools in Bhutan. *Journal of ICT in Education*, *9*(1), 77–85. https://ojs.upsi.edu.my/index.php/JICTIE/article/view/6133/3432

Nikian, S., Nor, F. M., & Aziz, M. A. (2013). Malaysian teachers' perception of applying technology in the classroom. *Procedia Social and Behavioral Sciences*, *103*, 621-627. https://doi.org/10.1016/j.sbspro.2013.10.380

- Olumorin, C. O., Babalola, E. O., Ashaolu, S., & Omolafe, E. V. (2022). Students' attitude towards the utilization of Google Classroom for learning. *Indonesian Journal of Educational Research and Technology*, 2(3), 213–222. <u>https://ejournal.upi.edu/index.php/IJERT/article/view/44828/18607</u>
- Piaralal, S. D., Tahir, M. H. M., Adnan, A. H. M., Shah, D. S. M., & Shak, M. S. Y. (2023). The use of Google Classroom among secondary school teachers. *Journal of Nusantara Studies*, 8(1), 310– 332. <u>https://journal.unisza.edu.my/jonus/index.php/jonus/article/view/741/409</u>
- Rauf, A. A., & Swanto, S. (2020). Attitudes and technology integration among ESL secondary school teachers in Sabah. *Malaysian Journal of Social Sciences and Humanities*, *5*(12), 280-287. <u>https://msocialsciences.com/index.php/mjssh/article/view/566/432</u>
- Rahayu, T., Syafril, S., Pahrudin, A., Aini, N. R., & Puspasari, V. (2019). Use of Frog VIe in science learning. *Journal of Physics: Conference Series*, 1155 (1), 1-5. <u>https://iopscience.iop.org/article/10.1088/1742-6596/1155/1/012089/pdf</u>
- Rosni, N. M., Isa, K., & Alias, A. A. (2022). Penerimaan pelajar pelajar di Semenanjung Malaysia terhadap penggunaan aplikasi Google Classroom sepanjang tempoh perintah kawalan pergerakan. *Journal of Borneo Social Transformation Studies*, *8*(1), 113–124. https://doi.org/https://doi.org/10.51200/jobsts.v8i1.4167
- Shaharanee, I. N. M., Jamil, J., & Mohamad Rodzi, S. S. (2016). The application of Google Classroom as a tool for teaching and learning. *Journal of Telecommunication, Electronic and Computer Engineering*, *8*(10), 5-8.

https://repo.uum.edu.my/id/eprint/20521/1/JTEC%208%2010%202016%205%208.pdf

- Statistics Solutions. (2021, August 2). Sample Size Formula Statistics Solutions. <u>https://www.statisticssolutions.com/dissertation-resources/sample-size-calculation-and-sample-size-justification/sample-size-formula/</u>
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in science education, 48*(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478. https://doi.org/10.2307/30036540
- Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157-178. <u>https://doi.org/10.2307/41410412</u>
- Wagner, W. E. (2015). Using IBM SPSS statistics for research methods and social science statistics. USA: SAGE Publications.
- Zakaria, M., Ahmad, J. H., Bahari, R., Hasan, S. J., & Zolkaflil, S. (2021). Benefits and challenges of adopting Google Classroom in Malaysian university: Educators' perspectives. *Ilkogretim Online: Elementary Education Online*, *20*(1), 1296–1304. <u>https://doi.org/10.17051/ilkonline.2021.01.123</u>

# An Exploratory Study on Effects of WeChat-Assisted Self-Regulated English Writing Instruction on EFL Learners' Motivation

#### Ying ZHAO<sup>a\*</sup>, Di ZOU<sup>b</sup>, Haoran XIE<sup>c</sup>, Pei LIU<sup>d</sup>

<sup>a</sup>Department of English Language Education, The Education University of Hong Kong, Hong Kong SAR, China

<sup>b</sup>School of Graduate Studies, Lingnan University, Hong Kong SAR, China <sup>c</sup>Department of Computing and Decision Sciences, Lingnan University, Hong Kong SAR, China

<sup>d</sup>School of science and technology, Hong Kong Metropolitan University, Hong Kong SAR, China

#### \*s1142504@s.eduhk.hk

Abstract: Scholars suggested using social networking to assist second language and foreign language learning. Researchers in the present study designed and developed a Mini Program on the WeChat platform to assist university EFL learners in improving self-regulated English writing. This exploratory study aims to investigate the effects of WeChat-assisted Self-Regulated English Writing Instruction on university EFL learners' motivation in English writing. A quasi-experimental design with forty-five second-school-year university students was conducted for two weeks. Participants were randomly divided into the control group, the traditional group and the WeChat-assisted group. They took a pre- and post-questionnaire survey on English writing motivation before and after the intervention. Descriptive statistic analysis was used to show the status quo of the participants' motivation, paired sample t-test was used to analyze the differences in motivation within each group before and after the intervention, and ANOVA analysis was used to compare differences in motivation between groups after the intervention. Both the traditional approach and the WeChat approach generated positive impacts on students' overall motivation, self-efficacy and interest. Students using the traditional approach outperformed those using the WeChat-assisted approach but the results did not reach a significant level. Implications for future studies are suggested.

**Keywords:** self-regulated language learning; motivation; foreign language learning; technology-enhanced language learning

#### 1. Introduction

The significance of Self-Regulated Learning (SRL) is increasing with the technological tools advance. In the domain of English as a second language and foreign language learning

(ESL and EFL), previous researchers have conducted much research on investigating the application of different technological tools in assisting SRL, such as digital games (e.g., Qiao et al., 2022) and mobile apps (e.g., Yang & Song, 2022). According to scholars' review papers in technology-enhanced language learning, they emphasized the using of social networks as a promising educational technology research trend in ESL and EFL (Chen et al., 2020; Zhang & Zou, 2022). However, WeChat, as a social networking application with over one billion monthly active users worldwide (Tencent, 2023), has not received much attention from researchers. In addition, English writing is one of the most complex and challenging skills for EFL learners (Graham et al., 2018). Therefore, the present study designed and developed a WeChat Mini Program to assist university EFL learners in improving self-regulated English writing.

# 2. Literature Review

# 2.1 Self-Regulated Learning in English Writing

The Self-Regulated Learning (SRL) has become a significant characteristic in the 21<sup>st</sup> century (Li et al., 2022). As for self-regulated learning in writing, Zimmerman and Risemberg (1997) laid the theoretical foundation by analyzing the writing process from an SRL perspective. Harris and Graham (2017) suggested that teachers should conduct the explicit instruction on self-regulated strategies for students. Graham et al. (2018) concluded the importance of self-regulated language learning (SRLL) strategies on writing and writing development across different writing contexts. Many researchers have confirmed the SRLL strategy's positive impacts on second or foreign language writing with empirical evidence (e.g., Teng et al.,2022). However, these researchers targeted young learners to investigate the impacts of SRLL strategies on English writing. Both Bai et al. (2020) and Teng (2020) highlighted the advantages of SRLL for young learners' English writing. Though many existing studies on the SRLL strategy use were on university students, they were from over a decade ago (Chang, 2005; Chien, 2012).

Therefore, more studies are desperately needed to use an intervention to investigate university students' SRLL in EFL writing contexts. Researchers and teachers needs to learn about the status quo of university EFL learners' SRLL. More importantly, continuous pedagogical interventions for lifelong learning can become possible.

#### 2.2 Motivation in English Writing

Motivation is generally conceptualized as the psychological incentives to generate actions (Barak et al., 2016). Dörnyei (2005) stressed the significance of motivation for students to achieve success in EFL. Studies on motivation in ESL/EFL drew much attention from researchers but studies specifically focused on motivation in English writing were relatively scant (Lee et al., 2018). Later, Guo and Bai (2019) suggested that self-efficacy and interest are crucial elements to EFL learners in self-regulated English writing, especially for young learners. However, not enough attention is paid to university EFL learners who transmitted from traditional middle school to a new and different EFL learning context at the university. Therefore, this exploratory study planned to investigate the effects of the

WeChat-assisted Self-Regulated English Writing Instruction (SREWI) on these variables.

Regarding self-efficacy in motivation, extensive research has shown that students with high motivation performed better in English writing (Bai et al., 2021; Bai et al., 2022). As for SRL, researchers also found the positive impacts of students' self-efficacy on self-regulated English strategy use. Regarding interest in motivation, Dörnyei and Ushioda (2013) conceptualized interest in the learning context as the level of a learner's engagement and participation driven by curiosity and challenges. Bai and Guo (2019) found a low level of interest in English writing among young learners. They called for more attention to research on interest in English writing. Bai et al. (2022) investigated effects of self-regulated English writing instruction by an e-learning approach on young learners' overall motivation, self-efficacy, and interest. Their study found a significant increase in interest after the e-learning intervention for young learners. Therefore, this exploratory study will involve university students to explore their motivation in English writing. The present study implemented the SREWI through a traditional approach and a WeChat-assisted learning approach. Two experimental groups of students received these two different intervention respectively. The primary research questions for the exploratory study included:

1. What is the status quo of university EFL learners' English writing motivation before the intervention?

2. Is there any difference in English writing motivation for students before and after the intervention in each condition?

3. Is there any difference in English writing motivation for students after the intervention between different conditions?

# **3.Research Methods**

#### 3.1 Participants

Convenience sampling was adopted for this exploratory study. Forty-five second-school-year university EFL learners aged from 19 to 24, majoring in civil aviation at a comprehensive university in mainland China, took part in the study voluntarily. The prerequisite for the students to start the professional training at the aviation academy is to meet the English language proficiency criteria. Therefore, it is of great necessity to improve their self-regulated English learning. This exploratory study investigated the effects of the SREWI on EFL learners' motivation in English writing.

#### 3.2 Design of the WeChat-Assisted Self-Regulated Writing Instruction Mini Program

Researchers designed a WeChat Mini Program on the SREWI based on the "Cyclical Phase Model" (Cleary & Zimmerman, 2012) and the "Self-Regulated Strategy Development Theory" (Harris & Graham, 2017). The major purpose of the program was to help students learn SRLL strategies for English argumentative writing. There were five lessons to teach students SRLL strategies for pre-writing, while-writing, and post-writing stages. As shown in Figure 1, this WeChat Mini program supported five main functions for learners, namely previewing the lesson by watching a short video clip, investigating learners' previous experience on SRLL by doing a survey, learning SRLL strategies by reading materials and

completing tasks, self assessing the study by taking a quiz, and monitoring self-study by visiting personal profile and leaderboard.



Figure 1. Functions of WeChat Mini Program

# 3.3 Procedures

The study adopted a quasi-experimental design. All the participants took a pre- and post-questionnaire survey on motivation before and after the experiment. Before the experiment, the researcher sent consent forms to fifty-five students from the same class, their College English teacher and the dean of the department to introduce the study's purposes and procedures, and explained that any participant could withdraw the experiment at any time of the study. Forty-five students signed the consent form and took part in the experiment. All forty-five students were divided into three groups randomly: a control group, a traditional group and a WeChat-Assisted group. All the control group, the traditional group and the WeChat-Assisted group had the regular College English course and took part in the pre- and post-survey. Both of the traditional group and the WeChat-Assisted group took the treatment of the SREWI after class for two weeks. However, the traditional group took the treatment by reading paper materials and the WeChat-Assisted group by using the WeChat SREWI Mini Program designed by the researchers.

The experiment began on June 2, 2023 and lasted for two weeks. The participants in the traditional group and those in the WeChat-assisted group could arrange their own study in an autonomous learning environment. As an exploratory study, the researchers did not set the requirement on learning time but requested the teacher to remind participants in the experimental groups to learn every two days. The post-survey was conducted on June 17, 2023. The teacher collected the questionnaires immediately after students completed them. Because one participant in the WeChat-assisted experimental group did not complete the post-survey, data from the other fourteen participants in this group, fifteen in the control group and fifteen in the traditional experimental group, was analyzed.

#### 3.4 The instrument for measuring students' motivation

This study used the motivation questionnaire adopted by Bai et al.'s (2022) research. Their study reported the high internal consistency of the questionnaire items. The questionnaire is a 5-point Likert scale and contains nine items in two subcategories: self-efficacy and interest. Participants rate each item from "1" ("completely disagree") to "5" ("completely agree"). In order to assess the reliability of the questionnaire, the data from the pre-survey was analyzed and the Cronbach' s alpha was 0.90. The result indicated that the reliability of the questionnaire was high and could be used for further analysis.

#### 4. Results

Descriptive statistic analysis was adopted to describe EFL learners' motivation status quo. Paired sample *t*-test was used to analyze the differences in EFL learners' motivation within the group before and after the intervention. ANOVA was used to compare EFL learners' motivation differences between different groups after the intervention.

#### 4.1 The Status quo of participants' English writing motivation before the intervention

Oxford (1990) classified self-reported data for a 5-Likert scale questionnaire into three levels: a high level for a mean in the range of 3.5-5.0, a medium level for 2.5-3.4, and a low level for 1.0-2.4. Table 1 shows that participants' overall motivation score was at a medium level, slightly over 3.00 (M=3.28, SD=.67). For the two subcategories, the writing self-efficacy score (M=3.53, SD=.70) was higher than the writing interest score (M=2.90, SD=.81).

Table 1. Results of participants motivation in English writing in the pre-survey						
		Min	Max	Mean	SD	Ν
	Motivation	1.78	5.00	3.28	.67	44
Pre \	Writing self-efficacy	1.60	5.00	3.53	.70	44
survey	Writing interest	1.00	5.00	2.90	.81	44

Table 1. Results of participants' motivation in English writing in the pre-survey

#### 4.2 English writing motivation before and after the experiment within each group

Table 2 illustrates the pre- and post-survey scores of participants' motivation in the control group. Participants' self-efficacy was at a lower high level (M=3.52, SD=.75) before the experiment and increased slightly (M=3.68, SD=.73) after the experiment, even though they did not receive any additional intervention. The increase in writing self-efficacy was not significant (t=-1.38, df=14, P=.19 > .05). Students' writing interest was at a medium level (M=2.87, SD=.88) before the experiment and increased (M=3.27, SD=.84) after the experiment. The writing interest score did not increase significantly (t=-1.99, df=14, P=.07 > .05). Participants' overall motivation was at a medium level (M=3.23, SD=.73) before the experiment and increase significantly (t=-1.99, df=14, P=.07 > .05). Participants' overall motivation was at a medium level (M=3.23, SD=.73) before the experiment and increased to a high level (M=3.50, SD=.71) after the experiment. However, the overall motivation increase was not significant (t=-2.08, df=14, P=.06 > .05).

Table 3 illustrates the pre- and post-survey scores of participants' motivation in the traditional experimental group. Participants' writing self-efficacy was at a high level before the intervention and increased after the intervention. The self-efficacy score increase was significant (t=-2.45, df=14, p<.05). Participants' writing interest was at a medium level

(M=2.98, SD=.94) before the intervention and increased (M=3.47, SD=1.05) after the intervention. The writing interest score improved significantly (t=-2.53, df=14, P<.05). As for the overall motivation, the change reached a significant level (t=-2.84, df=14, P<.05), increasing from a medium level (M=3.33, SD=.72) before the intervention to a high level (M=3.77, SD=.67) after the intervention.

Table 4 shows the pre- and post-survey scores of participants' motivation in the WeChat-assisted experimental group. Before the intervention, participants' self-efficacy in English writing was at a medium level (M=3.46, SD=.71) and increased to a lower high level (M=3.56, SD=.58) after the WeChat-assisted intervention. But the change was not significant (t=-.76, df=13, P=.46 > .05). Their writing interest was at a medium level (M=2.86, SD=.63) before the intervention and increased (M=3.36, SD=.74) after the WeChat-assisted intervention. The interest in English writing increased significantly(t=-2.76, df=13, P<.05). As for the overall motivation in English writing, the change was not significant (t=-.65, df=13, P=.53 > .05).

		Mean	SD	SE Mean	Paired <i>t</i> test		
					t value	df	Sig (two-tailed)
writing self-efficacy	pre	3.52	.75	.19	-1.38	14	.19
	post	3.68	.73	.19	_		
writing interest	pre	2.87	.88	.23	-1.99	14	.07
	post	3.27	.84	.22	—		
Motivation	pre	3.23	.73	.22	-2.08	14	.06
	post	3.50	.71	.19			

Table 2. EFL learners' motivation before and after the experiment in the control group

Table 3. Motivation before and after the intervention in the traditional experimental group

		Mean	SD	SE Mean	Paired <i>t</i> test		
					t value	df	Sig (two-tailed)
writing self-efficacy	pre	3.60	.69	.18	-2.45	14	.03
	post	4.01	.58	.15	-		
writing interest	pre	2.98	.94	.24	-2.53	14	.02
	post	3.47	1.05	.27	-		
Motivation	pre	3.33	.72	.19	-2.84	14	.01
	post	3.77	.67	.17	-		

 Table 4. Motivation before and after the intervention in the WeChat-Assisted group

		Mean	SD	SE Mean	Paired <i>t</i> test		
					t value	df	Sig (two-tailed)
writing self-efficacy	pre	3.46	.71	.19	76	13	.46
	post	3.56	.58	.16			
writing interest	pre	2.86	.63	.17	-2.76	13	.02
	post	3.36	.74	.20			
Motivation	pre	3.29	.59	.16	65	13	.53
	post	3.37	.47	.12	_		

#### 4.3 English Writing Motivation before and after the intervention between groups

One-way ANOVA was adopted to assess the homogeneity of variance before starting data analysis between groups. Data from the pre-survey was used to test the homogeneity of variance. The results in Table 5 show that there is no significant difference (F(2, 41)=.075, P=.928>.05) between groups in participants' motivation before the intervention. Therefore, the results demonstrated that the motivation data from the control and experimental groups before the experiment was homogeneous. As there were multiple variables in the study, MANOVA was used to generate adjusted means to compare differences in motivation in post-survey between groups.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.071	2	.035	.075	.928
Within Groups	19.344	41	.472		
Total	19.415	43			-

Table 5. Results of ANOVA test on pre-survey data

4.3.1 Differences in EFL learners' motivation in post-survey between the control group and the traditional experimental group

Adjusted means were used to compare the differences in EFL learners' motivation after the experiment in order to reduce the misinterpretation of the data with multiple variables. As it is shown in Table 6, the adjusted mean of the control group is 3.52 and 3.75 of the traditional experimental group. However, there was no significant difference (F(1,28)=1.57, P=.22 > .05) between the two groups. For the two subcategories of motivation, the control group's adjusted mean of self-efficacy is 3.69 and 4.00 for the traditional experimental group. But the difference in self-efficacy was not significant (F(1, 28)=3.38, P=.08 > .05). The adjusted mean of interest for the control group was 3.30 and 3.43 for the control group. However, the result shows no significant difference in interest between the two groups. Even though the traditional experimental group's adjusted means of overall motivation, self-efficacy, and interest in the post-survey were higher than those in the control group, none of the results reached a significant level.

Variable	Group	Ν	Mean	S.D.	Adjusted mean	F	Sig.
Motivation	Control Group	15	3.50	.71	3.52	1.5	.22
	Traditional Group	15	3.77	.67	3.75	7	
Self-efficacy	Control Group	15	3.68	.73	3.69	3.3	.08
	Traditional Group	15	4.01	.58	4.00	8	
Interest	Control Group	15	3.27	.84	3.30	.21	.65
	Traditional Group	15	3.47	1.05	3.43	-	

Table 6. Differences in motivation in the post-survey between the control group and the traditional experimental group

4.3.2 Differences in EFL learners' motivation in the post-survey between the traditional experimental group and the WeChat-Assisted experimental group

As it is shown in Table 7, the adjusted mean of motivation for the traditional experimental group is 3.71 and 3.44 for the WeChat-assisted group. However, there is no significant difference in EFL learners' motivation (F(1, 27)=2.12, P=.16 > .05) in English writing between the traditional experimental group and the WeChat experimental group. For the two subcategories of motivation, the traditional group's adjusted mean of self-efficacy is 3.97 and 3.60 for the WeChat-assisted group. But the difference is not significant (F(1, 27)=3.70, P=.07 > .05). The traditional experimental group's adjusted mean of interest is 3.49 and 3.34 for the WeChat-assisted group. However, the result did not show any significant difference between the two experimental groups (F(1,27)=.28, P=.60 > .05). According to the analysis, the adjusted means of overall motivation, self-efficacy, and interest of the traditional group were higher than those in the WeChat-assisted experimental group in the post-survey. However, the results were not significant.

					1	5 1	
Variable	Group	Ν	Mean	S.D.	Adjusted mean	F	Sig.
Motivation	Traditional Group	15	3.77	.67	3.71	2.12	.16
	WeChat Group	14	3.37	.47	3.44		-
Self-efficacy	Traditional Group	15	4.01	.58	3.97	3.70	.07
	WeChat Group	14	3.56	.58	3.60		-
Interest	Traditional Group	15	3.47	1.05	3.49	.28	.60
	WeChat Group	14	3.36	.74	3.34		-

Table 7. Differences in motivation in the post-survey of the two experimental groups

# 5. Discussion and Conclusion

Regarding the status quo of participants' motivation in English writing before the intervention, the overall motivation was at a medium level. It should be noted that the small sample size and the participants' major may influence the results in the status quo of motivation in our exploratory study. For our exploratory study, the students' self-efficacy in English writing was at a lower high level. This finding was consistent with Bai et al.'s study (2022) with young EFL learners. They found that young EFL learners' self-efficacy was already at a lower high level before the intervention. It showed that both young and university-level students reported a lower high level of self-efficacy in English writing. Besides, the writing interest for university EFL learners was at a medium level before the intervention. However, this finding differs from Bai and Guo's study (2019). They found that EFL learners' interest was at a low level. There might be two probable reasons for this difference. The major reason might be the occupational feature of the participants in our study, because the expectation of becoming a civil aviation pilot may stimulate a student's interest in English. Additionally, age might be an influential factor, because the participants in our study were university level students but those in Bai and Guo's study (2019) were primary students. Considering the results in self-efficacy and interest, it indicated that age may generate a more substantial influence on EFL learners' interest rather than self-efficacy.

Regarding the motivation in English writing before and after the experiment within each group, both an traditional approach and a WeChat-assised approach generated positive impacts on participants' overall motivation, self-efficacy, and interest. However, the increase generated by the traditional learning approach reached a significant level, which implied the effectiveness of the SREWI on facilitating students' motivation. The WeChat-assited approach improved students' interest in English writing significantly. This finding is consistent with previous research. For example, EFL learners' interest increased significantly after the intervention of a technology-enhanced SRL English writing instruction (Bai et al., 2022).

Regarding the motivation in English writing in the post-survey between groups, the SREWI could positively impact participants' motivation but the improvement was not statistically significant. Students who learned the SREWI in a traditional approach outperformed those who did not take the intervention. This finding implied the positive effects of the SREWI on students' English writing motivation. However, students who learned the SREWI in a WeChat-assisted approach did not perform better than those who took a traditional learning approach. As the final examinations were approaching when the study was to start, the researchers did not conduct an individual orientation session to explain the use of this WeChat Program in detail. Additionally, the small sample size and the short experiment period might attribute to the results.

Therefore, the present exploratory study found the positive effects of WeChat-assisted SREWI on EFL learners' motivation, self-efficacy, and interest in English writing. Students who used a traditional learning approach to learn the instruction performed better than others in motivation. The WeChat-assisted approach increased students' interest significantly. The findings implied that language teachers could use the applications in the market to design appropriate learning tasks and instructions to assist in teaching and facilitate students' motivation in learning (Bai et al., 2022).

Limitations of the present study and suggestions for future research will be suggested. First, the present exploratory study was lack of immediate and effective teacher support during the intervention. The implementation of a new technology learning tool requires effective teacher support (Zhang & Zou, 2022). Second, the small sample size and the short experiment period were limitations of this study. Due to these two factors, the findings of this exploratory study cannot be generalized to students in other learning contexts. So, future studies are suggested to experiment with a large sample size during a long learning period. Third, the design of the WeChat Mini Program impacted the results in the WeChat-assisted group. In order to improve the design and quality of this Mini Program, researchers collected data on students' using experience after the experiment and made improvements for the program. In future studies, researchers will further improve the program, investigate the effects of WeChat-assisted Self-regulated English Writing on other learning outcomes, and explore effective approaches for EFL learners to improve their English writing with the help of advanced technologies.

#### References

- Bai, B., & Guo, W. (2019). Motivation and self-regulated strategy use: Relationships to primary school students' English writing in Hong Kong. *Language Teaching Research, 25*(3), 378-399.
- Bai, B., Shen, B., & Mei, H. (2020). Hong Kong primary students' self-regulated writing strategy use: Influences of gender, writing proficiency, and grade level. *Studies in Educational Evaluation*, 65, 100839. https://doi.org/10.1016/j.stueduc.2020.100839
- Bai, B., Wang, J., & Nie, Y. (2021). Self-efficacy, task values and growth mindset: What has the most

predictive power for primary school students' self-regulated learning in English writing and writing competence in an Asian Confucian cultural context?. *Cambridge Journal of Education, 51*(1), 65-84.

- Bai, B., Wang, J., & Zhou, H. (2022). An intervention study to improve primary school students' self-regulated strategy use in English writing through e-learning in Hong Kong. *Computer Assisted Language Learning*, 35(9), 2265-2290.
- Barak, M., Watted, A., & Haick, H. (2016). Motivation to learn in massive open online courses: Examining aspects of language and social engagement. *Computers & Education, 94*, 49-60.
- Chang, M. M. (2005). Applying self-regulated learning strategies in a web-based instruction-An investigation of motivation perception. *Computer Assisted Language Learning*, *18*(3), 217–230. https://doi.org/10.1080/09588220500178939
- Chen, X., Zou, D., Cheng, G., & Xie, H. (2020). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of Computers & Education. *Computers & Education*, *151*, 103855.
- Chien, S. C. (2012). Students' use of writing strategies and their English writing achievements in Taiwan. *Asia Pacific Journal of Education*, *32*(1), 93–112. https://doi.org/10.1080/02188791.2012.655240
- Cleary, T. J., & Zimmerman, B. J. (2012). A cyclical self-regulatory account of student engagement: Theoretical foundations and applications. In S. L. Christenson, A. L. Reschly, & C. Wylie (eds.), *Handbook of research on student engagement* (pp. 237-257). Springer.
- Dörnyei, Z. (2005). The psychology of the language learner. Lawrence Erlbaum.
- Dörnyei, Z., & Ushioda, E. (2013). Teaching and researching: Motivation. Routledge.
- Graham, S., Harris, K.R., MachArthur, C., & Santangelo, T. (2018). Self-regulation and writing, in Handbook of Self-Regulation of Learning and Performance (pp.138-152), D. H. Schunk & J. A. Greene (eds.). Routledge.
- Guo, W. J., & Bai, B. (2019). Effects of self-regulated learning strategy use on motivation in EFL writing: A comparison between high and low achievers in Hong Kong primary schools. *Applied Linguistics Review*, *13*(1), 117–139. doi:10.1515/applirev-2018-0085
- Harris, K. R., & Graham, S. (2017). Self-regulated strategy development: Theoretical bases, critical instructional elements, and future research. In *Design principles for teaching effective writing* (pp. 119-151). Brill.
- Lee, I., Yu, S., & Liu, Y. (2018). Hong Kong secondary students' motivation in EFL writing: A survey study. *TESOL Quarterly, 52*(1), 176–187
- Li, X., Xia, Q., Chu, S. K. W., & Yang, Y. (2022). Using gamification to facilitate students' self-regulation in e-learning: A case study on students' L2 English learning. *Sustainability, 14*(12), 7008. https://doi.org/10.3390/su14127008
- Qiao, S., Chu, S. K. W., Shen, X., & Yeung, S. S. S. (2022). The impact of an online gamified approach embedded with self regulated learning support on students' reading performance and intrinsic motivation: A randomized controlled trial. *Journal of Computer Assisted Learning, 38*(5), 1379-1393. https://doi.org/10.1111/jcal.12684
- Shen, B., & Bai, B. (2022). Chinese university students' self-regulated writing strategy use and EFL writing performance: Influences of self-efficacy, gender, and major. *Applied Linguistics Review*, (0). https://doi.org/10.1515/applirev-2020-0103
- Tencent. (2023). Number of active WeChat messenger accounts Q2 2011-Q4 2022. Retrieved from: https://www.statista.com/statistics/255778/number-of-active-wechat-messenger-accounts/.

- Teng, F. (2020). Young learners' reading and writing performance: Exploring collaborative modeling of text structure as an additional component of self-regulated strategy development. *Studies in Educational Evaluation, 65*.
- Teng, M. F., Wang, C., & Zhang, L. J. (2022). Assessing self-regulatory writing strategies and their predictive effects on young EFL learners' writing performance. *Assessing Writing*, *51*, 100573. https://doi.org/10.1016/j.asw.2021.100573
- Yang, Y., & Song, Y. (2022). Understanding primary students' self-regulated vocabulary learning behaviours on a mobile app via learning analytics and their associated outcomes: a case study. *Journal of Computers in Education*, 1-30. https://doi.org/10.1007/s40692-022-00251-x
- Zhang, R., & Zou, D. (2022). Self-regulated second language learning: a review of types and benefits of strategies, modes of teacher support, and pedagogical implications. *Computer Assisted Language Learning*, 1-38. https://doi.org/10.1080/09588221.2022.2055081
- Zimmerman, B., & Risemberg, R. (1997). Becoming a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology, 22*, 73–101. https://psycnet.apa.org/doi/10.1006/ceps.1997.0919

# A Step toward Characterizing Student Collaboration in Online Knowledge Building Environments with Machine Learning

Alwyn Vwen Yen LEE<sup>a\*</sup>, Chew Lee TEO<sup>b</sup> & Aloysius ONG<sup>c</sup>

<sup>abc</sup>National Institute of Education, Nanyang Technological University, Singapore \*alwyn.lee@nie.edu.sg

**Abstract:** Existing research has substantial progress in uncovering outcomes of collaborative learning in recent years, but more attention can be directed towards the better understanding of collaborative learning processes via quantitative frameworks and methods. Through the use of knowledge building as a collaborative learning pedagogical approach, it is possible for researchers to glean deeper insights into aspects of students' collaboration within authentic learning environments. In this paper, the multimodal approach of data collection and analysis was conducted with a proposed conceptual analytical framework that can characterize constructs of collaborative activities in a knowledge building classroom using machine learning methods. The application in a pilot is discussed along with how this conceptual development can offer a summary of new insights into students' individual and group collaborative trajectories during learning tasks.

**Keywords:** Computer-supported collaborative learning, student collaboration, knowledge building, multimodal learning analytics, machine learning

#### 1. Introduction

Since the conceptualization of computer-supported collaborative learning (CSCL) in the late 1980s, the field and related studies of collaborative activity among a group of individuals for the purpose of learning and supported by information and communication technologies (Suthers, 2012) have subsequently made significant advancements and inroads into the development of technology-enhanced tools and theories to support the collaborative learning process (Hernández-Sellés et al., 2020). A plethora of frameworks, methods and techniques exists to measure student collaboration in terms of qualitative and quantitative indicators that are found within authentic learning environments (e.g., Martinez et al., 2003; Mercer et al, 2009). Within different forms of collaborative learning and pedagogical approaches, Knowledge Building (Scardamalia & Bereiter, 2003) has emerged as a principle-based approach where collaboration is a means for community knowledge advancement and acquisition, with feasible frameworks and techniques to evaluate the quality of both online and offline face-to-face discourse (e.g., Lee & Tan, 2017).

The measure of student collaboration is however multi-faceted and research studies tend to focus on specific notions of collaborations (e.g., engagement; Halpin et al., 2017) to avoid over-stretching resources. With the advent of virtual learning environments (VLEs) and the affordance of metrics accelerated by digital technology and new norms, recent CSCL research and analyses (Jeong et al, 2019) have been able to produce outcomes of collaborative learning from CSCL research that are more quantifiable and tangible, but this is to the detriment of more emphasis that should be placed on the analysis and understanding of collaborative learning processes. Further, heavy reliance on educators' expertise and experience is still expected in practice (Ben-Peretz, 2001), to observe, analyze, and identify the different extents and levels of student collaboration, and therefore the use of virtual

learning environments contributes to an additional digital layer of surveillance and monitoring that educators need to handle as compared to in-person learning. Knowledge building related studies were also of no exception and the critical importance of better understanding the collaboration processes with educational outcomes as by-products has constantly been emphasized but are often times the other way around in most completed studies.

To address these trends, emergent technologies can be tapped to investigate the multidimensions of collaborative activities that influence the quality of collaborative process of learning, but with lesser burdens on the educator and without sacrificing the visibility and degree of understanding about collaborative activities. An emergent and explored way of better understanding the different dimensions of interaction and knowledge growth in a collaborative learning environment is to adopt a multimodal learning analytics (MMLA) approach to investigate a myriad of multimodal data including physiological data (heart rate, skin conductance); physical movement (poses), verbal utterances, tonalities and overlaps of voices (Di Mitri, 2018). This approach is wide-ranging in nature and its use is briefly explained in this paper as the underlying data collection approach.

As part of an eventual goal to obtain a comprehensive prediction of the level of collaboration in an authentic CSCL environment, the design and development of a conceptual multimodal analytical framework is proposed, by conducting the following:

- 1. First, to categorize different constructs of collaboration into various analytical levels and structures to replicate the possible interactions patterns that exist in a collaborative learning environment (i.e., student-student group; student group—teacher).
- 2. Next, multimodal data in terms of metrics and what can be measured, will be parsed into the proposed micro, meso, or macro levels for further analysis, and also considering the complexity and scale of the data involved.
- 3. Finally, the framework will aggregate the measurements of the various modalities and attempt to integrate them into a single indicator to represent the level of collaboration for the measured entity, that is, the group when undergoing group work, and the class during whole-class instruction.

In this paper, apart from presenting and discussing the design and development of a conceptual analytical framework that is currently being tested in several knowledge building classes, we also report some preliminary findings from a completed pilot case study that used machine learning algorithms and the proposed analytical framework to evaluate and verify the quality of collaboration and collaborative discourse within a knowledge building and CSCL environment. We also seek to answer the following research question: *How are aspects of collaborative learning processes in a knowledge building environment, whether offline (face-to-face) or online, characterized and visualized with the aid of an analytical framework to determine the extent of collaboration in a CSCL environment?* 

# 2. Background and Literature Review

# 2.1 Knowledge Building and relevant VLE (Knowledge Forum)

Knowledge building is an approach that moves away from traditional instruction-based learning towards the empowerment of students as members of a knowledge building community, allowing them to take charge of their own and the community's knowledge creation processes and the advancement of knowledge (Scardamalia & Bereiter, 2003). Studies on knowledge building have found significant, positive impacts on the development of students' metacognition, collaboration, and epistemic inquiry processes for knowledge acquisition (Yang et al., 2020) and is often supported by the Knowledge Forum, an online discourse platform that facilitates inquiry, discourse, knowledge advancement, and collective cognitive responsibility (Chan & Chan, 2011). On the Knowledge Forum, students engage in online discourse through writing and reading of notes, as well as building-on to notes of other students in class-sized groups or smaller working groups during group work (see Figure 1).



*Figure 1.* Spatial area on the Knowledge Forum for students to post, share, and build on each other's ideas.

The value in which the knowledge building approach brings to a CSCL environment is the emphasis of valuing student agency and allowing students to figure out gaps, the challenge in addressing the gaps, and the way forward. As such, the analysis of content on the Knowledge Forum is focused on how students displayed agency and how they approach the problem, how they expand the problem, and how they manage collaboration, more than the completion of the task (Zhang et al., 2009) and this can be done in several ways, including the analysis of epistemic emotions (Teo et al., 2022). It is posited that by being able to identify, match, and conceptualize the mentioned processes into clearer, and if possible, more discrete indicators of knowledge building and therefore collaboration, it is possible to then triangulate the quality of face-to-face discourse with online discourse to attempt a better understanding of the nuances of collaboration process in CSCL environments that is less dependent on the final performance and outcomes of assigned tasks.

# 2.2 Collaboration in a Virtual Knowledge Building Environment

Opinions vary largely when asked to describe a collaborative classroom, considering the multifaceted aspects of a CSCL environment. From the knowledge building perspective, collaboration is seen as the process of students collectively figuring out the incremental steps to improve ideas as well as enacting the steps. Students increasingly take on higher-level responsibilities (e.g., defining gaps mid-way through the inquiry, determining the required experiments and resources, figuring out what additional information might be needed). As more students get involved in the collaborative process, they also begin to value the collective effort, not just tracking of individual progress, and the class accumulates a diversity of ideas, connections across these ideas, with new and novel ideas surfacing that invites further questioning. Students can work in small groups and navigate different class-based or group-based discourse spaces on the Knowledge Forum (see Figure 1), with or without the teacher, who acts as a co-constructor rather than a leader and provider of knowledge. Figure 2 shows snapshots of a class-based or group-based discussion with the respective sharings and ideas reflected on the Knowledge Forum, as possible setups for encouraging and enhancing collaboration in a knowledge building environment.

# 2.3 Tapping on Multimodal Learning Analytics for Characterization

Multimodal Learning Analytics (MMLA) refers to the collection, analysis, utilization, and integration of multimodal data to study, characterize, and inform student learning behaviors (Rodríguez-Triana et al., 2018). Previous work from Di Mitri et al. (2018) has conceptualized a classification framework to conduct extensive literature review on empirical studies in MMLA, using the findings to propose a taxonomy of multimodal data for learning, with additional research also examining the significance and influence of multimodalities in the collaborative process (Rodríguez-Triana et al., 2018).



Class View – Main page which allows all students (when front-facing) to post and share their ideas

Group View – Group page where students from same group can gather to post and share their ideas

*Figure 2.* Snapshots of class lesson (top left) and group discussions (bottom left), and screenshots (right) of their created and shared notes on the Knowledge Forum.

In all, findings from these studies were able to identify learning analytics modalities that could be used to support and enrich learner collaboration, including the use of speech, body poses, and facial expressions. As a result, multimodal data can theoretically be used to classify and inform different dimensions of learning, but the major impediment in wide-scale adoption lies in the need to convincingly weave an integrative picture that includes all necessary collaboration process that happened in a CSCL environment, in addition to the complex benchmarking effort required to ensure the validity of multimodal analysis of collaborative processes beyond the product or final artefacts. This study seeks to identify the specifics in the mentioned gaps and also attempts to address several of the difficulties in implementation.

# 2.4 Proposed Design of Conceptual Analytical Framework

Educators and researchers have long viewed teaching and learning from a variety of perspective, some of which include the sociological perspective (e.g., Jarvis, 2007) that is concerned with structure and functions of the society and educational processes that are situated within; the social psychology perspective that studies interactions between teachers, students, and other stakeholders (e.g., Gehlbach, 2010); Vygotsky and Piaget's postulations that social interaction is fundamental to cognitive development, albeit with different foci on social basis of mind and individual as starting point respectively (Rogoff, 1999); and Engeström's Activity Theory (2000) as a framework for analyzing and redesigning work. Among the many other perspectives that take a manifold of view of knowledge that analyzes interaction levels and learning processes, we consider that within an increasingly volatile and uncertain education landscape, there can be a perspective that builds on evolutionary realism, one which is based on Dopfer's micro-meso-macro architecture (2004) that is motivated by the need to handle emergent and highly complex situations as a population, a structure, and a process of rules. This adopted micro-meso-macro concept for this study is a multilevel classification and analytical framework that was successfully utilized by Lee et al. (2022) to understand and study behaviors of and relations of society on different levels, and to explain interaction patterns within an ecosystem.

In this concept, the micro is the individual, the meso are parts of the society (e.g., groups, organizations), and the macro represents the society as a whole, encompassing political, economic, social, and other factors that affects all its members. Through this framework, behavioral patterns and interactions within and between the different levels can be examined and analyzed in greater detail. Within this study, knowledge building, as an approach and pedagogy, handles theories, ideas, and practices in a structure with agents (knowledge builders in this study) and the multilevel structure of this framework can be adopted to understand ways in which different modalities can be classified to facilitate the

identification of different patterns of collaborative activity within and across different interacting members in authentic CSCL environments.

# 3. Design and Development of Analytical Framework

# 3.1 Conceptualization of Analytical Framework

The analytical framework examines the various constructs of students' collaboration in a knowledge building environment, including elements from both in-person and virtual learning environment. The stages of data organization, classification, analysis, and visualization are:

- (i) Collection of data from different modalities Students' collaborative activities in both offline (face-to-face) and online (Knowledge Forum) environments are captured using a range of sensing modalities, including video cameras, audio recorders, physiological sensors, server data logs, and surveys. Features are then generated from the visual, verbal, physiological, textual, and emotional data.
- (ii) Sense-making The micro-meso-macro concept as explained in section 2.4 is adopted to make sense of the different data modalities and to subsequently inform the interactions at the following levels and pairings.
  - Micro level Individual students interacting with artefacts and information.
  - Meso level Student collaboration in a group is seen as the meso-structure where interactions take place within a group of students.
  - Macro level Interactions within or between student groups and teacher.
  - Micro-meso pair Students' collaboration in a group shaped by individual students' speech, actions, and behaviors as the microstructure.
  - Meso-macro pair Student engagement between groups (meso-structure) can be shaped by their interactions or the teacher, who represents part of the macro-structure.
- (iii) Relating constructs of students' collaboration with the micro-meso-macro levels of interactions Table 1 shows how different analyses are categorized at either of the two levels of interactions and also based on the constructs of collaboration.

	Constructs of collaboration					
	Affective component	Social component	Shared epistemic component	Pedagogical & environmental component		
Level 2: Student group- teacher (meso- macro)	Analyze students' emotions in response to teachers'	Conduct network analysis to obtain collaborative	Analyze teachers' assistance to students in discussing	Analyze received instructions, students' interactions with		
Note: Teacher- independent collaboration may occur at this level	instructions	measures and indices Analyze	their thoughts	devices and learning artifacts in the physical space		
Level 1: Student-student group (micro-meso)	Analyze students' expressions towards challenges or collaboration within groups	communications and actions at both individual and community level	Analyze implicit and explicit knowledge artefacts that shape the collaboration	Analyze students' interactions with each other, with the devices, and on the virtual discourse space		

Table 1. How different analyses are categorized at two levels of interactions and based on constructs of collaboration

It is to be noted that at level 1, collaborative activities strictly occur without teachers and is only between students, while it is possible for students in level 2 to continue collaborating with each other with the teacher as an observer (not an active co-constructor). The analyses within each level also remain interrelated and may possibly influence other analysis within their own levels and across both levels. The interactions are also not restricted to a certain area of effect or influence, with each analysis potentially able to influence and affect the quality of students' collaborative learning processes, such as how the analysis of students' emotions in response to teacher's instructions may still affect students' collaboration within groups.

(iv) Designing visualization and feedback mechanisms to transfer the sense-making process to students – Table 2 shows how various metrics can be conceptualized across the modalities and micro-meso-macro levels.

Levels (referring to) Modalities	<b>Micro</b> (Students working on their own, conducting individual reflections, with focus on quality of individual notes on Knowledge Forum)	<b>Meso</b> (Individual students interacting with their groups, building on notes within their groups, with possible teacher interactions)	Macro (Student interacting across groups and co- constructing with their teacher in whole-class discussions using class-based view on Knowledge Forum)
Face-to-face discourse (verbal)	_	Semantic patterns of speech within group	Semantic patterns of speech across group
Online discourse (on Knowledge Forum)	Complexity of individual ideas and contribution across all views	Types of scaffold used within the group's Knowledge Forum view	Types of scaffold used within the class's Knowledge Forum view
Knowledge Forum activity	Number of created and modified notes, and scaffolds used	Number of build-on notes and read notes within the group's Knowledge Forum view	Number of build-on notes and read notes within the class's Knowledge Forum view
Motoric actions	_	Hand gestures, body movements, and head poses during group discussions	Body movements when interacting and discussing between groups
Electrodermal activity	Individual skin conductance level	Skin conductance when interacting within group	Skin conductance when interacting between groups
Emotions	Epistemic and non- epistemic (universal, basic) emotions		_
Teachers' actions			Teachers' pedagogy (context), speech, and actions

 Table 2. Conceptualized list of metrics across the modalities and micro-meso-macro levels

# 3.2 Pulling Modalities Together to Characterize and Measure Student Collaboration

Several machine learning algorithms were found to be appropriate and feasible for making sense of the different modalities. Through this study, a pilot was conducted and implemented to conduct sense-making and interpretation processes of discourse (online and face-to-face), physiological data, and epistemic emotions. To decide on an appropriate choice of methods

for the pilot, prior tests (Hartmann et al., 2019) helped to shortlist text classification methods for unstructured textual discourse, of which linear regression (LR), neural network (NN), support vector machine (SVM), and Naïve Bayes (NB) made the cut for further testing in this study. We have also taken into consideration that this shortlist of methods were widely surveyed and used in sentiment detection and classification of content categories (Hartmann et al., 2019). Convolutional neural networks (CNN) were then used to analyze student expressions, together with digital signal processing of data from physiological sensors, to determine students' epistemic emotions (Lee et al., 2023).

With the above-mentioned methods, the pilot study will be able to identify specific or combination of modalities that contribute to collaborative learning and also the extent of collaborative knowledge building activity within a CSCL environment. In short, the constructs of students' collaboration are characterized through the combined use of the analytical framework with multimodal data and various machine learning methods, so that student collaboration during knowledge building activities can be measured and characterized to levels of knowledge building.

# 4. Pilot Case Study

#### 4.1 Settings, Participants, Equipment

To obtain data for framework development and machine learning, multimodal data was collected from a Primary 5 (Grade 5) class in a Singapore school, where 20 students attended a 2.5-hour Social Studies lesson "Man and His Environment" that was crafted as a debate on the pros and cons of staying or leaving an assigned area in Southeast Asia, ranging from the Mekong river, Anak Krakatoa (Indonesian volcano), or the Sarawak rainforests. Students were encouraged to debate the present problems and potential benefits of the geographical area they were assigned to and to craft a possible solution for their predicament. Over 50 hours of individual video and audio data were collected with students' consent, in line with local ethics review and adherence to regulations for data and privacy protection. This is in addition to online Knowledge Forum data that records students' actions, activities and learning artifacts.

During the lesson, the following array of sensors were deployed: H6 audio recorder for recording verbal discourse, XiaoMi360 camera for capturing student movements and poses around a table during collaboration, GoPro camera to provide a third-person view of collaborative activities between groups, Empatica E4 wristband for students who gave additional consent, and pen-and-paper emotion surveys. Altogether, online Knowledge Forum data, verbal, physiological, and emotional data were collected and transcribed, and prepared as inputs for processing in the analytical framework with machine learning methods.

#### 4.2 Data Analysis and Findings

While working with Knowledge Forum notes and verbal data, these were broken down into discourse turns before analysis for content of speech and the type of verbal moves made by the student. The Knowledge Forum data was analyzed for content and student trace activity on the platform (e.g., number of notes read, posted, build-on) were also tracked and consolidated. Other collaborative indexes of the student's contribution on the Knowledge Forum were also extracted using the Knowledge Building Discourse Explorer (KBDeX, Oshima et al, 2012) and networks were analyzed to uncover structures of knowledge building.

The integrated set of textual data and features (selected K best for training: K=4000 for verbal, K=3000 for Knowledge Forum) was then readily used to create a machine learning model for prediction purposes. To mitigate overfitting since no new data was collected or generated, convolutional neural network was chosen for fixed input and output sizes, with regularization applied to the model to improve generalizability. Large language models (e.g., BERT) were attempted but unfeasible for smaller setups. Model training was conducted using verbal and online discourse labels, coded by two knowledge building experts based on the knowledge building levels. The levels are casual talk (CT, for verbal data) or non-content (NC,

for Knowledge Forum data); question level 1 (Q1) for simple fact-seeking questions; explanation level 1 (E1) for simple explanations without elaborations; question level 2 (Q2) for expanded explanation-seeking questions; and explanation level 2 (E2) for explanations with elaborations, evidence, and justifications. These levels are encoded from 0 to 4 accordingly. In this model of five dropout layers (to avoid over-fitting) and one gaussian noise layer (to add randomness for robustness), the training accuracy is 0.789 with a test accuracy of 0.601.

Visual data (video) consisting of motoric actions taps on the OpenPose library to analyze various skeletal movements of students and teacher when he appears in frame. For video data that capture more details of students' expressions, convolutional neural networks (CNNs) were implemented to identify one of the following nine epistemic emotions: Activeness, anxiety, boredom, confusion, curiosity, enjoyment, frustration, interest, and surprise (Lee et al., in press). Physiological markers including electrodermal activity and heart rate data were analyzed using peak-detection using the E4 software. Seven self-reporting surveys were administered between tasks throughout the lesson and the responses related to epistemic emotions were also analyzed using IBM Statistical Package for Social Sciences (SPSS) software for descriptive analysis and triangulation.

From the case study, we were able to preliminarily characterize several aspects of student collaboration using the analytical framework, which were triangulated with the teacher's instructions, and actions. For example, these characterizations include how greater student movements relates with an increase in student responses, while an increase in student communications signify better quality of contributions in speech and on the Knowledge Forum. Knowledge building contributions on the Knowledge Forum were also found to be expanded and better quality (Q2 and E2) than communicated verbally. From the analysts' point of view, several of such instances throughout the 2.5-hour lesson were identified where these constructs of student collaborations were prominent. However, limited data for training and testing may also mean that the same aspects of student collaboration are challenging to reproduce in different contexts and by other students. Nonetheless, the current findings are supportive of the framework's application in examining interaction patterns within a collaborative knowledge building environment.

# 4.3 Aggregation and Visualization of Findings

Using the above findings, it was also essential to aggregate measurements and provide a visual interface with simplified information on an interactive dashboard (see Figure 3). This dashboard with video playback and seeking functionality also provides an approximate measure called the "Fusion score" that incorporates measures obtained from the different modalities with adjustable weightages so that the end-user, be it a teacher or analyst, can understand the extent of collaboration for a group or class. With a score of 0 to 1, the prototype's fusion score is aligned to three major stages of the knowledge building process, starting with idea sharing activities that exhibit scores closer to 0, while idea improvement activities will score higher, followed by activities with rise above that will score closer to 1.

# 4.4 Limitations

Among several limitations, we address two significant limitations that have been briefly mentioned thus far and are currently being worked on. The first pertains to the limited datasets that can be used for training models, testing, and refining of the analytical framework. A larger dataset of five lessons, considered a large trove of data considering the multimodality nature, is currently considered for subsequent use, as such an inclusion will be beneficial in testing the validity and reliability of the larger system. The other limitation refers to the methodical and practical challenges that can affect the sustainability and scalability of this MMLA study, which are also highlighted by Yan et al (2022), with recommendations to ensure that MMLA innovations that used discontinued technologies are constantly upgraded and caution is taken in ensuring minimal systemic bias while using MMLA innovations.



*Figure 3.* Interactive dashboard with video playback and seeking functionality, and indicators that provide insights about extent of collaboration in a knowledge building lesson.

# 5. Conclusion

CSCL is a well-established field with many technological and theoretical advancements over the past few decades, contributing to concerted efforts directed towards the better understanding of collaborative learning processes. Through the use of knowledge building as a collaborative learning pedagogical approach, a concentrated focus on the learning processes allows researchers to glean deeper insights into aspects of students' collaboration in authentic learning environments. The use of MMLA and the proposed analytical framework brings together multiple modalities to inform learning behaviors and to characterize student collaboration with the aid of machine learning methods. Preliminary findings from a pilot case study has shown that several constructs of student collaboration can be characterized and the limitations in current form can be duly addressed with sufficient research and time.

# Acknowledgements

This study was funded by the Ministry of Education, Singapore, under the Education Research Funding Programme (OER 19/19 TCL). The views expressed in this paper are the author's and do not necessarily represent the views of the host institution.

# References

- Ben-Peretz, M. (2001). The impossible role of teacher educators in a changing world. *Journal of teacher education*, *52*(1), 48-56.
- Chan, C. K. K., & Chan, Y.-Y. (2011). Students' views of collaboration and online participation in Knowledge Forum. *Computers & Education, 57*(1), 1445-1457.
- https://doi.org/10.1016/j.compedu.2010.09.003 Di Mitri, D., Schneider, J., Specht, M., & Drachsler, H. (2018). From signals to knowledge. A concentual model for multimodal learning analytics. *Journal of Computer Assisted Learning* 
  - conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning,* 34(4), 338–349. https://doi.org/10.1111/jcal.12288
- Dopfer, K., Foster, J., & Potts, J. (2004). Micro-meso-macro. *Journal of evolutionary economics, 14*, 263-279. https://doi.org/10.1007/s00191-004-0193-0
- Engestrom, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics, 43*(7), 960-974.
- Gehlbach, H. (2010). The social side of school: Why teachers need social psychology. *Educational psychology review*, *22*, 349-362. https://doi.org/10.1007/s10648-010-9138-3

Halpin, P. F., von Davier, A. A., Hao, J., & Liu, L. (2017). Measuring student engagement during collaboration. *Journal of Educational Measurement*, *54*(1), 70-84.

Hartmann, J., Huppertz, J., Schamp, C., & Heitmann, M. (2019). Comparing automated text classification methods. *International Journal of Research in Marketing, 36*(1), 20-38. https://doi.org/10.1016/j.ijresmar.2018.09.009

- Hernández-Sellés, N., Muñoz-Carril, P., & González-Sanmamed, M. (2020). Interaction in computer supported collaborative learning: an analysis of the implementation phase. *International Journal of Educational Technology in Higher Education, 17*(1), 1–13. https://doi.org/10.1186/s41239-020-00202-5
- Jarvis, P. (2007). *Globalization, lifelong learning and the learning society: Sociological perspectives.* Routledge.
- Jeong, H., Hmelo-Silver, C., & Jo, K. (2019). Ten years of Computer-Supported Collaborative Learning: A meta-analysis of CSCL in STEM education during 2005–2014. *Educational Research Review, 28*, 1-17. https://doi.org/10.1016/j.edurev.2019.100284
- Lee, A. V. Y., & Tan, S. C. (2017). Promising ideas for collective advancement of communal knowledge using temporal analytics and cluster analysis. *Journal of Learning Analytics, 4*(3), 76-101. http://dx.doi.org/10.18608/jla.2017.43.5
- Lee, A. V. Y., Teo, C. L., & Tan, S. C. (2022). Rethinking teaching and learning with preschoolers: Professional development using knowledge building and a 3M analytical framework. *International Journal of Educational Research Open*, *3*, 100147.
- Lee, A. V. Y., Teo, C. L., Yuan, G., Ong, A., Lim, R. E. C., Bounyong, S., Juliano, F., & Zhao, A. M. (2023, in press). *Towards recognition of students' epistemic emotions in a student knowledge building design studio.* In *Proceedings of the ISLS Annual Meeting 2023*. Montreal, Canada: International Society of the Learning Sciences.
- Martinez, A., Dimitriadis, Y., Rubia, B., Gómez, E., & De la Fuente, P. (2003). Combining qualitative evaluation and social network analysis for the study of classroom social interactions. *Computers & education*, *41*(4), 353-368. https://doi.org/10.1016/j.compedu.2003.06.001
- Mercer, N., Littleton, K., & Wegerif, R. (2009). Methods for studying the processes of interaction and collaborative activity in computer-based educational activities. In K. Kumpulainen, C. E. Hmelo-Silver, & M. César (Eds.), *Investigating classroom interaction* (pp. 27-42). Brill. https://doi.org/10.1163/9789087907624\_004
- Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge Building Discourse Explorer: a social network analysis application for knowledge building discourse. *Educational Technology Research and Development, 60*(5), 903–921. https://doi.org/10.1007/s11423-012-9265-2
- Rodríguez-Triana, M. J., Prieto, L. P., Martínez-Monés, A., Asensio-Pérez, J. I., & Dimitriadis, Y. (2018, March). The teacher in the loop: Customizing multimodal learning analytics for blended learning. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK '18)* (pp. 417-426). New York, NY: Association for Computing Machinery. https://doi.org/10.1145/3170358.3170364
- Rogoff, B. (1999). Cognitive development through social interaction: Vygotsky and Piaget. In P. Murphy (Ed.), *Learners, learning and assessment* (pp. 69-82). Paul Chapman Publishing.
- Scardamalia, M., & Bereiter, C. (2003). Knowledge building environments: Extending the limits of the possible in education and knowledge work. In A. DiStefano, K.E. Rudestam, & R. Silverman (Eds.), *Encyclopedia of distributed learning*. Thousand Oaks, CA: Sage Publications.
- Suthers, D. D. (2012). Computer-supported collaborative learning. In N. M. Seel (Ed.), *Encyclopedia* of the Sciences of Learning. New York, NY: Springer.
- Teo, C. L., Ong, A., & Lee, V. Y. A. (2022). Exploring students' epistemic emotions in knowledge building using multimodal data. In A. Weinberger, W. Chen, D. Hernandez-Leo, & B. Chen. (Eds.), *Proceedings of the 15th Computer-Supported Collaborative Learning: CSCL 2022* (pp. 266-273). Hiroshima, Japan: International Society of the Learning Sciences.
- Yan, L., Zhao, L., Gasevic, D., & Martinez-Maldonado, R. (2022, March). Scalability, sustainability, and ethicality of multimodal learning analytics [Online]. In *Proceedings of the 12th International Conference on Learning Analytics and Knowledge (LAK '22)* (pp. 13-23).
- Yang, Y., van Aalst, J., & Chan, C. K. (2020). Dynamics of reflective assessment and knowledge building for academically low-achieving students. *American Educational Research Journal*, 57(3), 1241-1289.
- Zhang, J., Scardamalia, M., Reeve, R., & Messina, R. (2009). Designs for collective cognitive responsibility in knowledge-building communities. *Journal of the Learning Sciences, 18*(1), 7–44. https://doi.org/10.1080/1050840080258167

# Student Engagement Detection: Case Study on Using Peer-to-Peer Emotion Comparison with Context Regularization

**Geyu LIN, Manas GUPTA, Cheryl Sze Yin WONG & Huayun ZHANG** Institute for Infocomm Research (I<sup>2</sup>R), Agency for Science Technology and Research (A\*STAR), Singapore

{lin\_geyu, manas\_gupta, cheryl\_wong, zhang\_huayun}@i2r.a-star.edu.sg

**Abstract:** This paper describes a method to automatically assess participants' engagement in online education. Similar to emotion recognition, student engagement can be subjective. Hence, it is challenging to obtain large-enough and consistent ground-truth engagement labels for automatic student engagement. We propose an unsupervised method that could detect abnormal engagement states using peer-to-peer emotion correlation analysis in different modalities. Without any human engagement labeling, this zero-shot method accurately pinpoints the abnormal student engagements in our experiment. Modality-dependent engagement prediction also suggests possible distractions on the student's device.

Keywords: Student Engagement, Virtual Learning, Multimodality

#### 1. Introduction

Market for online education keeps expanding in recent years with the development of hardware, such as internet infrastructures, personal digital devices, etc. and the progress of software, such as online education tools and online education theories. Coronavirus sped up this procedure when schools all over the world had to shut down their physical classes and move classes online (Zheng et al., 2021). Many remote education programs outlast the pandemic. Schools and students prefer online learning to some extent because of the flexibility and convenience it offers.

Due to the intrinsic difference between real classroom and online classroom, both teachers and students are facing new challenges for online education. Without the proper classroom management and discipline restriction, students may find it hard to resist the temptation of food, drink, and other entertainment within their reach. Without face-to-face interaction with their students, teachers may not be able to gauge the effectiveness of varying teaching strategies on engaging the students. Many teachers reported they lose confidence in their ability to meet students' needs. Hence, the use of automatic student engagement detection (Baker et al, 2004) can help teachers manage the online classroom and identify the students who are disengaged. Automatically monitoring and diagnosing student engagement variation could become a teacher assistant for quality online education.

However, training machine learning models for accurate student engagement detection would require accurate ground-truth engagement labels (Khan et al., 2022). The accurate ground-truth labels could be hard to obtain due to subjective perspectives and differing cultures (Ocumpaugh et al. 2014). Hence, an unsupervised method to detect abnormal engagement states is proposed in this paper.

We examine the feasibility of analyzing students' engagement during the online classes by using learning context regularization to find out the out-of-sync student using correlation of emotion prediction of webcam video with learning contexts' emotion prediction.

# 2. Problem

Student engagement is correlated with the student emotion states and it is reflected by the student's behaviour response in the class and his/her academic improvement before and after the learning procedure. Multimodal data collected in virtual classroom, including individual performance in class quiz, average time answering the questions, activities in the classroom, what the student is watching, what the student is listening, eye contact, facial and body movement, can be collected and carefully labelled to build deep learning models detecting student engagement.

A fundamental challenge to deep learning models is the reliance on supervised learning and by extension good data - data that is large-enough and consistently labelled. Small, and inconsistent data often lead to serious performance degradation in real-world. While this is not unique to engagement detection, the lack of a widely accepted measurement for student engagement level makes it even more difficult to collect and label high quality data for engagement detection task, in addition to individual difference between students and subjective bias of human labelers.

In this study, an alternative way is adopted for detecting out-of-sync students. Instead of training deep learning model using well labelled engagement data, peer-to-peer emotion comparisons are conducted in different modalities to pinpoint abnormal engagement states in the student group.

#### Learning Context Regularization Context Regularized Emotion Synchronized Segmentation :----:--Ś video emotion state - : : p p Learning Context Regularization :--Context Regularized Engagement Detection Q o emotion state :--0.8 0.6 Learning Context Regularization - : .: . : : video otion state . : Ŋ -

# 3. Methodology

Figure 1. Block Diagram for Proposed Method

The proposed framework for detecting possible disengagement consists of three components, namely synchronized segmentation to determine the learning context, emotional states based on multiple modalities of data and the concept of peer-to-peer correlation in the same context. This is being elaborated in the following sections.

# 3.1 Synchronized Segmentation

Experiments are conducted in asynchronous settings. Students attend the virtual class at their own convenience using a pre-recorded lecturing video. By attending the virtual class, they agree that their learning procedures, including their webcam recording, their screen

activities, and audio output on their learning devices, will be saved in a multi-stream video file and it will be used for research purposes after anonymization. Student video must be segmented for engagement variation analysis along with time. Two synchronization methods are adopted for student video segmentation.

# 3.1.1 Fixed-Window-Length Segmentation

Student videos are synchronized by the instant starting time of lecturing video and chopped without overlapping. All student videos are segmented into fixed 10-sencond clips without considering the continuance of the lecturing content.

# 3.1.2 Lecturer-Speech-Content Based Segmentation

The lecturer's speech in the pre-recorded video is transcribed using openai/whisper-large-v2 (Alec et al., 2022) speech recognition model. This model was trained on 680K hours of multilingual and multi-task transcribed speech data collected from the web. While this model does produce highly accurate transcription for lecturer's speech, the corresponding timestamps it provided are insufficient and sometimes inaccurate. To obtain accurate timestamps for student video segmentation, another HMM-ANN (Dahl et al., 2012) (Hidden-Markov Model – Deep Neural Network) hybrid model was adopted for force-alignment between the audio stream in student video and the whisper transcription of the lecturing video. Student videos were then segmented into varying-length clips according to the sentence-end punctuation in the whisper transcription. Doing in this way, student videos are synchronized by lecturer-speech-content and student video clips are associated with relatively complete lecturer sentences. The average length of student video clips is around 10 seconds.

# 3.2 Multimodal Emotion States

In an online learning environment, there can be multiple modalities of data that can be captured, such as the learner behavior through video and learning context through video and audio, which can be converted into text. Hence, we leverage on the Emolysis toolkit (Ghosh et al., 2023) for video and text emotion recognition. The video modality model used for prediction is the High-Speed face Emotion (HSE) model (Savchenko et al., 2022), architecture based on EfficientNet (Tan & Le, 2018), pre-trained on AffectNet (Mollahosseini et al., 2019). The text modality model used for prediction is a pre-trained RoBERTa module (Liu et al., 2019) followed by a three-layer Deep Neural Network (DNN) that maps the latent RoBERTa features to label space. The RoBERTa is fine-tuned and DNN is trained with the CMU-MOSEI dataset (Zadeh et al., 2018).

The models predict the valence, arousal, as well as the emotion states, namely fear, anger, joy, sad, disgust, surprise, trust, anticipation and none. In our analysis, the various emotion states may not be relevant for student engagement, thus we focus on the valence and arousal values. One of the challenges in multimodality data is the variance in prediction frequency for different modalities of data. In our case, it would be the webcam video, screen video and audio (transcript to text). Hence, the calculated average values in each clip obtained using synchronized segmentation would be utilized.

# 3.3 Peer-to-Peer Correlation and Engagement in Context

Learning context is an important factor to consider when analyzing the correlation of webcam video emotion prediction between students. Students might have different reaction times as the lecture flows, hence analyzing the emotion correlations of students directly might not work well.

The learning contexts between different students should be highly correlated because they should be watching the same video. With this fact, we can pick out students that are, a) did not open the video, b) watching different video and c) not listening to the video by only analyzing the correlation of the context. By adding an additional modality of webcam video, we can identify the students that were not focusing on the lecture through making valence and arousal scores on each student.

In our paper, we have experimented with two different types of learning contexts. First, the screen recording of the lecture itself. Second, the text content generated from the audio of the video. These two factors serve as regularizations when we evaluate the correlation between students to pick out outliers from students.

We make prediction for the arousal and valence score of the context, then take an average of webcam videos' emotion prediction and context emotion prediction to get final arousal and valence score to perform correlation evaluation. With learning context regularization on webcam videos' emotion prediction, we can identify out-of-sync student easily using correlation on the combined arousal and valence score between students.

# 4. Experiments

# 4.1 Multimodal Data Collection

We use the OBS Studio open-source recording tool to collect multimodal data. The tool allows us to collect the webcam of the student, the video of the lecturer, the audio coming from the lecturer and the audio generated by the student for each recording session. The various modalities are synced by default. We extract the relevant modalities from the raw data and process them individually for running our experiments.

#### 4.2 Pilot Data Analysis

We created two outliers for evaluation to represent different types of distractions that the student may have to illustrate the feasibility of evaluating engagement level in an online learning setting.

#### 4.2.1 Student Watching Unrelated Content

Participant id003 does not open the lecture video on screen. Hence id003 should be a representative of watching unrelated content during lecture. This participant should be identified as an outlier when using the screen video and webcam video.

#### 4.2.2 Student Not Focusing on The Lecture

Participant id002 is a representation for students that are having the lecture on screen but are sleeping or looking around. This participant should be identified as an obvious outlier using screen video and audio.

# 4.3 Result

#### 4.3.1 Pairwise Correlation Coefficient (PCC) Matrix

PCC matrix consists of the pairwise correlation coefficient (PCC) of each of the participants. The (i, j)-th entry of the PCC matrix represents the PCC of prediction score of participants *i* and participant *j*. We use heatmap to visualize the matrix, the color varies from red to

green, representing the PCC from lower to higher values. Hence, if a row or column for a participant consists of many red grids, this participant should be identified as an outlier.

#### 4.3.2 Engagement Score

We assign an **engagement score** to each participant's arousal and valence prediction using the following method:

- a) Suppose we have n participants, each participant have l clips.
- b) Duplicate participants *i* 's series (n 1) times.
- c) Concatenate all other participants' series except participant i
- *d)* Calculate the PCC of two series obtained from step (b) and step (c). Assign the PCC as the engagement score for participants *i*

The engagement score is a summarized score about how a participant's prediction emotion score correlates with others. This is used as a unified measure to represent how a student is engaged during the lecture.

#### 4.3.3 Regularization Factor

The regularization factor refers to the factor that is highly correlated under the experiments setting. The regularization factors can simply be learning contexts, as discussed in section 3.3. We have two regularization factors:

- a) Screen video emotion score.
- b) Audio modality emotion score.

#### 4.3.4 Overall Score Calculation

Given the emotion score prediction (arousal, valence) from video (student and lecturer), audio and text modality. We apply average on student video arousal with each of the regularization factor and both regularization factors, then analyze the engagement score and PCC matrix to see whether we can identify the outlier from calculation. Also, we compare two segmentation techniques, a) sentence segmentation, b) uniform 10s segmentation, for webcam video and audio modalities analysis and webcam video, screen video and audio modalities analysis and webcam video.



# 4.3.5 Screen Video with webcam video

Engagement score (pcc): arousal two screen out

Engagement score (pcc): valence two\_screen\_out

Figure 2. PCC matrix and engagement score (screen + webcam) using content-based segmentation

The above are the engagement score and PCC matrix of arousal and valence for screen video and webcam video. In this section, we examine the result using screen video as regularization factor.

From the confusion matrix of valence, we can observe that the id003 is the most obvious outliner because id003 did not open the screen of instructed video. The emotion prediction of screen video will differ for id003 compared to other participants. However, from the confusion matrix of arousal, it might not be obvious that id003 is an outlier. With the help of engagement score, we can observe that id003 is a clear outlier with negative correlation with others. This shows the importance of providing a consolidated score for better interpretation.

Hence, using screen video and webcam video, we can identify outlier described in section 4.2 by considering the PCC matrix and the engagement score of valence and arousal.

#### 4.3.6 Audio With Webcam Video



Figure 3. Results for Audio with Webcam Video

The above shows the engagement score result and PCC matrix of webcam video with audio emotion as regularization factor using fix-window-length segmentation. We can observe that the outlier is not obvious in this case, because the uniform 10s segmentation cannot have a good regularization compared to sentence level segmentation due to the incomplete sentence information. Some sentence might be trimmed, such that the model cannot predict the emotion of the learning context accurately.

With content-based segmentation, we can clearly identify id002 as an outlier from both PCC matrix and engagement score of both arousal and valence. Hence, this suggests that with a good regularization factor, we can identify the outlier by comparing the correlation between different students and using only webcam video of student's face and audio emotion prediction.

#### 4.3.7 Screen Video, Webcam Video and Audio



Figure 4: Results for Audio with Webcam Video and Screen Video

From the engagement score and PCC matrix generated from fixed-window-length segmentation, it's hard to conclude the outlier because the engagement score are comparable for arousal and id001 has lower engagement score for valence than both id002 and id003. Also, from the PCC matrix, we might conclude that id001 is also an outlier besides id002 and id003.

From the result of combing three modality with content-based segmentation, we can differentiate both outlier (id002 and id003) from engagement score of arousal and valence. Even though id001 seems also have outlier behavior from engagement score of valences, the outlier behavior of arousal prediction is much more obvious. We still can identify id002 and id003 as outlier by PCC matrix of both arousal and valence. This proves that the context regularization helps identify different kinds of distractions that students may have during online learning sessions.

# 5. Conclusion

In the paper, we demonstrate two feasibilities of analyzing engagement under online learning settings, by analyzing the context regularized arousal and valence prediction as follows:

- a) Record the screen video and webcam video together, the screen video emotion prediction served as a regularization factor to help to identify student that is less engaged who might be watching other unrelated content (e.g., id003) or not listening carefully (e.g., id002).
- b) With only the students' face, the audio emotion prediction output served as an effective regularization factor because audio modality should be uniform unless student is listening to something else. Furthermore, we could identify students who are not listening carefully through a regularized score between webcam video emotion prediction and audio emotion prediction.

The content-based segmentation provides a more reliable context regularization due to the content dependency on the text modality. Using fixed-window-length segmentation, the text will be trimmed randomly. This will result in the unreliable prediction on text modality. Hence with content-based segmentation creates a more reliable context regularization when we evaluate the correlation between students.

#### Acknowledgements

We would like to thank Siti Maryam Binte Ahmad Subaidi for coordinating the data collection and post processing, Chen Hao, Jeremy Wong, Lily Hoang Huong Giang and Tan Chin Tuan (alphabetical order) for helpful discussion of this case study. We would also like to thank the Artificial Intelligence, Analytics Informatics (AI3) and Institute for Infocomm Research (I2R), A\*STAR for supporting this work.

#### References

- Alec Radford, Jong Kim, Tao Xu, Greg Brockman, Christine Mcleavey, and Ilya Sutskever (2022), Robust Speech Recognition via Large-Scale Weak Supervision. Technical Report, OpenAI.
- Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004, April). Off-task behavior in the cognitive tutor classroom: When students" game the system". In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 383-390).
- Dahl G.E., Yu Dong, Deng Li, and Acero A. (2012), Context-Dependent Pretrained Deep Neural Networks for Large Vocabulary Speech Recognition, IEEE Trans. Audio, Speech, and Language Proceeding, 33-42.
- Ghosh, S., Cai, Z., Gupta, P., Sharma, G., Dhall, A., Hayat, M., & Gedeon, T. (2023). Emolysis: A Multimodal Open-Source Group Emotion Analysis and Visualization Toolkit. arXiv preprint arXiv:2305.05255.
- Khan, S. S., Abedi, A., & Colella, T. (2022). Inconsistencies in the Definition and Annotation of Student Engagement in Virtual Learning Datasets: A Critical Review. arXiv preprint arXiv:2208.04548.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). ROBERTA: A robustly optimized BERT pretraining approach. arXiv (Cornell University). <u>https://doi.org/10.48550/arxiv.1907.11692</u>
- Ocumpaugh, J., Baker, R., Gowda, S., Heffernan, N., & Heffernan, C. (2014). Population validity for educational data mining models: A case study in affect detection. British Journal of Educational Technology, 45(3), 487-501.
- Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2019). AffectNet: A database for facial expression, Valence, and arousal computing in the wild. IEEE Transactions on Affective Computing, 10(1), 18–31. https://doi.org/10.1109/taffc.2017.2740923
- Savchenko, A. V., Savchenko, L. V., & Makarov, I. (2022). Classifying emotions and engagement in online learning based on a single facial expression recognition neural network. IEEE

Transactions on Affective Computing, 13(4), 2132–2143. <u>https://doi.org/10.1109/taffc.2022.3188390</u>

- Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. International Conference on Machine Learning, 6105–6114. http://proceedings.mlr.press/v97/tan19a/tan19a.pdf
- http://proceedings.mlr.press/v97/tan19a/tan19a.pdf Zadeh, A. B., Liang, P. P., Poria, S., Wang, Z., & Morency, L. (2018). Multimodal Language Analysis in the Wild: CMU-MOSEI Dataset and Interpretable Dynamic Fusion Graph. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). https://doi.org/10.18653/v1/p18-1208
- Zheng, M., Bender, D. & Lyon, C. Online learning during COVID-19 produced equivalent or better student course performance as compared with pre-pandemic: empirical evidence from a school-wide comparative study. BMC Med Educ 21, 495 (2021). https://doi.org/10.1186/s12909-021-02909-z

# Effectiveness of Synchronous and Asynchronous Online Learning in Bloom's Taxonomy Framework

Cheryl Sze Yin WONG<sup>a</sup>, Lily HOANG<sup>b</sup> & Chin Tuan TAN<sup>a\*</sup>

<sup>a</sup>Institute for Infocomm Research (I2R), Agency for Science Technology and Research (A\*STAR), Singapore
 <sup>b</sup>Institute of High Performance Computing (IHPC), Agency for Science Technology and Research (A\*STAR), Singapore
 \*{tan chin tuan}@i2r.a-star.edu.sg

Abstract: Synchronous and asynchronous Learning is usually highly-structured, with course activities, assessments, and technology requirements explicitly written and scheduled. It allows great flexibility and accessibility for learners with strict alignment to the learning outcomes and module learning objectives. However, it is not clear that such synchronous or asynchronous learning is effective in different level of courses which may require different level of cognitive process, knowledge dimensions and combination of both. In this paper, we realigned meta-analysis data by Zeng et al. 2023 in cognitive process and knowledge dimensions under revised Bloom's Taxonomy framework. We proposed a metric to compute cognitive complexity level on the re-aligned data. Our proposed metric was able to show that synchronous classes are more suitable for basic cognitive domain learning of factual knowledge and conceptual knowledge, while asynchronous classes are more suited for higher cognitive domain learning of procedural knowledge and metacognitive knowledge. Likewise. our further grouping in educational level were also able to show that asynchronous classes are more effective for learning at higher education levels when the subjects of learning are usually in higher knowledge dimensions. Both observations are congruent in expected trend of cognitive complexity.

Keywords: synchronous learning, asynchronous learning, Bloom's taxonomy

#### 1. Introduction

With the rapid development of the internet and its technologies, online learning has become increasingly popular due to the convenience it offers. Students can stay in the comfort of their homes to learn and gain access to a wider variety of materials. Online learning can be typically divided into two categories – asynchronous and synchronous learning. Synchronous learning is the learning environment where the teacher and student are online and interact at the same time – this typically describes video conferences or webinars. On the other hand, asynchronous learning allows learners to learn at their own time and pace, interact with each other over temporally delayed time periods – through learning management systems (LMS) such as Coursera, consisting of forums and pre-recorded lectures.

The evaluation of online students' learning can be categorized into (1) Learner's perception, (2) Learner's Process and (3) Learner's Product (Hew et al., 2004). Learner's perception can generally be obtained through the learner post-survey on if the course was useful and easy to follow. Learner's process, on the other hand, is more complex, as one attempts to track various types of learning such as cognitive processes, meta-cognitive processes, social construction of knowledge, collaboration processes and problem-solving processes among the learners (Hew et al., 2004). Lastly, the Learner's product measures whether the learner has met the learning outcome through an assessment tool. Many studies have investigated the effectiveness of synchronous and asynchronous learning – evaluating

whether one is better than the other using the Learner's product. (Zeng & Luo, 2023) performed a meta-analysis on these studies – they gathered and filtered studies that compared synchronous and asynchronous online learning with an assessment tool. They concluded that an asynchronous learning environment was shown to be more effective in achieving better learning outcome or as least as good as synchronous learning. However, some studies also indicate that synchronous learning may be preferred in learner perception (Fabriz et al., 2019; Amer, 2018).

When designing courses and lessons, teachers typically utilize the Bloom's taxonomy framework (Bloom, 1956; Anderson and Krathwohl, 2001) to write the learning outcomes that the lesson aims to achieve. Bloom's taxonomy framework consists of different levels of cognitive processes - *remember, understand, apply, analyze, evaluate, create (from low to high level of cognition)*. The success of the lesson is typically measured when the proposed learning outcomes are met. Therefore, we seek to explore if the mode of online learning (synchronous or asynchronous) is dependent on the levels of cognition that the lesson aims to achieve. This could help teachers design their online courses better. In this preliminary study, we propose to further expand the meta-analysis conducted in (Zeng & Luo, 2023), using revised Bloom's taxonomy (Anderson and Krathwohl, 2001) for cognitive profiling.

# 2. Background

#### 2.1 Synchronous and Asynchronous Online Learning

Online learning environments can be classified based on synchronicity into two types: synchronous and asynchronous environment (Ebner & Gegenfurtner, 2019). In the synchronous online learning environment, the teacher and students gather on an online platform at the same time, in which instruction and possibly "real-time" interaction takes place. On the other hand, in an asynchronous environment, learning materials are prepared by the teacher and made accessible to students, who can then peruse the contents and learn at their own time and pace. Both synchronous and asynchronous environments offer advantages and drawbacks. Synchronous learning allows for immediate clarification of doubts and brainstorming since everyone is simultaneously present, but the learning pace is invariable among all participants, possibly leading to some students finding the pace too fast or slow. Whereas asynchronous learning offers teachers and students the freedom to work at their own pace and preferred time, but without the ability to interact in real-time with others.

Past studies have reached different conclusions as to whether synchronous or asynchronous is more effective at promoting learning. Kubey et al. (2001) showed that learners in an asynchronous environment performed better academically than those in a synchronous one. Perera and Richardson (2010) suggested that asynchronous approaches are more effective due to the larger amount of learning materials offered. Conversely, Somenarain et al. (2010) showed that synchronous online learning is better at improving students' conceptual understanding in a biology course. Other studies concluded that synchronous and asynchronous approach produce the same outcome in terms of students' conceptual understanding and academic performance (Duncan et al., 2012; Lim et al., 2022).

#### 2.2 Bloom's Taxonomy

Original work on Bloom's Taxonomy (Bloom, 1956) is basically a hierarchical paradigm that divides into cognitive process namely, Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating, in descending order of complexity for curriculum development and assessment to give teachers the tools they need to scaffold learning experiences and monitor their students' development throughout different cognitive skills. After revision in 2001, the revised Bloom's taxonomy (Anderson and Krathwohl, 2001) included four cognitive dimensions: factual knowledge, conceptual knowledge, procedural knowledge, and metacognitive knowledge in parallel with cognitive processes that students acquire and develop towards mastery of a subject.
Bloom's Taxonomy has been an excellent guide in planning for student learning and learning outcomes. In a recent work by Zeng and Luo (2023), they investigated the effectiveness of synchronous compared to asynchronous online learning by conducting a meta-analysis of previous studies published between 2002 and 2022 across multi-disciplines based on the effect sizes computed in synchronous and asynchronous online learning environments. Their result shows that asynchronous learning is slightly more effective than synchronous learning in promoting students' knowledge. However, the overall effect size does not show a clear trend of change across disciplines and education levels. In this paper, we attempt to re-examine the study under the Bloom's Taxonomy with our proposed metric to highlight the effectiveness of asynchronous and synchronous classes in delivering context of different cognitive complexity.

## 3. Methodology

We re-aligned the asynchronous vs synchronous learning data in Zeng and Luo (2023) in cognitive process dimension and knowledge dimension in a model created by Rex Heer in 2012 (Heer, 2012). We also attempt to group the effect size with our proposed metric for cognitive knowledge (which will be elaborated in subsequent section) to see the effectiveness of asynchronous/synchronous learning in acquiring knowledge at different complex levels and establish the appropriate learning setup for different educational levels.

# 3.1 Proposed Metric for Cognitive Complexity Level

	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual	List (1)	Summarize (2)	Respond (3)	Select (4)	Check (5)	Generate (6)
Conceptual	Recognize (2)	Classify (3)	Provide (4)	Differentiate (5)	Determine (6)	Assemble (7)
Procedural	Recall (3)	Clarify (4)	Carry out (5)	Integrate (6)	Judge (7)	Design (8)
Metacognitive	Identify (4)	Predict (5)	Use (6)	Deconstruct (7)	Reflect (8)	Create (9)

Figure 1. The cognitive complexity level metric based on revised Bloom's Taxonomy

The cognitive complexity level metric was designed with the revised Bloom's Taxonomy is shown in Figure 1. The lowest cognitive level of 1 is given to the cognitive process of listing and the highest cognitive level of 9 is allocated to the cognitive process of creating.

For each study, we considered the following factors when assigning the cognitive complexity level.

- The educational level of the participant
- The disciplinary field of the participant
- The assessment tool used for determination of achievement

The cognitive complexity level for each course is the sum of cognitive processes present. This provides an indicative score of the cognitive complexity for the given course. There was a total of three raters for the identification of cognitive complexity level, the average score between the three raters were taken.



Figure 2. Effect size vs Averaged cognitive complexity score.

The relationship between the effect size and averaged cognitive complexity score is illustrated in Figure 2. The pink region represents the area of lower cognitive complexity score of 30 and below while the green region represents the area of higher cognitive complexity score of 30 and above. The legend reflects the educational level of the different studies. From Figure 2, the K-12 studies generally have courses that falls below cognitive complexity of 30 while the higher education and professional training studies have varied levels of cognitive complexity in the courses. The grey vertical portion of the plot reflects the region of uncertainty where the effect sizes fall below +/-0.2. Many studies fall in that region, indicating that there may not be conclusive evidence that synchronous or asynchronous learning is better. However, there is a trend suggesting studies with higher cognitive complexity score are better delivered in asynchronous classes.

#### 4. Discussion

A good and well-thought-out course design is essential. Synchronous and asynchronous courses are usually highly-structured. The course activities, assessments, and technology requirements should all be explicitly written and scheduled, while also allowing for accessibility for learners with disabilities. They should also be aligned with the learning outcomes and module learning objectives.

Our proposed metric for cognitive complexity level on the studied data that was re-aligned in cognitive process and knowledge dimensions exhibits a trend that suggests synchronous classes are more appropriate for delivering content which requires mostly basic cognitive processing on factual knowledge and conceptual knowledge; while asynchronous classes could be better suited for procedural knowledge and metacognitive knowledge, which require higher cognitive process in learner. Likewise, our further grouping in educational level were also able to show that asynchronous classes are more effective for learning at higher education levels when the subject of learning are usually in higher knowledge dimensions. Both observations are congruent in the expected trend of cognitive complexity.

The effect sizes from (Zeng & Luo, 2023) were calculated based on the outcome of assessments. However, many of the assessment tools that were shared are multiple choice quizzes that seem to be standardized assessments in traditional learning environment. Studies showed that math and reading can be improved using digital formative assessments (See, et al, 2021). Likewise, Ohio state synchronous vs asynchronous class data in K–12 schools delivering online education, showed that these students performed worse in standardized assessments than those in traditional charter and traditional public schools (Ahn

and McEachin, 2017). In a review of 52 asynchronous class vs synchronous class, Yan et al (2020) found that personal and contextual factors more than just teacher ability and motivation in conducting formative assessment. A tightly integrated support system is a crucial factor. (See, et al, 2021). Bloom's Taxonomy is basically a core concept for categorizing cognitive skills that emphasize the necessity of developing learners' critical thinking and higher-order cognitive skills and are deemed to expanded upon inclusion of digital technology. Educators can use the taxonomy to develop engaging and relevant learning experiences that match the demands of the digital age by recognizing their commonalities, distinctions, and relative contributions. Digitalizing Bloom's Taxonomy utilizes technology for better learning outcomes by lining up the cognitive levels of Bloom's Taxonomy with digital actions and skills.

#### Acknowledgements

We would like to thank the Artificial Intelligence, Analytics & Informatics (AI3) and Institute for Infocomm Research (I2R), A\*STAR for supporting this work.

#### References

- Amer, A.A.M. K. (2018). COMPARING THE EFFECTS OF PRE-RECORDED LECTURE AND LIVE ONLINE LECTURE ON LEARNING AND SENSE OF COMMUNITY.
- Anderson, L. W. and Krathwohl, D. R., et al (Eds..) (2001) A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Allyn & Bacon. Boston, MA (Pearson Education Group)
- Ahn, J., and McEachin, A. (2017). Student Enrollment Patterns and Achievement in Ohio's Online Charter Schools.Educational Researcher, 46(1), 44–57. DOI: https://doi.org/10.3102/0013189X17692999.
- Armstrong, P. (2010). Bloom's Taxonomy. Vanderbilt University Center for Teaching. Retrieved July 19, 2022 from https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/.
- Berry, S. (2017). Educational Outcomes of Synchronous and Asynchronous High School Students: A Quantitative Causal-Comparative Study of Online Algebra 1. Northeastern University ProQuest Dissertations Publishing, 2017. 10633522.
- Bloom, B. S. (1956). "Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain." New York: David McKay Co Inc.
- Carnahan, C. (2012). The Effects of Learning in an Online Virtual Environment on K-12 Students.
- Constantine, M.B. A study of individual learning styles and e-learning preferences among community health aides/practitioners in rural Alaska. Ph.D. thesis, TUI University.
- Duncan, K., Kenworthy, A., & McNamara, R. (2012). The effect of synchronous and asynchronous participation on students' performance in online accounting courses. Accounting Education, 21(4),
- Ebner, C., & Gegenfurtner, A. (2019). Learning and Satisfaction in Webinar, Online, and Face-to-Face Instruction: A Meta-Analysis. Frontiers in Education, 4,
- Emmanouilidou, K., Derri, V., Antoniou, P., & Kyrgiridis, P. (2012). Comparison between synchronous and asynchronous instructional delivery method of training programme on in-service physical educators' knowledge. Turkish Online Journal of Distance Education, 13(4), 193-208.
- Fabriz, S., Mendzheritskaya, J., & Stehle, S. (2021). Impact of synchronous and asynchronous settings of online teaching and learning in higher education on students' learning experience during COVID-19. Frontiers in Psychology, 12. <u>https://doi.org/10.3389/fpsyg.2021.733554</u>
- Heer, R.(2012) A Model of Learning Objectives based on A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives (Iowa State University, Center for Excellence in Learning and Teaching ; www.celt.iastate.edu/teaching/RevisedBlooms1.html).
- Hew, Khe & Liu, Shijuan & Martinez, Ray & Bonk, Curt & Lee, Ji-Yeon. (2004). Online Education Evaluation: What Should We Evaluate?. Association for Educational Communications and Technology.
- Hill, S. (2009). An investigation of the impact of asynchronous online learning on student achievement. Capella University ProQuest Dissertations Publishing, 2009. 3387212.

- Khodaparast, F., & Ghafournia, N. (2015). The Effect of Asynchronous/Synchronous Approaches on English Vocabulary Achievement: A Study of Iranian EFL Learners. *English Language Teaching*, *8*, 117-127.
- Kubey, R. W., Lavin, M. J., & Barrows, J. R. (2001). Internet use and collegiate academic performance decrements: Early findings. Journal of Communication, 51(2),
- Le, K. (2022). Pre-Recorded Lectures, Live Online Lectures, and Student Academic Achievement. Sustainability, 14(5), 2910. <u>https://doi.org/10.3390/su14052910</u>
- Lim, C. L., She, L., & Hassan, N. (2022). The Impact of Live Lectures and Pre-recorded Videos on Students" Online Learning Satisfaction and Academic Achievement in a Malaysian Private University. International Journal of Information and Education Technology, 12(9), 874–880.
- Ogbonna, C. G., Ibezim, N. E., & Obi, C. A. (2019). Synchronous versus asynchronous e-learning in teaching word processing: An experimental approach. South African Journal of Education, 39(2), 1-15.
- Perera, L., & Richardson, P. (2010). Students' use of online academic resources within a course web site and its relationship with their course performance: An exploratory study. Accounting Education: An International Journal, 19(6), 587–600
- Saif, N., Niotis, K., Dominguez, M., Hodes, J. F., Woodbury, M. J., Amini, Y., Sadek, G., Scheyer, O., Caesar, E., Hristov, H., Knowlton, N., Lee, P., McInnis, M., & Isaacson, R. S. (2020). Education Research: Online Alzheimer education for high school and college students. Neurology, 95(16), e2305–e2313. https://doi.org/10.1212/wnl.00000000009859
- See, B., Gorard, S., Lu, B., Lan, D., Siddiqui, N. (2021) Is technology always helpful?: A critical review of the impact onlearning outcomes of education technology in supporting formative assessment in schools, Research Papers in Education, DOI:10.1080/02671522.2021.1907778
- Somenarain, L., Akkaraju, S., & Gharbaran, R. (2010). Student perceptions and learning outcomes in asynchronous and synchronous online learning environments in a biology course. MERLOT Journal of Online Learning and Teaching, 6(2), 353–356.
- Suliman, M., Ta'an, W., Abdalrhim, A., Tawalbeh, L. I., & Aljezawi, M. (2022). The impact of online synchronous versus asynchronous classes on nursing students' knowledge and ability to make legal and ethical decisions. Nurse Education Today, 109, 105245. https://doi.org/10.1016/j.nedt.2021.105245
- Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, H. (2021). A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment. Assessment in Education: Principles, Policy & Practice, 28(3), 228-260.
- Zeng, Hang & Luo, Jiutong. (2023). Effectiveness of synchronous and asynchronous online learning: a meta-analysis. Interactive Learning Environments. 1-17. 10.1080/10494820.2023.2197953.
- Zhu, X., Shek, D. T. L., & Chan, C. H. M. (2021). Promoting Service Leadership Qualities and Well-Being among University Students through an Online Course during COVID-19 Pandemic. International Journal of Environmental Research and Public Health, 18(15), 8162. <u>https://doi.org/10.3390/ijerph18158162</u>

# WIPP

# Task-based Robot-assisted learning to support L2 Speaking Practice

Cheng-Yueh JAO<sup>a</sup>, Hui-Chin YEH<sup>a\*</sup>, Shih-Hsien YANG<sup>d</sup>, Ming-Chang WU<sup>c</sup> & Chen-Fu WANG<sup>d</sup>

<sup>a</sup> Department of Applied Foreign Languages, National Yunlin University of Science and Technology, Taiwan

<sup>b</sup> Department of Applied Foreign Languages, National Formosa University, Taiwan <sup>c</sup> Graduate School of Technological and Vocational Education, National Yunlin University of Science and Technology, Taiwan

<sup>d</sup> Graduate School of Design, National Yunlin University of Science and Technology, Taiwan

#### \*hyeh@yuntech.edu.tw

Abstract: This study examined the impact of task-based robot-assisted interactive presentation creation on the students' speaking skills. The study used a paired-sample t-test to compare students' English speaking scores before and after the intervention, with robot-assisted interactive presentation creation as the independent variable and English speaking skills as the dependent variable. The results of the study showed a significant improvement in students' English speaking skills after the intervention. The study's results provide evidence for the effectiveness of robot-assisted instruction in enhancing English speaking skills among students. The interactive nature of the presentations created with the help of robots may have contributed to the significant improvements observed in this study. By engaging students in the creation of their own presentations, the intervention may have increased their motivation to improve their English speaking skills.

Keywords: Robot-assisted language learning, task-based learning, English speaking

#### 1. Introduction

The productive skills enable language learners to produce language in spoken and written forms. These language skills are an integral part of communication as they proffer EFL learners the ability to harness their language knowledge into real-life practice. However, many Asian language learners pay more attention to their receptive skills that are listening and reading rather than productive skills, namely, speaking and writing because of the nature of school curriculum (Putri et al., 2017). Syzenko and Diachkova (2020) explained that compared with receptive skills, productive skills and other critical thinking competence like reasoning and problem-solving tend to be neglected in the EFL curriculum in Asia. A lack of opportunities for real language use render learners' acquisition of extensive knowledge about the target language (know what), while they are not strong when using the language in a contextual way (know how). Hsieh (2016) added that most EFL teaching materials are based on textbooks aiming at preparing EFL learners for a variety of language examinations instead of focusing on real-life connection. This means that grammatical correctness takes precedence over pragmatic accuracy because of the supremacy of fundamental syllabus (Blake, 2016). As a result, many EFL learners tend to be passive learners who only gain information or knowledge without engaging with the received information or the previous learning background. They tend not to interact with others, share opinions, or contribute to a dialogue for a more thorough discussion (Fu & Yang, 2019). To encourage learners to be more proactive, creating authentic learning situations where learners can develop their productive language skills is the key. To achieve this, one of the prevalent solutions to building EFL learners' productive language skills is artificial intelligence (AI). AI has caught researchers and educators' attention to integrate it into educational fields in recent years since AI specialists are craved in the employment market. Ministry of education in Taiwan has proposed an AI

and emerging technology project in 2019 to cultivate students' AI competencies by encouraging teachers to incorporate robots with school subjects. Robot-assisted learning could provide essential affordances for knowledge construction and learner-centered pedagogy due to the varied and flexible characteristics of robots. Through the assistance of robots, teachers could easily promote personalized and adaptive learning in their teaching, enabling students to effectively master target knowledge with focused remediation. Although robots offer various benefits for learners and teachers to facilitate their learning, the challenges and difficulties of integrating technology with language learning still are worth further exploration (Tafazoli & Gomez-Parra, 2017), especially in developing learners' productive competence. Through a task-based instructional design using codes and robots, students are expected to present their ideas in both speaking and writing while applying their critical thinking and problemsolving competence. By doing so, more opportunities are created to assist students in building their productive competence while working with robots. This fills the gap of what has been missing in the current EFL curriculum in most of the Asian countries, a lack of developing productive language competence. This study thus applied a task-based instructional design with the integration of robots to engage students in creating interactive presentation through coding and multimedia design so as to facilitate students' productive skills. The guided questions are shown as follows. (1) What were the effects of task-based instruction using codes and Dash to perform the interactive presentations on EFL students' speaking skills?

#### 2. Method

#### 2.1 Participants

The participants will be chosen in a two-credit undergraduate course for first year students entitled "Oral Communication" at the University of Science and Technology. This course is part of the required core curriculum for students from the Department of English. Approximately 60 first-year students are expected to register in the "Oral Communication" course, with English proficiency levels ranging from 500 to 600 – following TOEIC scores. This range of scores indicates that they could read and write simple texts and hold a basic conversation about familiar topics or those of personal interests. Each student's willingness to participate in this study will be confirmed each year by their signing of a consent form, as well as keeping them informed about the research scope and data to be collected. All related documents will be submitted to the Human Subject Committee for approval.

#### 2.2 Research design

The instructor, who serves as the principal investigator, will offer a required course "Oral Communication" which will run for 18 weeks. Besides teaching English skills, the course aims to develop students' coding skills to perform robot activities. In addition to the course, students will receive in-class teaching and after-class online coding practice for six weeks. In the first week, learners will be introduced to coding content and the Blockly app tool, and be informed about the RALL contentcreating project. From the second week onwards, students will first focus on hands-on practice with programming in order to build up their coding skills and complete robot activities. They will work individually and in groups to get familiar with the coding practice including sequencing and debugging. They will work on some exercises provided by the code.org workspace to complete a variety of robot activities. During the 7th and 8th weeks, students will start to plan their rundown of the storytelling for interactive presentations. They will adapt a story and twist a bit on its content to make something original and show the story through manipulating Dash robots by coding. They will complete a project planner worksheet and a digital story planning sheet. After the group discussion, they will start to work on their coding to manipulate Dash robots for the story presentation part. During this process, they may need to debug their codes if errors occur in their coding. They will write down their story, make some recording to manipulate the robots etc. By the eighth week, students will present their story through the RALL project to demonstrate their speaking skills. The instructor will also evaluate the performance based on the rubrics for speaking skills and give feedback.

#### 2.2.1 Data collection and analysis

The collected data will include 60 students' productive language competence on their individual initial speaking skills based on the designed rubric assessment (pre-test on speaking skills) and their productive language competence on their individual speaking skills after learning coding based on the designed rubric assessment (post-test on speaking skills). A paired-sample t-tests will be performed. The t test was intended to compare the mean of pre-and post-tests for speaking skills. Thus, the purpose of these the paired-sample t test will reveal the mean difference between pre-and post-tests for speaking tasks.

#### 3. Results

#### 3.1 Students' improvement in English speaking skills after performing interactive presentation

The study utilized a paired-sample t-test to investigate the impact of robot-assisted interactive presentation creation on students' English speaking skills. The study's independent variable was the intervention of robot-assisted interactive presentation creation, while the dependent variable was students' English speaking skills. The results of the study indicate a significant improvement in students' pre- and post-intervention English-speaking scores in terms of accuracy (p < .00), fluency (p < .00), and overall speaking ability (p < .00). Furthermore, the effect size was found to be large in accuracy (d = 1.65), fluency (d = 1.62), and overall English-speaking ability (d = 1.75), suggesting a significant improvement in students' English-speaking accuracy, fluency, and overall ability after the intervention. These findings provide strong support for the hypothesis that robot-assisted interactive presentation creation has a positive impact on students' English-speaking skills.

#### 4. Discussion and conclusion

The objectives of this study are to explore the effects of robot-assisted multimodal composition on EFL students' productive competences through engaging them in creating multimodal presentation with the aid of robots and coding. The results revealed that the students made significant progress in developing fluent and accurate speaking performance, which provide a coherent line of argument from educational practices in developing productive competences through multimodality coupled with robot technology, which has been missing from the literature. Therefore, this study can inform EFL educators and researchers ways to engage EFL learners in practicing using the target language for productive output through the affordance available in social robots.

#### 5. References

- Blake, R. (2016). Technology and the four skills. Language Learning & Technology, 20(2), 129–142. http://dx.doi.org/10125/44465
- Fu, J. S., & Yang, S.-H. (2019). Exploring How YouGlish Facilitates EFL Learners' Speaking Competence. Educational Technology & Society, 22 (4), 47-58.
- Hsieh, H. C. (2016). High school students' topic preferences and oral development in an English-only shortterm intensive language program. English Language Teaching, 9(9), 116–133.
- Putri, N. L. P. N. S., Artini, L. P., & Nitiasih, P. K. (2017). Project-based learning activities and EFL students' productive skills in English. Journal of Language Teaching and Research, 8(6), 1147-1155.
- Tafazoli, D., & Gómez Parra, M. E. (2017). Robot-assisted language learning: Artificial intelligence in second language acquisition. Current and future developments in artificial intelligence, 1, 370.
- Syzenko, A. & Diachkova, Y. (2020). Building cross-cultural competence in a foreign language through technology-enhanced project-based learning. Amazonia Investiga, 9(27), 411-418.

# Designing an IoT-based Biorobotic Complex Board Game <Eurasia Channel>

#### Song-Lin CHEN<sup>\*</sup>, Ju-Ling SHIH, Shaun-Wen CHEN

Department of Network Learning Technology, National Central University, Taiwan \*ccssll120061203@gmail.com

**Abstract:** <Eurasia Channel> integrates Internet of Things (IoT) technology to create a captivating and intricate complex board game experience for learners. Set against the backdrop of ancient China's Silk Road, the game offers players a diverse range of tasks while delving into the historical knowledge of the Silk Road. The role of IoT technology is crucial, facilitating the transmission of a wide array of data between objects and servers, enabling the study of learners' historical thinking performance during their engagement with the game. This game also places significant emphasis on interaction and entertainment, skillfully blending elements from large-scale complex board gaming to enhance the learning process. <Eurasia Channel> seamlessly merges education, technology, and entertainment, opening a gateway for learners to explore both ancient history and modern technology in an engaging manner.

Keywords: Internet of Things, Complex Board Game, Silk Road, Biorobotic.

## 1. Introduction

IoT technology remains underutilized in STEM education and game-based learning (Mavroudi et al., 2012). This study seeks to innovate traditional board games by incorporating IoT, enhancing player experience. In formal education, textbooks dominate, yet technology prompts varied learning tools, including board games (Paris & Yussof, 2012). Board games traditionally lack dynamism, but IoT envisions seamless interconnection, adding dynamism through sensors and microcontrollers (Pai, 2017). This study aims to develop a complex board game called <Eurasia Channel> that merges education, technology, and entertainment, bridging ancient history and modern tech for learners.

# 2. Related work: Internet of Things in Education

The Internet of Things (IoT) is a technology pioneered by the MIT Auto-ID Center. The core functionality of IoT revolves around communication between devices (Pai, 2017). All objects utilize RFID as their unique identifier (Haller, Karnouskos, & Schroth, 2009). Existing literature predominantly focuses on the technical aspects of IoT (Gong, 2016). However, IoT has profound implications for the education sector, bringing forth numerous benefits (Pai, 2017), such as: (1) enhancing the learning experience, (2) improving operational efficiency, (3) reducing costs, (4) ensuring reliability, and (5) addressing security concerns. Therefore, IoT technology is a crucial and efficient asset in the field of education. In this study, we have designed a board game that leverages IoT technology to enhance engagement and interactivity.

# 3. Game Design

<Eurasia Channel> is based on the design by Chen, Shih, and Chen (2022), which takes inspiration from the ancient Silk Road map of China. In this board game, players start on a journey using biorobotic horses to complete specific tasks and progress through the game

(as shown in Figure 1). Each team represents distinct ethnicities, and individual players assume unique roles, engaging in trading various items or constructing stations to earn coins and emerge victorious. To enhance the interactive experience, IoT technology is seamlessly integrated into the game, responsible for recording all transaction records among players, tracking object locations and movement paths, and triggering specific tasks and scores through MQTT signals sent to a server. This integrated approach empowers players to easily control the biorobotic horses, while enjoying the game's interactivity, ultimately elevating immersion and underscoring the indispensability of IoT technology in educational gaming.



Figure 1. Map of < Eurasia Channel >

With the integration of Internet of Things (IoT) technology, players can effortlessly maneuver the biorobotic horses by using a terminal. As the biorobotic horse traverses designated areas, sensors transmit specific signals to the server via Message Queuing Telemetry Transport (MQTT) (Hunkeler, Truong, & Stanford-Clark, 2008). Subsequently, players receive targeted missions or scores. Upon successful completion of in-game tasks, the game map engages in interactive exchanges with players (as shown in Figure 1 錯誤! 找不 到參照來源。). During gameplay, players have the opportunity to acquire knowledge in mechanics, electronics, geography, and history. <Eurasia Channel> incorporates Internet of Things (IoT) technology through the implementation of the MQTT protocol. Upon network establishment, RFID sensors are utilized to scan embedded RFID tags on the Silk Road map (Figure 1 錯誤! 找不到參照來源。). Each tag's unique ID triggers distinct events displayed on terminals. The server records individual player checkpoints, data, movements, statuses, resources, and paths.

Engaging in < Eurasia Channel> enables players to partake in interdisciplinary learning. The integration of IoT technology allows the server to capture various data types, facilitating communication among different components. This enhances player engagement and interactivity, ultimately fostering the development of historical thinking skills. With all RFID sensors in place on the map, IoT devices can transmit tag IDs to specific channels on the MQTT Broker. The process begins with the biorobotic horses' movement on the map. Upon reaching a designated area, the RFID scanner continuously scans the path beneath the robot. Detection of a tag ID prompts the ESP32 to publish data to the designated Broker channel. Subsequently, the server subscribes to this channel, retrieves the data, and stores it in the database. Ultimately, users can access the database to control the biorobotic horses in accordance with the obtained data. System structure is as shown in Figure 2.



## 4. Excepted Result

By integrating Internet of Things (IoT) technology into the complex board game, players are afforded the opportunity to learn about ancient cultural history and cultivate historical thinking while accomplishing tasks. The inclusion of these in-game tasks serves to enhance players' motivation to learn, facilitating a continuous state of engagement for learners throughout the process of playing and learning. Teachers can effortlessly record all essential information, including task completion, scores, and the paths of each team. Beyond collecting these quantitative data points, we can further delve into various aspects related to players and the game. For instance, we can explore the impact of playing <Eurasia Channel> on players' historical thinking abilities.

## Acknowledgements

This study is supported in part by the National Science and Technology Council (previously known as Ministry of Science and Technology) of Taiwan, under MOST 108-2511-H-008 - 016 -MY4 and NSTC 112-2410-H-008 -010 -003.

# References

- Chen, S. W., & Shih, J. L., & Chen, Y. M. (2022, November). Four-bar Linkage Quadruped Biorobotic Instructions: Gamified Design and Development. *The 30th International Conference on Computers in Education (ICCE 2022)* Kuala Lumpur, Malaysia.
- Gong, W. (2016). The internet of things (IOT): What is the potential of the internet of things (IOT) as a marketing tool?. Bachelor's thesis, University of Twente. https://purl.utwente.nl/essays/70018
- Grgić, K., Špeh, I., Heđi, I. (2016). A web-based IOT solution for monitoring data using MQTT protocol. *International Conference on Smart Systems and Technologies (SST)*. https://doi.org/10.1109/sst.2016.7765668
- Haller, S., Karnouskos, S., Schroth, C. (2009). The internet of things in an enterprise context. *Lecture Notes in Computer Science*, 14–28. https://doi.org/10.1007/978-3-642-00985-3\_2
- Hunkeler, U., Truong, H. L., Stanford-Clark, A. (2008). Mqtt-S—A publish/subscribe protocol for Wireless Sensor Networks. 3rd International Conference on Communication Systems Software and Middleware and Workshops (COMSWARE '08). https://doi.org/10.1109/comswa.2008.4554519
- Mavroudi, A., Divitini, M., Mora, S., Gianni, F. (2018). Game-Based Learning for IOT: The Tiles Inventor Toolkit. *Advances in Intelligent Systems and Computing*, 294–305. https://doi.org/10.1007/978-3-319-75175-7\_31
- Pai, S. S. (2017). *IOT Application in Education*. IJARnD. https://www.ijarnd.com/manuscript/iot-application-in-education/
- Tengku Paris, T. N., Yussof, R. L. (2012). Enhancing grammar using board game. *Procedia Social and Behavioral Sciences*, 68, 213–221. https://doi.org/10.1016/j.sbspro.2012.12.221

# Application of ChatGPT in the Role-play Game of Modeled United Nations

Zhen-Hung TSAI \*, Ju-Ling SHIH, Geng-De HONG Graduate Institute of Network Learning Technology, National Central University, Taiwan

\*b7258ter@gmail.com

**Abstract:** When Artificial intelligence has emerged, it has changed the area of education dramatically. In this study, a game about Modeled United Nations was designed with ChatGPT as one of the players. The study analyzed how ChatGPT will affect the students' resolutions about the international issues in the group discussion and negotiation. An API to ChatGPT is used to record the students' interaction with ChatGPT, and a quantitative analysis of interactions and survey is used to collect the students' actions and feedback about the activity. The research has expected the students to learn how to use AI technology correctly and have a better understanding about the global issues. With the help of AI, the students are also expected to make better decision with AI, and have the ability of problem-solving while working as a team.

Keywords: AI Chatbot, Model United Nations (MUN), Game Development

## 1. Introduction

Artificial Intelligence (AI) has played a key role in the digital transformation nowadays. The emergence of ChatGPT not only has the revolutionary potential for Education, but also changes the way how people use AI. With its comprehensive online database and powerful Natural Language Processing model, ChatGPT can understand the various questions and provide corresponding answers for users. They can learn how to ask ChatGPT questions precisely to get the background knowledge and how to use all the information provided from ChatGPT to solve the questions.

In this study, a game about Modeled United Nations (MUN) was designed to incorporate international issues and ChatGPT that allows players to use ChatGPT as a consultant. This study was to analyze when the students face the conflict issue, how they retrieve information from ChatGPT and propose resolutions with ChatGPT's guidance and directions.

## 2. Related work:

## 2.1 ChatGPT

The rapid development of Natural Language Processing has brought lots of new applications for education. Students can interact and start a conversation with ChatGPT and artificial intelligence. They must think and explore about the questions in their own words and receive the responses from ChatGPT to personalize the information and making it their own to achieve the meaning of learning (Mohammadreza, et al., 2023). ChatGPT can be applied on the invention of educational Question Answering System. It can achieve a more intellectual system by providing corresponding answers according to the questions.

Therefore, in this study, we expect the students to view ChatGPT as their learning companion or consultant, who they can work with while facing the international issues. The

students will use ChatGPT to find the directions of solutions and help them to make the reasonable decisions.

# 2.2 Modeled United Nations

Modeled United Nations (MUN) can elevate students' learning abilities include speaking, negotiation skills, and problem-solving ability. Besides, it can help students have a better understanding on global affairs. During the process of MUN, students will play in the roles of ambassadors and discuss current international issues according to the rules and structure of the United Nations (Daniel, 2003). This study was designed for a game in the form of MUN. While setting conflicts between countries, the student will play as the delegates of each country to solve conflicts with other countries by negotiations and conferences.

# 3. Modeled United Nations Game Design

This study was designed in the form of MUN. In the game, students will be the representatives of different countries. Each country consists of three members and ChatGPT as the fourth. Every student will play as a different role in the country to discuss their country's approach to the issues presented. Conflict issues with opposed standpoints will be presented in the game so that students can have different perspectives toward the issues. During the game, the students can use ChatGPT any time to retrieve information and possible resolutions toward the issue. They would form problem-solving directions from the responses of ChatGPT and through the discussion between group members. Then, they will make declarations for the country they represent. The gaming flow of this study is as Figure 1.

Before the game starts, the rules of the game and the operation of the system will be introduced to the students. First, the players will be divided into groups. The background knowledge about each country will be given to each group. Then, the group discussion will start and students can ask for advice by consulting ChatGPT. After gaining response from ChatGPT, the students will have to digest the suggestions of ChatGPT, integrate it with their ideas, and propose resolutions for the issues. The presentation process will begin once every group has finished integrating their ideas completely. Every group will debate and negotiate about the resolutions on the issues with other groups. Resolutions will be decided during the process of voting at the end of the game. After that, the instructor will ask the students' feedback about the game and provide their own perspectives about the MUN activity.

The study will observe how the students enquire ChatGPT of the issues and what kind of the decisions they make from the responses of ChatGPT by using their interaction records. With the design of the game, the students can not only enhance their understanding about international issues, they are also expected to grow critical thinking and collaboration abilities.



Figure 1 Game flow

## 4. Research design

This study is to realize the interaction between students and ChatGPT during the process of MUN. It is designed to discuss about students' cooperation with ChatGPT and analyze how ChatGPT influence their decisions. The study will collect research data by video recordings. At the same time, by using an API to connect with ChatGPT will the researchers save the conversation records in the form of CSV files. At the end of the study, students need to fill in a quantitative survey to record the students' feeling about the activity and their thoughts

about consulting with ChatGPT. The records will be used to analyze and evaluate the interactive effect between students and ChatGPT in the MUN game.

# 5. Expected Results

This study is designed to help students learn how to discuss with AI and internalize knowledge in the response from ChatGPT to help them make decisions through the process of the MUN game. While interacting with ChatGPT, students can learn more knowledge and perspectives, and consider about the various possibilities of the issues. Playing as a teammate or consultant, ChatGPT can provide students valuable information and advices to help them have a better understanding of the issues during the game. Moreover, these learning experience and skills can help students to gain advantages in the game and have important effects for students' learning in the future

# 6. Conclusion

In this study, students are expected to learn not only how to debate on international issues and make decisions with their teammates, but also how to use AI correctly and combine advices from ChatGPT and others to propose resolutions for the issue. AI is also expected to help students cooperate with others in a more efficient way and get critical answers from ChatGPT. Through the practical participation of MUN, the students can discuss the resolutions for the international issues as a team which can promote the students' retention of information and their impression of learning.

# Acknowledgements

This study is supported in part by the National Science and Technology Council (previously known as Ministry of Science and Technology) of Taiwan, under MOST 108-2511-H-008 - 016 -MY4.

## References

Rospigliosi, P. (2023) Artificial intelligence in teaching and learning: what questions should we ask of ChatGPT?, Interactive Learning Environments, 31(1). 1-3.

https://doi.org/10.1080/10494820.2023.2180191

- Jamian, L. S., Sidhu, G. K., & Aperapar, P. S., (2013). Managerial decision styles of deans in institutions of higher learning. Procedia Social and Behavioral Sciences, 90(2013), 278-287. https://doi.org/10.1016/j.sbspro.2013.07.092
- McIntosh, D. (2001). The uses and limits of the Model United Nations in an international relations classroom. International Studies Perspectives, 2(3), 269-280. http://www.jstor.org/stable/44218169
- Taecharungroj, V. (2023). "What can ChatGPT Do?" Analyzing early reactions to the innovative AI Chatbot on twitter. Big Data and Cognitive Computing., 7(1), 35. MDPI AG. Retrieved from http://dx.doi.org/10.3390/bdcc7010035

# IOT-integrated SDGs Scenario-based Games in the Classrooms

#### Ching-Chieh LIN\*, Ju-Ling SHIH, Yu-Hao LU

Graduate Institute of Network Learning Technology, National Central University, Taiwan \*andy19940705@gmail.com

**Abstract:** The Internet of Things (IoT) technology has found widespread application in various facets of daily life, including education. Through the integration of IoT technology and authentic open data into educational materials, students can engage in more interactions and develop a sense of participation. This, in turn, enhances their learning interests and motivation. Additionally, students can gain a more lifelike and tangible learning encounter in the classroom, particularly through scenario-based gaming. This research endeavors to utilize the Sustainable Development Goals (SDGs) as content for the creation of IoT-integrated classroom games. These games aim to immerse participants in the intricate web of social and environmental conflicts. Employing an exploratory learning approach, the goal is to augment students' awareness of SDG issues and cultivate multiple representation literacies.

Keywords: IOT-integrated Games, Multiple Representations, Data Visualization

## 1. Introduction

In today's society, IoT technology is increasingly prevalent, touching various life aspects, including education. Unlike passive learning from traditional materials, IoT-integrated educational resources promise a more practical, engaging, and interactive learning experience. This study aims to create an IoT-integrated SDGs scenario game via online platforms and IoT firmware. Visualized data from diverse sources will aid seamless gameplay. Analyzing students' cognitive recognition and outcomes will form a basis for multidimensional representation analysis. IoT devices and sensors will visualize relevant data, fostering an immersive learning atmosphere. This enables active involvement in content-rich scenarios, aligning with real-world applications.

## 2. Related Work

## 2.1 Multiple Representations

"Representation" serves both as a learning medium and a contemplation tool. In the game, we analyze students' varied representations through data visualization, aligning with their preferences. This fosters comprehensive, multi-dimensional comprehension, linking new knowledge with real-life scenarios. Students find more enjoyment and satisfaction in learning. Chen, Chen, Zuo, & Hou (2023) highlight that multiple representations stimulate thinking and learning. Offering diverse representations promptly supports learning, enhancing understanding of surroundings and cognitive growth. This study employs a game format with multiple representations, visually showing students' diverse patterns of representation during gameplay.

## 2.2 IOT

In the changing educational landscape, traditional teaching falls short of reform needs. IoT and Artificial Intelligence (AI) technologies n education lay the foundation for progress (Yu & Mi, 2023). Merging education and tech drives efficient learning, improves quality. IoT

enables personalized education (Tripathy, Mishra, & Dash, 2021), offering real-world application through sensors and devices. This enhances comprehension, lowering costs, boosting efficiency. This study aims to use IoT for a context-based classroom game, amplifing students' practical experience, engagement, and digital skills. The goal is to prepare students for the digital future.

# 2.3 Data Visualization

Duke, Brodlie, Duce, & Herman (2005) propose that visualization initiates when individuals possess data they wish to explore and interpret. Visualization involves converting information from one representation form to another (Chen & Golan, 2015). Data visualization can render information more compelling (Pandey, Manivannan, Nov, Satterthwaite, & Bertini, 2014). Moreover, research has indicated a robust connection between data visualization and the humanities (Manovich, 2010). In fact, graphical representation can be more precise and impactful compared to traditional statistical calculations (Tufte, 2001). Hence, our aim is to utilize data visualization techniques to assist students in attaining a lucid comprehension of diverse data points. We intend to convert the multifaceted data representations generated by students into intuitive and understandable charts, images, or visual presentations. This approach will aid students in achieving a deeper understanding of the learning material and heighten their learning outcomes.

# 3. Game Design

This game utilizes task design, game prompts, data exploration, and technological assistance from IoT to facilitate diverse forms of communication, interaction, cooperation, and coordination among students. The game involves students taking on specific roles, including Farmer, Conservation Group, Hunter, Government, and Developer. These five members constitute a group, each role equipped with unique physical cards.

The game's cards have two main purposes. The Inquiry card lets students interact with map areas using IoT icons, exploring diverse data like water resources, land details, forest distribution, population structure, and animal ecology. Personalized data exploration strategies can be formed based on their interpretation of game missions. The Action cards are to initiate interaction with other players. It acts as the initial condition for engaging with others. Each card is equipped with a sensing device, and corresponding sensing zones are established in the game environment. When students use these cards to interact with the sensing zones, the screens they possess will promptly exhibit numerical values or dynamic scene changes. This grants students an intuitive comprehension of the broader impact of their actions within the game (see Figure 1).



#### 4. Expected Result

This study uses IoT tools to enhance students' tech experience, boost motivation, and explore diverse data representations within SDG-framed gameplay analysis, uncovering multiple literacies. Possible research methods for this study include a combination of qualitative and quantitative approaches using in-depth interviews and focus groups to gather insights into students' perceptions, experiences, and motivations within the IoT-integrated game; at the same time using surveys and gameplay analytics to provide statistical data on students' engagement patterns, learning outcomes, and problem-solving strategies. Data from these methods will be synthesized to gain a comprehensive understanding of students' behaviors and outcomes.

This research aims to elevate learning motivation through gamification and leverage data visualization for students' scientific insight. Gathering multiple representations showcases interactive engagement with data and peers in-game, comprehensively analyzing learning behaviors, problem-solving, and creativity. This contributes to personalized instruction and diverse learning, fostering self-reflection and growth. Efforts seek continuous improvement and educational optimization.

#### Acknowledgements

This study is supported in part by the National Science and Technology Council (previously known as Ministry of Science and Technology) of Taiwan, under MOST 108-2511-H-008 - 016 -MY4 and NSTC 112-2410-H-008 -010 -MY3.

#### References

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. Learning and instruction, 16(3), 183-198.
- Alsawaier, R. S. (2018). The effect of gamification on motivation and engagement. The International Journal of Information and Learning Technology, 35(1), 56-79.
- Aparicio, M., & Costa, C. J. (2015). Data visualization. Communication design quarterly review, 3(1), 7-11.
- Chen, M. F., Chen, Y. C., Zuo, P. Y., & Hou, H. T. (2023). Design and evaluation of a remote synchronous gamified mathematics teaching activity that integrates multi-representational scaffolding and a mind tool for gamified learning. Education and Information Technologies, 1-27.
- Chen, M., & Golan, A. (2015). What may visualization processes optimize?. IEEE transactions on visualization and computer graphics, 22(12), 2619-2632.
- Duke, D. J., Brodlie, K. W., Duce, D. A., & Herman, I. (2005). Do you see what I mean? [Data visualization]. IEEE Computer Graphics and Applications, 25(3), 6-9.
- Manovich, L. (2010). What is visualization?. paj: The Journal of the Initiative for Digital Humanities, Media, and Culture, 2(1).
- Pandey, A. V., Manivannan, A., Nov, O., Satterthwaite, M., & Bertini, E. (2014). The persuasive power of data visualization. IEEE transactions on visualization and computer graphics, 20(12), 2211-2220.
- Sin, K., & Muthu, L. (2015). Application of big data in education data mining and learning analytics–a literature review. ICTACT journal on soft computing, 5(4), 1035-1049.
- Tufte, E. (2001). Graphical excellence. The visual display of quantitative information, 13-52.
- Wu, A., Wang, Y., Shu, X., Moritz, D., Cui, W., Zhang, H., ... & Qu, H. (2021). Ai4vis: Survey on artificial intelligence approaches for data visualization. IEEE Transactions on Visualization and Computer Graphics.
- Yu, H., & Mi, Y. (2023). Application Model for Innovative Sports Practice Teaching in Colleges Using Internet of Things and Artificial Intelligence. Electronics, 12(4), 874.

# Pupils' perceived immersion, attitudes, and learning effectiveness in virtual field trips: A comparison between immersive and projective environments

## Kun-Hung CHENG<sup>ab\*</sup>

<sup>a</sup>Graduate Institute of Library and Information Science, National Chung Hsing University, Taiwan <sup>b</sup>Innovation and Development Center of Sustainable Agriculture, National Chung Hsing

University, Taiwan \* khcheng@dragon.nchu.edu.tw

**Abstract:** Focusing on the implementation of virtual field trips in elementary school classroom, this study aimed to explore pupils' affective perceptions and cognitive performance in either immersive environments (using VR headsets, n=26) or projective environments (using traditional projection screen, n=26) through a quasi-experimental design. The results showed that the pupils exhibited stronger perceived immersion and positive attitudes toward virtual learning and achieved better learning effectiveness in the immersive setting than in the projective setting. We are continuing to expand this work for more understandings of immersive virtual field trips in elementary classrooms.

Keywords: Virtual reality, field trip, science learning, immersion

# 1. Introduction

Virtual field trips are one of the educational applications of virtual reality (VR) technology for instructors to lead learners to explore learning content in virtual environments (Caliskan. 2011). Several studies have examined the pedagogical possibility of immersive virtual field trips in classrooms (Cheng & Tsai, 2019; Han, 2020). Notably, some studies reported that learning science by immersive VR may result in attentional dispersion and cognitive overload for learners (Makransky et al., 2019). Compared with learning science by immersive VR, students showed less affective perceptions (e.g., motivation, interest, and engagement) but performed better when using traditional medium such as slideshows (Parong & Mayer, 2018). Similar findings were also documented in Parong and Mayer's study (2021). However, these studies were conducted in the context of research lab rather than in the practical fields such as classrooms. Therefore, focusing on the implementation of virtual field trips in elementary school classroom, this work aimed to explore students' affective perceptions and cognitive performance in either immersive environments (using VR headsets) or projective environments (using traditional projection screen). Specifically, a series of comparison of pupils' perceived immersion, attitudes, and learning effectiveness in immersive and projective virtual field trips will be implemented in this study.

# 2. Method

## 2.1 Participants

There were 52 elementary school students from two classes in fifth grade (26 students for each class) invited to participate in this study. The gender distribution of the pupils was balanced, that is, there were 13 males (50%) and 13 females (50%) for each class. About half of the pupils had had experiences in the usage of VR-related applications, indicating that the

novelty effects of virtual technology the participants perceived may not interfere the results of the current study.

# 2.2 Experimental design and procedure

A quasi-experimental study was conducted to compare the pupils' perceived immersion, attitudes, and learning effectiveness in immersive virtual field trips with those in projective virtual field trips. While the pupils in the experimental group (n=26) were involved in the virtual learning activities by individually wearing head-mounted displays (HMDs, plastic cardboard used in this study) for freely interacting with the virtual learning elements and navigating the virtual scenes, the pupils in the control group (n=26) engaged in the virtual learning activities by passively observing the virtual learning elements and scenes through a projective screen in front of a classroom. The pupils in the two research groups received the same science learning content of solar energy from a VR application for implementing virtual field trips. Two teachers with experiences in using the VR application were invited to implement the learning activity in the experimental and control groups, respectively. They reached a consensus in exploiting the same learning content and instructional strategies. In other words, the possible extraneous variables were controlled in this study.

The learning activities involves three science lessons (approximately 40 minutes for each lesson). While the first two lessons were instructed in a traditional way, the third lesson was instructed by virtual field trips. Before the virtual learning activity began, the pupils were required to complete a pre-test for assessing their prior knowledge in terms of solar energy. When the virtual field trips finished, the pupils had to complete a post achievement test and respond to two surveys. One instrument was adapted from Jennett et al.'s *Immersive Experience Questionnaire* (IEQ) (2008) including the subscales of (1) *basic attention*, (2) *temporal dissociation*, (3) *transportation*, (4) *emotional involvement* and (5) *enjoyment*. The other one was adapted from the questionnaire including the subscales of (1) *usefulness*, (2) *visual appeal*, (3) *satisfaction*, and (4) *intention* used in Cheng et al.'s study (2019).

## 3. Results

## 3.1 Comparison of the pupils' perceived immersion

To understand the pupils' perceptions of immersion in the virtual learning activities, a series of independent t-test was conducted to examine whether they perceived immersion in the immersive and projective virtual field trips in different degrees. According to Table 1, the pupils' perceived basic attention (t=3.43, p<0.01), transportation (t=2.04, p<0.05), emotional involvement (t=2.57, p<0.05), and enjoyment (t=3.37, p<0.01) in the immersive setting were stronger than the pupils' perceptions in the projective setting with large effect sizes.

	Immersive VR Mean (SD)	Projective VR Mean (SD)	t-value	Effect size
Basic attention	4.53 (0.60)	3.74 (1.01)	3.43**	0.95
Temporal dissociation	3.26 (1.00)	3.09 (1.15)	0.57	0.16
Transportation	3.86 (0.95)	3.27 (1.13)	2.04*	0.57
Emotional involvement	4.05 (0.76)	3.45 (0.91)	2.57*	0.72
Enjoyment	4.31 (0.69)	3.54 (0.93)	3.37**	0.94

Table 1.	Comparison	of students'	' perceived	immersion	between	the two	groups
----------	------------	--------------	-------------	-----------	---------	---------	--------

\**p* < .05, \*\**p* < .01

## 3.2 Comparison of the pupils' attitudes

This study conducted a series of independent t-test to examine the differences of the pupils' attitudes towards the virtual field trips when engaging in immersive and projective environments. The results in Table 2 show that, with large effect sizes, the pupils in the

immersive learning environments exhibited more positive attitudes towards virtual field trips than the pupils in the projective learning environments did to a significant level. Specifically, the pupils considered that the immersive virtual field trips were usefulness for their science learning (t=3.18, p<0.01) and could appeal their visual attention (t=3.19, p<0.01). Also, they were satisfied with the virtual learning experiences (t=2.91, p<0.01) and were inclined to learn science by immersive virtual technology (t=2.92, p<0.01). In other words, compared with the projective setting in the classroom, the immersive virtual field trips may benefit the pupils' attitudes more.

	Immersive VR Mean (SD)	Projective VR Mean (SD)	t-value	Effect size
Usefulness	4.53 (0.73)	3.84 (0.84)	3.18**	0.88
Visual appeal	4.46 (0.75)	3.65 (1.05)	3.19**	0.89
Satisfaction	4.50 (0.74)	3.80 (0.98)	2.91**	0.81
Intention	4.56 (0.63)	3.92 (0.93)	2.92**	0.81
*				

	Table 2. Cor	nparison of	pupils' a	attitudes	between	the tw	o groups
--	--------------	-------------	-----------	-----------	---------	--------	----------

\**p* < .05, \*\**p* < .01

#### 3.3 Comparison of the pupils' learning effectiveness

To examine whether the pupils in the two experimental groups academically performed in a different level, this study conducted an analysis of covariance (ANCOVA) with the pre-test score as the covariance. The significant difference in the pupils' learning effectiveness between the immersive and projective VR settings was found (F=5.40, p<0.05). Specifically, the pupils engaging in the immersive virtual field trips performed (adjusted mean=26.00, standard error=0.89) better than those learned science by observing virtual learning elements through a projection screen (adjusted mean=22.89, standard error=0.89).

## 4. Conclusion

Similar to the findings of past studies regarding learners' affective perceptions in immersive VR environments (Makransky et al., 2019; Parong & Mayer, 2018), this study did verify that the pupils exhibited stronger perceived immersion and positive attitudes toward virtual learning in the immersive setting than in the projective setting. However, in contrast to those studies (Makransky et al., 2019; Parong & Mayer, 2018), the pupils in this study were cognitively benefited more by actively exploring in the immersive learning environments than by passively observing the virtual learning content (e.g., panoramic scenes) through traditional projection screen. To understand what factors may influence the pupils' cognitive performance during immersive virtual field trips, we are continuing to collect more research data in elementary classrooms for the examination of structural relationships among the affective (e.g., curiosity or motivation) and cognitive (e.g., conceptual knowledge or transfer of learning) variables. Qualitative data such as teachers' perceptions of immersive virtual field trips in classrooms was also our attempt of virtual learning research in the future.

# **References (selected)**

- Cheng, K. H.\*, & Tsai, C. C. (2019). A case study of immersive virtual field trips in an elementary classroom: Students' learning experience and teacher-student interaction behaviors. *Computers & Education, 140*, 103600.
- Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225-236.
- Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785-797.
- Parong, J., & Mayer, R. E. (2021). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning*, *37*(1), 226-241.

# Ethical and Privacy Concerns in Artificial Intelligence Dialogue Systems: Do Students in Higher Education Really Care About Them?

Chunpeng ZHAI<sup>\*</sup>, Santoso WIBOWO & Lily D. LI School of Engineering and Technology, CQUniversity, Australia \*c.zhai@cqu.edu.au

Abstract: This work-in-progress paper investigates on whether students in higher education really care about the ethical and privacy concerns in the use of artificial intelligence (AI) dialogue systems. Based on semi-structured interviews of 37 undergraduate students, this study uncovered a variety of viewpoints. The study found that despite their awareness of ethical and privacy concerns in the use of AI dialogue systems, the majority of students did not put much effort in addressing these issues. Around 75% of the students prioritized efficient and effective university assignment completion with the help of Al dialogue systems over these ethical and privacy concerns. The study also revealed that cultural differences influenced students' attitudes towards ethical and privacy concerns. Specifically, students from high-context cultures, where a significant portion of communication relies on shared experiences, implicit cues, and non-verbal elements, showed less concern compared to those from low-context cultures, where their communication relies heavily on explicit messages, direct language, and precise information. This study recommends several strategies to address privacy and ethical concerns, including the development of comprehensive data management policies, the use of data anonymization and minimization techniques, and the enforcement of stringent security measures to prevent data breaches.

Keywords: Dialogue systems; higher education; ethical concerns; privacy, students

## 1. Introduction

Artificial intelligence (AI) dialogue systems, also known as conversational agents or chatbots, are software mechanisms that mimic human dialogue. These systems have been integrated with various sectors, such as e-commerce, healthcare, finance, and notably, education (Adamopoulou & Moussiades, 2020). Their use in academia has recently gained significant attention, promising to transform pedagogical practices and student learning (Zhai & Wibowo, 2023). Al dialogue systems offer various advantages, such as automating tasks, aiding in data analysis, and fostering innovative learning and evaluation methods (Song & Xiong, 2021). This allows for more effective learning, empowering students to better understand complex and creative information (Zhai & Wibowo, 2023). Additionally, these systems can simplify online learning, extending educational opportunities to students in remote locations, thus promoting a more equitable and inclusive educational landscape (Bhat et al., 2023). Despite these benefits, there are significant concerns associated with using AI dialogue systems in academia, primarily ethical issues and privacy concerns (Song & Xiong, 2021). Given that AI systems can potentially introduce bias and discrimination in educational settings, and even with safeguards like blockchain technology and end-to-end encryption, users might still face threats like malicious input, profiling, contextual attacks, and data breaches (Adamopoulou & Moussiades, 2020). This underscores the importance of conducting in-depth research into the ethical and privacy concerns of AI dialogue systems in education (Zhai & Wibowo, 2023). In this work-in-progress paper, a case study was conducted to better understand these privacy

and ethical concerns and outline potential strategies for dealing with these concerns. This study attempts to answer the following research questions:

- 1) What are the privacy and ethical concerns associated with the use of AI dialogue systems in higher education?
- 2) Do students care and act accordingly to ethical and privacy concerns in the use of Al dialogue systems in higher education?

#### 2. METHOD

This study aims to better understand the privacy and ethical concerns in the use of MACHE-Bot, an Al dialogue system developed by Zhai and Wibowo (2023), for facilitating students in higher education with personalized, just-in-time feedback or assistance in their learning process. The acronyms of MACHE-Bot stand for multidimensional approach encompassing culture, humor, and empathy, which is culturally sensitive dialogue system, uniquely designed to engage users through humor and empathetic interactions. A total of 37 university students from various ages, genders, and cultural backgrounds were selected from a university in Queensland, Australia. The students interacted with the MACHE-Bot for a period of two weeks, which was followed by a semi-structured interview. These interviews were used to gather insights on students' experiences, attitudes, and understanding of the system, aspects which are hard to capture using quantitative methods. The participants' feedback was subsequently analyzed via NVivo Transcription and a three-tiered coding system, a structured approach to analyzing qualitative data, revealing key themes about their experiences and perceptions, particularly towards the ethical and privacy concerns related to the use of Al dialogue systems.

## 3. PRELIMINARY FINDINGS AND DISCUSSIONS

1) What are the privacy and ethical concerns associated with the use of AI dialogue systems in higher education?

Privacy concerns associated with AI dialogue systems in the education sector are crucial for safeguarding the integrity and confidentiality of data. Based on the study, these concerns include data privacy and ethical implications. Data privacy is a significant concern as AI dialogue systems often require access to personal information and conversations. This raises questions about how data is stored, used, and protected. Unauthorized access or data breaches could result in privacy violations and potential misuse of sensitive educational information (Song & Xiong, 2021). Additionally, biases and discrimination pose concerns in AI dialogue systems. These systems can inherit biases from training data, potentially resulting in discriminatory or unfair outcomes. In the education sector, where equal opportunities and inclusivity are paramount, higher education institutions must address biases and ensure fairness. Institutions must establish protocols to ensure proper data handling, secure storage, and compliance with privacy regulations to protect student and user data (Adamopoulou & Moussiades, 2020). Ethical implications of data usage, participants believe that higher education institutions should consider the ethical implications of collecting and analyzing data through AI dialogue systems. Transparency in data collection practices, ensuring informed consent from participants, and responsible handling of data are crucial elements to ensure ethical practices (Song & Xiong, 2021). Figure 1 illustrates the variety of ethical and privacy issues related to the use of AI dialogue systems in higher education.

2) Do students care and act accordingly to ethical and privacy concerns in the use of Al dialogue systems in higher education?

The research findings show that while students demonstrated an awareness of these concerns, most of them did not show significant action to address these issues. Around 75% of the students indicated that their primary objective was to complete assignments effectively and efficiently, regardless of ethical and privacy issues surrounding AI usage. Participant 11

states that "we know the ethical issues, but honestly, I never really thought about how I could actively avoid ethical problems in my study, I mean, my university assignment. Completing assignments effectively is important to me." Furthermore, there was an interesting cultural distinction, with students from high-context cultures appearing to be less concerned about these issues compared to those from low-context cultures. Participants perceived these issues as secondary to academic success and considered paraphrasing and citation to be sufficient to avoid any potential ethical infringements (Song & Xiong, 2021). For example, participant 8 stated that " I am from China, in our culture, getting high marks and study in good university, is the only option, we rarely think about ethical or privacy issues." And Participant 30 said that" in India, we have rote learning, and compete each other in study to get to the top position, using others' data and ideas that is normal in our culture." This might be attributed to the emphasis these cultures place on achievement and competition, leading students to sideline issues they perceive as less directly linked to their academic success (Zhai & Wibowo, 2023).



FIGURE 1. Ethical and privacy concerns associated with the application of AI dialogue systems in higher education

## 4. EXPECTED CONTRIBUTION AND FUTURE WORK

This work-in-progress paper examines on whether students in higher education really care about ethical and privacy concerns in the use of AI dialogue systems. A total of 37 undergraduate students was interviewed on the ethical and privacy concerns in the use of AI dialogue systems. The study found that the majority of students are aware of ethical and privacy concerns in the use of AI dialogue systems. However, they did not put much effort in addressing these issues. To fully leverage AI's potential in education and research, it is critical for educators and researchers to scrutinize the ethical and privacy consequences of AI dialogue systems, promoting their responsible and transparent usage. To expand on these findings and gain a more comprehensive understanding, further research will be conducted using a mixed-method approach which will enable the incorporation of both quantitative and qualitative data collection methods, providing a more holistic view of the effectiveness and implications of AI dialogue systems in education.

#### References

- Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. IFIP international conference on artificial intelligence applications and innovations,
- Bhat, J. R., AlQahtani, S. A., & Nekovee, M. (2023). FinTech enablers, use cases, and role of future internet of things. *Journal of King Saud University-Computer and Information Sciences*, 35(1), 87-101. <u>https://doi.org/10.1016/j.jksuci.2022.08.033</u>
- Song, X., & Xiong, T. (2021). A survey of published literature on conversational artificial intelligence. 2021 7th International Conference on Information Management (ICIM),
- Zhai, C., & Wibowo, S. (2023). A WGAN-based dialogue system for embedding humor, empathy, and cultural aspects in education. *IEEE access*. <u>https://doi.org/10.1109/ACCESS.2023.3294966</u>

# Resistance Training Support System with Pose Estimation

#### Koki YAMADA<sup>a</sup>, Naka GOTODA<sup>b\*</sup> & Ryota AKAGI<sup>a,c\*</sup>

<sup>a</sup>Graduate School of Engineering and Science, Shibaura Institute of Technology, Japan
<sup>b</sup>Integrated Center for Information, Kagawa University, Japan
<sup>c</sup>College of Systems Engineering and Science, Shibaura Institute of Technology, Japan
\*gotoda.naka@kagawa-u.ac.jp(N.G.), rakagi12@shibaura-it.ac.jp(R.A.)

Abstract: Squat training is highly effective for improving lower limb muscle function. However, when this type of training is performed alone without a personal trainer or physical therapist, it can lead to inefficiencies and injuries among trainees. Thus, we aim to develop a system that provides visual and real-time feedback on the correct posture of a lone squat trainee using pose estimation. In this study, we propose a function that demonstrates the correct posture for each squatting discipline in the form of a line-segment posture representation and warns the trainee when he/she assumes an incorrect posture. Squatting motion in the sagittal plane was captured using a camera connected to a personal computer, and the coordinates of the acromion, hip, knee, and ankle joints were detected using MoveNet. The joint angles were calculated from the detected coordinates and the correct posture was determined according to the individual's body shape. The color of the formed line-segment posture representation changed to indicate the difference between the trainees' actual and correct postures. In future research, we plan to assess the displacement of key points detected by MoveNet using optical technologies to confirm the reliability of this system's pose estimation. Afterwards, we will evaluate the effectiveness of this system in determining whether a trainee can move appropriately.

**Keywords:** squat training, pose estimation, MoveNet, visualization support, awareness support

#### 1. Introduction

Resistance training mainly focuses on improving muscle activity and is highly effective in improving muscle strength, power, and hypertrophy (Erskine et al., 2010). Resistance training is used not only in sports but also for health maintenance and rehabilitation, thus making it popular among athletes and the general population.

For effective resistance training results, they must be performed under appropriate program designs, proper progression, and motion to meet the objectives of the people undergoing resistance training. Having a personal trainer or other instructors with expertise in resistance training can make it more effective (Ratamess et al., 2008). However, it is also true that many people who perform resistance training are unable to have an instructor because of financial constraints and geographical issue (because they live in rural areas). It is difficult for them to assess objectively their posture while performing resistance training alone. Additionally, resistance training performed by adopting an incorrect posture can cause injuries such as tendon rupture and ligament damage (Bengtsson et al., 2018). Therefore, trainees must have a method that can address the problem associated with the lack of instructors.

Recently, several applications utilizing posture estimation were studied (Stenum et al., 2021). Pose estimation can detect "key points," such as the eyes, nose, and joints from video images captured by cameras by using machine learning. To the best of our knowledge, there have been no previous studies on the evaluation of posture during squat training, which helps improve both the flexion and extension torque of the knee and hip

(Akagi et al., 2020). This study proposes a system that visually teaches the correct posture to young people performing squat training alone who are adept with the use of basic personal computers (PCs) to perform squat training safely.

# 2. Functional requirements and quantitative body motion estimation

To train people undergoing squat training alone in the correct posture, it is essential to provide them with an environment wherein they can visually assess both their own and the correct postures in real-time. In addition, it is necessary to determine whether the squatting motion of the participants is correct. Moreover, because posture is determined by the joint angles involved in the squatting motion, these joint angles must be calculated as required. Therefore, the functional requirements of the squat-training support system are as follows,

- A) Assessment the body motion in real-time
- B) Visual instruction of the correct posture
- C) Calculation of the joint angles

To satisfy B) and C) during squat training, it is necessary to capture body motions. Therefore, we focused on the pose estimation library "TensorFlow.js" which can detect feature points of the body using machine learning from video images captured from a single common camera.TensorFlow.js can complete the inference process and render the processing results on the front-end side, thus enabling a faster response time. Hence, we employed MoveNet, a pretrained model in TensorFlow.js, to quantitatively capture motion from the perspective of A) real-time performance. MoveNet can detect 17 key points in the body. The human skeleton was evaluated based on these key points.



*Figure 1.* Seventeen key points defined on MoveNet and set of two vectors at each key points to calculate joint angles

## 3. Visual teaching of correct posture

In this system, a camera (connected to PC) records video images of the trainee from the side while he or she is training at 30 fps. A display connected to a PC was placed in front of the trainee so that he/she could check his/her posture and the content of the instructions in real-time.

The position X-Y coordinates of the trainees' shoulders, hips, knees, and ankle joints were obtained using MoveNet. First, the system automatically calculates the length between the shoulder and hip joints, thigh length, and lower leg length when trainee is in the standing position. In this case, each length is calculated using two vectors a, b, as shown in Figure 1 and formula (1) (6: shoulder, 12: hip, 14: knee, and 16: ankle joints are included in *i* and *j*). As shown in Figure 3, the ideal line-segment representations between the shoulder and hip joint, hip joint and knee joint, and knee joint and ankle joint in the squat motion are displayed by p5.js that is JavaScript library for draw in browser pages (B)), considering the trainee's

calculated length between the shoulder and hip joint, thigh length and lower leg length. The trainees performed the squatting motion to match the presented ideal line-segment representation. Each joint angle ( $\theta$ ) during the movement was calculated sequentially using the inner product as formula (2) (C)). If the trainee's posture matched the ideal line-segment representation, the color of the ideal line-segment representation changed to red to allow the trainee to know that he/she assumed the correct posture.

$$|\mathbf{a}|, |\mathbf{b}| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$
(1)

$$\theta = \cos^{-1} \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|} \tag{2}$$



#### 4. Future development

In future research, we plan to conduct an error verification test to assess the deviation between the position coordinates of the key points detected by MoveNet and the original joint position coordinates of the body using motion information captured using optical technologies to confirm the reliability of the system's posture estimation. Subsequently, we plan to evaluate the effectiveness of this system in determining whether a trainee can move appropriately according to the ideal line-segment posture representations.

#### Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP23K02735 (to NG). We would like to thank Editage (www.editage.com) for the English language editing.

#### References

- Akagi, R., Sato, S., Hirata, N., Imaizumi, N., Tanimoto, H., Ando, R., Ema, R., & Hirata, K. (2020). Eight-Week Low-Intensity Squat Training at Slow Speed Simultaneously Improves Knee and Hip Flexion and Extension Strength. *Frontiers in physiology*, *11*, 893.
- Bengtsson, V., Berglund, L., & Aasa, U. (2018). Narrative review of injuries in powerlifting with special reference to their association to the squat, bench press and deadlift. *BMJ open sport* & *exercise medicine*, *4*(1), e000382.
- Erskine, R. M., Jones, D. A., Williams, A. G., Stewart, C. E., & Degens, H. (2010). Inter-individual variability in the adaptation of human muscle specific tension to progressive resistance training. *European journal of applied physiology*, *110*(6), 1117–1125.
- Ratamess, N. A., Faigenbaum, A. D., Hoffman, J. R., & Kang, J. (2008). Self-selected resistance training intensity in healthy women: the influence of a personal trainer. *Journal of strength and conditioning research*, 22(1), 103–111.
- Stenum, J., Cherry-Allen, K. M., Pyles, C. O., Reetzke, R. D., Vignos, M. F., & Roemmich, R. T. (2021). Applications of Pose Estimation in Human Health and Performance across the Lifespan. Sensors, 21(21), 7315.

# Evaluation of Simulators to Promote the Understanding of Bioaccumulation among Elementary School Students

Yuka MATSUYAMA<sup>a</sup>, Shinichi KAMIYAMA<sup>b</sup>, Hideo FUNAOI<sup>c</sup> & Tomokazu YAMAMOTO<sup>a\*</sup>

<sup>a</sup>Graduate School of Education, Hyogo University of Teacher Education, Japan <sup>b</sup>Osaka University of Health and Sport Sciences, Japan <sup>c</sup>Soka University, Japan \*ttyamamo@hyogo-u.ac.jp

**Abstract:** This study introduced four simulators developed by Matsuyama et al. (2023) into sixth-grade science classes to educate students about bioaccumulation of pufferfish toxin. The simulators helped students understand how the concentration of pufferfish toxin increases as predators approach the top of the food chain. Results of a survey of "students' usability" indicated that all 12 of the questions received a positive rating. In "students' understanding," the number of creatures that students thought would be affected by DDT (dichlorodiphenyltrichloroethane) bioaccumulation increased in part pre- and post- transfer test. Particularly, the answers of aquatic creatures showed a significant improvement, and it was found that the simulators helped students understand "the food chain and bioaccumulation in the water."

Keywords: Computer Simulation, Scratch, Food Chain, Bioaccumulation

#### 1. Introduction

Bioaccumulation, which is closely related to food issues, is a problem that should be solved by using the knowledge of the food chain learned in elementary school. One bioaccumulation example is pufferfish toxin which is unnoticeable, complex, pervasive, and long-term phenomenon. Thus, directly observing or capturing its entirety is difficult (Itoi et al., 2015). Nonetheless, simulation supports students understanding by simplifying the events and adjusting scale and period of time. (Chang et al., 2019; Shingai et al., 2020). Furthermore, incorporating game elements and scaffolding into the simulation encourages students' deep exploration (Kukkonen et al., 2014). However, no simulators have been developed in Japanese elementary schools to learn the food chain. This study introduced the simulators for elementary school students to examine the bioaccumulation of pufferfish toxin. The simulator program comprises "Scratch," which is a teaching tool that can easily be used in various classroom settings, and they were loaded and run locally.

This study aimed to clarify the effectiveness of the simulators for learning bioaccumulation by introducing them to sixth-grade science classes and surveying students' usability and understanding of the simulators.

## 2. Features of the Simulators and Their Use

Four types of simulators (A, B, C, D) were introduced in the science class "connections between creatures" for 43 sixth-grade students. The simulator-based class comprised two hours wherein students are educated about the food chain and bioaccumulation in five creatures, including pufferfish, crabs, starfish, clams, and marine bacteria.

Figure 1 presents the screens of Simulators. Simulator A in Fig. 1 (left) helped the students understand the relationship between the puffer fish as a top predator and the food

chain. Clicking on a button with the creature's name allows the student to move it freely using arrow keys, enabling students to examine the "eat–eaten" relationship among the five creatures. Thereafter, using Simulators B and C, the students learned that all the toxins possessed by the prey were absorbed by the predator. Simulator D in Fig. 1 (right) helped the students understand that "the accumulation of poison depends on what and how much the predator preys on, and that the amount of toxins concentration increases as the creature approaches the top of the food chain" with classifying the amount of toxins by creature type.



Figure 1. The Screens of Simulators A (left) and D (right)

# 3. Analysis

The analysis included 43 sixth-grade students. The survey on "students' usability" was conducted with a total of 12 questions, which were answered using a four-point scale: 1 for "not applicable," 2 for "somewhat not applicable," 3 for "somewhat agree," and 4 for "agree." For all questionnaires, the mean and standard deviation were calculated.

In the test concerning "students' understanding," transfer tests were performed. The students were asked to answer nine items, such as lake water, underwater, and terrestrial creatures, assuming the range of DDT's exposure influence. Between the pre- and post-test, the total number of responses for each student was analyzed by Wilcoxon's signed rank test, whereas the number of students in each answer item was analyzed by the McNemar test.

# 4. Results and Discussion

Table 1 presents the results of "students' usability." For all questions, the average score was 3 points or more (more than a majority of positive opinions). This revealed that all four types of simulators were easy for students to use.

Table 1. Contents, Mean, and Standard Deviation of the Questionnaire on "students' u	usability"
--	------------

	Questions	Mean	S.D.
(1)	The simulator was easy to use.	3.79	0.41
(2)	When the creatures in the simulator ate other creatures, a sound	3.44	0.73
	reverberated, making it easy to identify which creatures to pay attention to.		
(3)	When the creatures in the simulator ate other creatures, the amount of	3.84	0.43
	poison was displayed, making it easy to identify which creatures I must pay		
	attention to.		
(4)	When the creatures in the simulator ate other creatures, the color of the	3.63	0.72
( )	creatures changed, making it easy to identify which creatures to notice.		
(5)	By moving the creatures in the simulator freely, I was able to confirm the	3.53	0.66
( )	mechanism of the food chain while learning independently.		
(6)	By using the simulator, I learned how invisible things, such as poison in	3.77	0.42
( )	living creatures, are transferred.		
(7)	By using the simulator, I was able to consider the circumstance in its entirety	3.58	0.58
. ,	and gain a broad perspective.		

(8) By using the simulator, I was able to consider what was occurring for an	3.37	0.72
extended period of time while capturing a series of situations.		
(9) By using the simulator, I was able to observe from various perspectives	3.47	0.58
what is happening in difficult-to-observe environments, such as in the sea.		
(10) While operating the simulator, I learned the preys and predators of all five	3.93	0.25
creatures (pufferfish, crab, starfish, clam, and marine bacterium).		
(11) While operating the simulator, I learned how pufferfish poison accumulates	3.72	0.50
in its body from my knowledge of the food chain mechanism.		
(12) By operating the simulator, I learned that poison is transferred from the	3.74	0.44
body of the being eaten-creature to the body of the eating-creature through		
the "eat-eaten" relationship.		

Before and after learning, the overall number of responses that the students believed included DDT tended to increase (Z=-1.849, p<.10). Figure 2 exhibits the number of students who were able to answer all the nine items. The findings demonstrate that the number of students who could answer only for lake shellfish (p<.01) and lake plankton (very small creatures) (p<.05) significantly increased. Thus, it can be inferred that the "food chain and bioaccumulation in water" expressed in the simulators improved students' understand of bioaccumulation.

In the future, we plan to analyze children's decision making using their knowledge of bioaccumulation to evaluate the effectiveness of the simulator. In addition, differences in effectiveness between using and not using the simulator need to be examined.



Figure 2. Change in the number of students who responded through learning

# Acknowledgements

This study was supported by JSPS KAKENHI Grant Number JP23H01022.

## References

- Chang, H. Y., Wen, C. T., Chang, M. H., & Liu, C. C. (2019). Pedagogy designs to augment the impact of computer simulations. In M. Chang, H-J. So, L-H, Wong, J-L. Shih, F-Y. Yu. (Eds), 27th International Conference on Computers in Education, Conference Proceedings Volume 2. (pp.58-60). APSCE.
- Itoi, S., Kozaki, A., Komori, K., Tsunashima, T., Noguchi, S., Kawane, M., & Sugita, H. (2015). Toxic Takifugu pardalis eggs found in Takifugu niphobles gut:Implications for TTX accumulation in the pufferfish. *Toxicon, 108*, 141-146.
- Kukkonen, J. E., Kärkkäinen, S., Dillon, P., & Keinonen, T.(2014). The effects of scaffolded simulationbased inquiry learning on fifth-graders' representations of the greenhouse effect. *International Journal of Science Education*, 36(3), 406-424.
- Matsuyama, Y., Kamiyama, S., Funaoi, H. & Yamamoto, T. (2023, September 23-24). Development of the simulators for learning "the pufferfish-toxin bioaccumulation" using Scratch: Application in the sixth grade "Connections among Creatures". [Paper presentation]. The 73rd annual SJST conference 2023, Kochi, Japan.
- Shingai, Y., Aoki, R., Takeda, Y., Kusunoki, F., Mizoguchi, H., Sugimoto, M., Funaoi, H., Yamaguchi, E., & Inagaki, S. (2020). Learning support game system for sustainable development goals education: effects of improved user interface on Satoyama forest management learning. *Proceedings of the 12th International Conference on Computer Supported Education, 1*, 428-435.

# App-Infused Preschoolers' Storyline Concept-Driven Numerical Curriculum Design

#### Ruei-Cheng YEN<sup>a\*</sup> & Ben CHANG<sup>a</sup>

<sup>a</sup> Graduate Institute of Learning and Instruction, National Central University, Taiwan \*110127001@cc.ncu.edu.tw

**Abstract:** For preschoolers, Apps can facilitate them to learn mathematical concepts, such as number recognition, counting, comparison, mixture problem, combination, etc. However, most of the mathematic Apps currently are drill-and-practice design which provides numerous questions and answers matching exercises for preschoolers, that leads them paying attention on doing exercises instead of concepts acquisition. We will argue that for the preschoolers, mathematical concept acquisition is more important than drill-and-practice exercises. Therefore, to response, for three to five-year-old preschoolers, we developed a mathematical concept learning App named Kid+ which aims to build up their conceptual understanding and cognitive ability. By integrating storyline design with situated learning, the preschoolers can use the App in a concept-embedded role-play story. As three preliminary preschool experts' using feedbacks, the result revealed that the App might help the preschoolers navigating through the storyline to learn the mathematical concepts embedded. However, further evaluation of preschoolers is needed.

Keywords: preschool, mathematics, gamification, conceptual awareness

#### 1. Introduction

For preschoolers, even a two-year-old toddler can show some math abilities (Harris, 2017), but for them, mathematical learning is not just a combination of complex formulas. They may have a hard time gaining math skills, but with their parents' guidance, the more support they receive, the more learning achievement they can earn (Skwarchuk, 2009). Educational Apps are choices for the preschoolers to learn math, but most of them are not instructive enough. Only few Apps provide feedback about user's learning effectiveness, while others use the drill-and-practice strategy and only give back answers, which is relatively cheaper and easier to develop (Callaghan, 2018). Using the correct method to build study habits and a learning environment is highly recommended, even though it may not be immediately effective but will be beneficial in the future (Anders, 2012).

Recently, National Council of Teachers of Mathematics (NCTM) has categorized preschool mathematics into five standards, namely "Numbers and Operation", "Algebra", "Geometry", "Measurement", and "Data Analysis and Probability" (NCTM, 2000). At the same time, preschoolers can be grouped by ages, and their math skills would show a significant gap between each age range due to their body and brain functionality (Chang, 2015). For example, toddlers under three years old can only recognize the difference in quantity between two distinct groups but can not tell the amount. A three-year-old child can count with his fingers but can't fully understand the meaning of numbers. Children aged four years old can easily read the number of a small amounts of objects without counting, and older children will not have difficulty calculating with small numbers (Harris, 2017). In addition, Chang (2003) and Chang (2009) detailed the abilities of children between three and five years old, and records have shown that they are capable of higher-level skills involving numbering or ordinal numbers, which are proportional to age. This means that younger children may be able to count successfully but may not understand the meaning of orders or comparisons. In addition, there are principles listed for instructors to follow while students delve into math, in order to ensure

that every student has the right to education, which includes fair treatment, continuous courses that suit students' concentration, support and challenges according to their demands, assessments at proper intervals, and most importantly the use of technology to increasing students' learning capacity (Cesarone, 2008), including number sense.

Number sense is not just the ability to calculate, but the skill of converting numbers into concepts. Numbers in different positions have various definitions, including switching and adjusting numbers. "Friendliness with number" preceded number sense. It's the sense of flexibly determining whether to comprehend or compute when one sees a number (Howden, 1989). In the meanwhile, other studies define number sense as a feeling of numbers. Wilson (2009) and Dehaene (2011) claimed that humans who have acquired number sense can roughly measure amounts by simply looking at a chunk and comparing them to another at the same time. People are commonly curious about how to teach number sense. Number sense must be built on a student's own knowledge and skills. Advanced knowledge and courses based on the student's natural learning development. Afterward, the fluency in calculating should be increased and concept understanding should be strengthened. Then opportunities for exploring, problem-solving, and communicating should be given. Finally, the major way that math presents itself should be introduced to the student (Griffin, 2014).

Therefore, the goal of this research is to build up a concept-driven math learning App which has a situated storyline where the preschoolers can explore the story on tablet. To design the App, a CEO-Awareness model which represents Concept learning, Experiencing, Objects & artifacts, and Awareness was applied. The App design and the preliminary results are elaborated.

## 2. Method

## 2.1 The Concept-Experience-Object Awareness Model

Artifacts such as digital aids occupy an important position in technology education. They allow users to use, create, and understand while looking for solutions and explanations of a subject, in order to gain experience during the progress (Sundqvist, 2016). On the artifacts, the learners can learn, apply, and achieve mathematical objectives serving as activities, contents, strategies, types, purposes, tasks, resources, and principles (Bell, 1993). To facilitate the mathematical learning design, a Concept-Experience-Object Awareness (CEO-Awareness) model was proposed. Figure 1 illustrates the CEO-Awareness framework in which there are four components including Concept Learning, Experiencing, Objects & Artifacts, and Awareness.



Figure 1. The Concept-Experience-Object Awareness (CEO-Awareness) model.

Technology development today makes children's cognitive skills better. The earlier the students experience multimedia devices, the more opportunities they get to think and explore. Preschoolers' math abilities with an abstract mindset would get enhancement only if the mobile learning game was well-designed (Kokkalia, 2016), and they could feel awareness by experiencing objects or artifacts afterward. As shown in Figure 1, to help the preschoolers learn concept-driven numerical curriculum systematically, an App named Kid+ is designed based on the concept of the CEO-Awareness model, which was implemented as Figure 2. In Figure 2, the "Concept" is displayed by dialogue, and several "Objects" will be displayed to assist the preschoolers to "Experience" the concept by manipulating the "Objects." Through

the activity on the App, we hope the preschoolers can aware the mathematical concepts embedded in the storyline by him/herself.



Figure 2. The CEO-Awareness model within the game.

## 2.2 Storyline

The story starts with a delightful morning at home. The main character, who is leading the learning content, is going on a picnic with her elder sister, assuming spending a peaceful and usual day, but it turns out facing adventures and quizzes during and after the way. The current demo of our system has ten rooms built in, including a main page, a select menu, seven stages, and a data presentation page. The seven stages contain three main concepts: number introduction in the first toward the third stage, one-to-one correspondence in the fourth and fifth, and the order of numbers in the sixth and seventh stages, which are shown as the previews for every stage with each concept in Figure 3.



*Figure 3.* The previews for every stage of the game.

It is an interactive game combined with exploration and situations inspired by television programs with similar concepts such as "Dora the Explorer" and "Micky Mouse Clubhouse". A goal, with three main lessons, would be told first at the beginning of each episode. The lessons are obstacles during the process of reaching the goal, with the main character using the word "we", guiding audiences to progress together without just being an audience (Ryan, 2010). There are three models, observing, collecting, and delivering, that could be added repeatedly to the lesson based on the current storyline in each episode. Additionally, partners or tools would show up as assistance (e.g. Dora's backpack, Micky's transportable device Toodles), providing comments and hints, making the audience feel like a helpful warrior (Dietrich, 2015).

## 2.3 Learning Materials

Avoiding arithmetic in lessons, learning aids are integrated with must-learn concepts from preschool to lower-grade elementary school. They are presented in a specific order, assisting students in understanding and exploring on their own. The contents are arranged

according to preschoolers' learning characteristics and are laid out with terms and icons they are familiar with. Moreover, they are imported with tutorials, explanations, and summaries (Lee, 2018). The chapters are subdivided into several difficulties, starting with basic ones and then increasing in complexity step by step.

#### 2.4 Implementation

We use GameMaker Studio 2 as the development foundation of the system and GameMaker Language (GML) as the visualized coding language. Furthermore, we choose joyful melodies as background music and higher pitch voice lines, providing more attraction and motivation for children (Sudarmilah, 2013). We also approach the composition from children's point of view (Cheng, 2021), including naivety and simple illustrations, bright pastel colors, and a 2-Dimension platformer layout.

The system was designed to operate on mobile devices, though it was built to be multiplatform. Since preschoolers have difficulty using mobile devices, we rearranged the controls to only involve touching and dragging. Besides, in order to provide more specific operations, objects that are interactable were marked with animation to indicate when they can be interacted with. However, due to the database not connecting to the system, researchers would need to manually write down the results after a user finishes all stages. Despite the inconvenience, it makes no difference without the function.

We follow the steps below to investigate the concept awareness of a participant. Conducted an initial clinic interview about the activities the participant engages in with parents indoors and outdoors. Noted the participant's preliminary knowledge of mathematics and mobile devices. Participants then began operating the system. Recorded the process and the end of the operation, including the frequency of questions asked, mistakes made, distractions during the progress, as well as the operating time of each stage and overall. Additionally, the frequency of screen touches was noted. Conducted a second clinic interview to gather feedback on the system. The surveys were modified from Keller's (2010) ARSC model to account for the cognitive abilities of preschoolers.

## 3. Preliminary Evaluation

The Kid+ App is still under construction but the prototype now is available for preliminary study. To collect the feedbacks from the preschool experts who have the rich experiences interacting with preschoolers, we invited two kindergarten teachers, one of whom is a kindergarten principal and the other one is a preschool teacher. Meanwhile, we also included an assistant professor who is an expert in the field of early childhood education. These experts were asked to evaluate the content in our system.

#### 3.1 Viewpoints from the preschool principle and the teacher

Based on the teachers' years of experience, a two-year-old toddler is able to operate a tablet, even if s/he could not quite understand the meanings. The abilities of preschoolers' can also be different due to family factors. Children may have a higher learning ability while older children may easily lose patience with basic tasks. Teachers believe that the illustration styles applied in our system are suitable for preschoolers, and the scale and colors of the layout are well-balanced. The story line is well-scripted, giving users a feeling of listening to a story. They believe that students will have a high interest in the game and will be likely to replay it. The game takes about thirty to forty minutes to complete, which aligns with the schedule of their kindergarten. If a child shows interest in the system, s/he can absolutely concentrate on it for at least thirty minutes.

The teachers also mentioned that the number of preschoolers participating in an experiment can affect the results, so it is recommended that each preschooler operates the system individually. The teachers imagined that instead of being quiet, kids would have lots of questions to ask. Therefore, recording is necessary. In addition, when surveying preschoolers, it should be done orally by asking questions. Older children may have more independent

thoughts, while others may only respond with yes or now, and require more guidance. If questions need to be presented on a paper, they should be multiple choices with visual cues allowing children to respond by pointing at.

## 3.2 The Preschool Experts' Opinions

Since there is evidence of a gap in preschoolers' learning ability some three-year-old kids can handle math problems beyond their ages, which need to be discussed later. In the system, the professor can introduce three concepts for kids to learn, which precisely cover the basic form the numbers and operations for preschoolers. Currently, it takes three to five minutes to complete a stage, with a total of seven stages, kids would be more concentrate if each concept can be complete within seven minutes. It is proper that adding difficulties to every concept, which can serve as checkpoints or save points in the game, in case some children is impatient to stay or want to quit.

Before conducting experiments, it is essential for children to familiarize themselves with the environment, and explanations must be told beforehand. Besides, while conducting surveys with preschoolers, they may have a hard time expressing their thoughts, and could possibly forget the details. Letting children discuss the question together would provide a better opportunity for them to accurately express their feelings.

## 4. Discussion and Conclusion

#### 4.1 System Operating

Kids under three would have difficulty using the system, due to their limited reading and hearing abilities, as well as their numerical knowledge is under about three. Kids between the age of three and four can operate the system smoothly, but they would need more prior knowledge before assigning objectives. For example, it would be beneficial to introduce what a sandwich is before asking them to collect sandwiches. Kids between the age of four and five are highly recommended to use and learn from the system. However, kids over five years old may feel the system too easy, resulting in potential impatience. It is suggested to add more choices, characters, images, and stages, with some aspects of the objects be modified such as scales, rotations, colors, etc. Moreover, improvements are needed in terms of animations.

Unlike the interactive and situated television programs such as "Dora the Explorer" and "Micky Mouse Clubhouse", give questions and problems during main character explores. Though they offer answers and choices simultaneously, they cannot be interacted with due to being TV programs. While playing a game, the game should provide instructions, choices and feedbacks, also, when it comes to a stage clearance, content review is unnecessary due to reciting can have a negative impact. Otherwise, it wouldn't feel like playing a game but rather listening to a lecture, and children will be bored. Also displaying a progress bar and a pause menu is important.

#### 4.2 Concept Understanding

The "Numbers and Operation" concept in our system seems to be too easy for children above the age of four, making it more suitable for children under four to operate and learn from. However, there are always some differences between them. We expect preschoolers who use our system can achieve effective learning outcomes in the early "Numbers and Operation" concept, and understand the three concepts in the game: counting, one-to-one correspondence, and the order of numbers. In traditional school settings, the three concepts are taught through students writing down the number or representing the quantities. Nevertheless, with the system, students can learn through the simultaneous presentation of numbers, patterns, and sounds, which is more efficient and facilitates their comprehension of the concepts.

#### 4.3 Awareness

The system was designed for mobile devices, by arranging the content for preschoolers. We expect to increase children's interest, so they can continue learning outside of school. Learning math is like learning language; it needs lots of efforts and time. The more parents and teachers engage preschoolers in math practice, the better their math abilities and motivation will become. This, in turn, enable them to develop number sense and unlock their potential in math for the future. Therefore, we expect preschoolers to develop self-awareness while learning concepts. Even without the presence of their parents, they can engage in mobile learning using a mobile device. By telling the story and showing hints constantly, we aim to reinforcing their ability to grasp the concepts they have learned. Figure 3 illustrates the concept awareness within the game This approach enables preschoolers to apply the knowledge when eventually go to school.

#### References

- Anders, Y., Rossbach, H. G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., & Von Maurice, J. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly*, 27(2), 231-244.
- Bell, A. W., Swan, M. B., Crust, R., & Shannon, A. (1993). *Awareness of Learning, Reflection and Transfer in School Mathematics*. Economic and Social Research Council.
- Callaghan, M. N., & Reich, S. M. (2018). Are educational preschool apps designed to teach? An analysis of the app market. Learning, *Media and Technology*, *43*(3), 280-293.
- Cesarone, B. (2008). Early childhood mathematics: Promoting good beginnings. *Childhood Education*, 84(3), 189-209.

Chang H. C., & Jong J. T. (2009). The development of one-to-one correspondence, counting and cardinality in 3 to 5 year-olds. *The Journal of Study in Child and Education*, *5*, 185-218.

- Chang L. C., Chu I. F., & Kau C. R. (2003). Study on preschoolers' basic learning skills. *Journal of Child Care, 1*, 171-184.
- Chang L. F. (2015). The development of the test of early number and operations. *The Journal of Study in Child and Education, 10*, 87-122.
- Cheng, M. C. (2021). Ten Brothers. Tien-Wei Publishing.
- Dehaene, S. (2011). *The Number Sense: How the Mind Creates Mathematics*. Oxford University Press.
- Dietrich, C., Buck, E., & Specht, A. (2015). Exploring the relationship between preschool-aged animated television and agriculture: A content analysis of Disney Junior's Mickey Mouse Clubhouse. *Journal of Applied Communications, 99*(4), 104-117.
- Griffin, S. (2004). Building number sense with Number Worlds: A mathematics program for young children. *Early childhood research quarterly, 19*(1), 173-180.
- Harris, B., & Petersen, D. (2017). *Developing Math Skills in Early Childhood*. Mathematica Policy Research.
- Howden, H. (1989). Teaching number sense. The Arithmetic Teacher, 36(6), 6-11.
- Keller, J. M. (2010). Motivational Design for Learning and Performance. Springer.
- Kokkalia, G., Drigas, A. S., & Economou, A. (2016). Mobile learning for preschool education. International Journal of Interactive Mobile Technologies, 10(4), 57-64.
- Lee, K. M., Kim, E. K., & Yoon, J. S. (2018). Weizmann Math Dictionary for Kids. Weison Books.
- NCTM. (2000). Principles and Standards for School Mathematics. NCTM.
- Ryan, E. L. (2010). Dora the Explorer: Empowering preschoolers, girls, and Latinas. *Journal of Broadcasting & Electronic Media, 54*(1), 54-68.
- Skwarchuk, S. L. (2009). How do parents support preschoolers' numeracy learning experiences at home? *Early Childhood Education Journal, 37*(3), 189-197.
- Sudarmilah, E., Ferdiana, R., Nugroho, L. E., Susanto, A., & Ramdhani, N. (2013). Tech review: Game platform for upgrading counting ability on preschool children. *In 2013 International Conference on Information Technology and Electrical Engineering (ICITEE) (pp. 226-231).* IEEE.
- Sundqvist, P., & Nilsson, T. (2018). Technology education in preschool: Providing opportunities for children to use artifacts and to create. *International Journal of Technology and Design Education*, 28(1), 29-51.
- Wilson, A. J., Dehaene, S., Dubois, O., & Fayol, M. (2009). Effects of an adaptive game intervention on accessing number sense in low-socioeconomic-status kindergarten children. *Mind, Brain, and Education, 3*(4), 224-234.

# A Proposal and Evaluation of Learning Advising using a Generative AI

Yasuomi TAKANO<sup>a\*</sup>, Taketo TSURUBE<sup>a</sup>, Haruki UENO<sup>b</sup> & Hiroshi KOMATSUGAWA<sup>a</sup>

<sup>a</sup>Graduate School of Science and Engineering, Chitose Institute of Science and Technology, Japan

<sup>b</sup>Faculty of Science and Technology, Chitose Institute of Science and Technology, Japan \* takano210@kklab.spub.chitose.ac.jp

**Abstract:** In this paper, we propose and evaluate a method in which learner's learning history and information about goal setting and reflection are provided as prompts to Large Language Models (LLM) that enables automated learning advice generated. The proposed approach allows us to generate advice messages that are relevant to the learner's learning properties, using learners' own sentences for reflections and goal settings. The generated advice is evaluated from the point of view whether it is similar to the actual teacher's one.

Keywords: Autonomous Learning, Flipped Classroom, Generative AI, GPT, LLM

## 1. Introduction

Japanese higher education has been increasingly focusing on autonomous learning to cultivate abilities of problem solving capable of handling unpredictable circumstances. We previously indicated effectiveness of an online flipped classroom model to foster the autonomous learning (Takano et al., 2022). However, there remains a problem that for those unfamiliar with self-guided learning, developing autonomy can be challenging. Ideally, educators should provide advice along learning situations of each student, but doing this weekly is practically difficult. Recent advancements in Large Language Models (LLM) like GPT have expanded the use of generative AI across various sectors, including education.

In this study, we investigate the feasibility of using generative AI for personalized learning advice. To obtain the advice, we incorporate data from the Learning Management System (LMS) and text information about students' goals and reflections. The generated advice is evaluated from the point of view whether it is similar to the actual teacher's one, taking into account the learner's properties and the blending of reflection and goal setting.

# 2. Class Model

The class model used in this study is an online flipped classroom model. In preparation, learners will work on the textbook, exercises, comprehension tests, and assignments. All the preparatory work will be done online. Learners' autonomy is important in this model. Learners can participate in group work during class time to discuss the results of their preparatory work while those with a low level of understanding can choose a "basic class" to supplement their knowledge. They are strongly recommended to reflect on their learning attitude and process after the class time. setting up the next week's goals. The learning advice automatically generated in this study is designed to assist in setting goals for the next week. We anticipate a process where learners, after reflecting on their learning, check the advice and set goals for the next week.
# 3. Automated Learning Advising

In this study, we propose a method in which learner's learning history and information about goal setting and reflection are provided as prompts to LLM that enables automated learning advice generated. The prompts sent to the LLM consist of three elements: the role of the conversation model, the learner's information, and the output format, as shown in Figure 1. The role includes teachers' instruction to provide advice on appropriate goal setting for a self-regulated learning. This is pre information used for the advising. The learner's information includes the learning history as well as goal setting and reflection that change, depending on learners. The output format includes the sentence definition of output form for the generated advice, with instructions based on four themes: "Introduction," "Current situation," "Potential for future growth," and "Towards setting goals for next week." The length of each theme should be around 2 to 3 sentences, and a brief description of what should be conveyed in each theme is provided.



Figure 1. Schematic of prompts given to LLM.

# 4. Evaluation and Results

The evaluation was conducted using data from a single session of the "C Programming" course held at University A in the 2021 academic year. This class was attended by 90 students. The advice sentences were generated by sending prompts, to the LLM, as shown in Figure 1. As the LLM, we utilized the dialogue-based language model gpt-3.5-turbo (temperature=0.7) provided by OpenAI. The verification was carried out from two perspectives: "Are the advice sentences considering the learner's characteristics?" and "Are the advice sentences considering the reflection and goal setting?"

First, learners were classified on basis of learners' learning features that are determined from the principal component analysis and K-means clustering using the six items listed in "Learning History" shown in Figure 1. As a result, four clusters representing the learning features were created: "Good Progress" "Textbook-intensive" "Exercise-intensive" and "Stagnant". Table 1 shows the relationship between each cluster and its features. The symbols in the table indicate the achievement level of each item, in the order of " $\bigcirc$ ,  $\bigcirc$ ,  $\triangle$ , x". In the automatic learning advising, it is required to generate advice to improve the items with low achievement in each cluster.

Next, we checked that the advice is generated in consideration of the learner's characteristics. The results of the analysis for each cluster (learning property) are shown in Table 2. The values in Table 2 show the percentage of advice that correspond to each learning history item. Comparing Table 1 with 2, we found that the advice is given more often to items at lower achievement, such as " $\triangle$ , x".

Cluster	Textbooks	Exercises	Test score	Assignments	Participation in GW	Contribution to GW
Good Progress	$\Delta$	$\triangle$	$\bigcirc$	Ø	Ø	Ø
Textbook-intensive	O	Δ	0	0	0	0
Exercise-intensive	$\Delta$	O	O	Δ	Δ	Δ
Stagnant	∆orX	∆orX	X	∆orX	×	×

Table 1. Learner characteristics for each cluster

 $\bigcirc$ :Very high achievement,  $\bigcirc$ :Highly accomplished,  $\triangle$ :Not very much,  $\times$ :Low achievement %GW stands for Group Work.

Table 2. Percentage of mentions of each item in automated study advising

Cluster	Textbooks	Exercises	Test score	Assignments	Participation in GW	Contribution to GW	Number of learners
Good Progress	78.00%	76.00%	20.00%	10.00%	22.00%	26.00%	50
Textbook-intensive	33.33%	100.00%	33.33%	0.00%	33.33%	33.33%	3
Exercise-intensive	71.43%	0.00%	0.00%	14.29%	42.86%	28.57%	7
Stagnant	90.00%	80.00%	16.67%	46.67%	86.67%	13.33%	30
Total	80.00%	72.22%	17.78%	22.22%	45.56%	22.22%	90

Next, we checked the percentage of advice considering the learner's information and the percentage of advice that is considered correct. The results are shown in Table 3. The results show that 72.22% (65/90) of the advice texts are considered to be appropriate for the goal setting and reflection. It was also found that 71.11% (64/90) of the advice texts are created in the flexible expression format with learners' intention, combining learners' goal setting/reflection with learning history. Furthermore, it is found that 97.78%(88/90) of the advices are considered appropriate expressions in the text. Finally, 70.00%(63/90) of the advice was a combination of learning history and goal setting/reflection, and there were no inappropriate expressions in the text.

Table 3. Percentage of Advice Considering Learner Information and Not Inappropriate.

Consideration of	Consideration of	Mixed advice	Not inappropriate	Mixed & not
learning history	goals and reflections		advice	inappropriate advice
98.89%	72.22%	71.11%	97.78%	70.00%

\*Mixed advice : Advice combining learning history and goals/reflection.

# 5. Conclusion and Future

In the proposed method, it was found that advice sentences can be generated considering the characteristics of the learner while also considering reflection and goal setting. It is considered that the proposed method can automatically generate advice sentences that are similar to those given by actual teachers. In the future, we plan to construct a learning support system incorporating automatic learning advising and evaluate its usefulness in actual classes. The research will be conducted in compliance with the regulation regarding personal information protection for LLM in our research institute, based on informed consent.

# References

Takano, Y., Kasahara, H., Maekawa, K., Ueno, H., Yamakawa, H., Komatsugawa, H. (2022). Proposal for a Flipped Classroom Model to Promote Autonomous Learning, World Conference on Computers in Education 2022, Poster, pp.116.

# The Role of Students' Help-seeking Profiles in the Al-assisted Game-based Learning

# **Ching-Huei CHEN**

Department of Industrial Education and Technology, National Changhua University of Education, Taiwan chhchen@cc.ncue.edu.tw

**Abstract:** This study aims to investigate the impact of students' help-seeking profiles on their seeking help from AI in a digital game-based learning environment. The study involved 102 middle school students who played online game and sought help from ChatGPT while solving physics problems. The results revealed that students' helpseeking profiles such as instrumental help-seeking were related to their help-seeking behavior in the game. The findings highlight the importance of considering individual differences, such as help-seeking profiles in designing AI-assisted learning environments to better support students' learning and help-seeking behavior.

Keywords: Help-seeking, AI, game-based learning, ChatGPT

### 1. Introduction

Digital game-based learning (DGBL) has gained popularity in recent years as a way to engage students in learning activities. With the advent of Artificial Intelligence (AI), DGBL has become more sophisticated, providing personalized learning experiences that cater to individual learners' needs. One important aspect of learning through DGBL is the role of help-seeking, which can greatly enhance students' learning experiences. Help-seeking is a critical process that enables learners to overcome challenges and achieve their learning goals. The concept of help-seeking has been extensively explored to date in the classroom setting or online learning environment, there remains a paucity of evidence on its impact in DGBL.

DGBL provides a unique environment for help-seeking, as there are a myriad of help sources and channels that can be different from those in traditional classroom settings or online learning. In a DGBL environment, learners may have access to various forms of help, such as in-game tutorials, hints, and feedback, as well as external resources like chatbots or intelligent agents. For instance, it was found that students with efficient help-seeking strategies were highly interactive with virtual characters and posters in DGBL (Xu et al., 2022). While students in traditional classrooms tend to seek informal help predominantly from their peers (Karabenick & Knapp, 1991), those engaged in DGBL exhibit varying preferences (Jimenez et al., 2014). The objective of this study was to investigate the underlying factors that drive GBL learners to seek assistance from AI, considering learners' approach to seeking help, including their attitudes, tendencies, and objectives. The results of this research could enhance our comprehension of GBL learners' AI help-seeking behaviors and suggest ways to promote the use of AI for seeking help in GBL environments.

# 2. Literature review

Help seeking is a critical component of effective learning and has been studied in the traditional classroom setting and online learning such as MOOC, Blackboard, or Forum (Kizilcec et al., 2017; Koc & Liu, 2016). Research on help seeking within digital game-based learning is still in its early stages of development. The environment of digital game-based

learning offers a distinctive opportunity for seeking assistance, as it presents a diverse range of help sources and channels that may not be present in traditional classroom or online learning settings. Therefore, it is valuable to investigate help-seeking in the context of digital game-based learning. In light of the existing research gaps in this field, the purpose of this study was to explore the connection between students' help-seeking profiles and their behaviors when seeking help in the context of digital game-based learning.

# 3. Method

# 3.1 Participants

A total of 102 7th grade students (49 females and 53 males) in two different middle schools took part in this study. The students were recruited from 6 classes, and showed no statistically significant difference in their previous experience in using games to learn content knowledge.

# 3.2 AI-assisted game-based learning

A digital game-based learning platform called *Summon of Magicrystal* designed to teach physics concepts was used. In the game, students take on the role of legendary warriors fighting against a dragon invasion. *Summon of Magicrystal* is also Al-assisted in that it incorporates machine learning algorithms that adapt to individual student needs. It uses ChatGPT, a state-of-the-art language processing algorithm, to provide real-time feedback and assistance during gameplay.

# 3.3 Instruments

To measure students' help-seeking profiles, a 13-item questionnaire was adapted from Karabenick (2003). The measurement included four subscales: help-seeking threat, avoidance of help seeking, instrumental help-seeking goal, and formal versus informal help seeking. The questionnaire was modified to refer to the AI-help instead of the class, and students rated their agreement on a 5-point scale. In addition to students' help-seeking profiles, game performance and ChatGPT behaviors were also retrieved from the system. Game performance included game score, gameplay time, and attempts. The game score refers to the points the player gains based on their performance in the game. Game points or score is accumulated by successfully completing levels, with advanced levels with higher point reward compared to basic levels. Gameplay time refers to the amount of time a player spends in playing the game. Attempts refers to the number of times a player tries to complete the game. ChatGPT behaviors included the numbers of questions asked and the time spent on the ChatGPT forum.

# 4. Results

Before discussing the relationships among the help-seeking profiles, we observed that the help-seeking profiles could be a little bit different from those in similar studies (e.g., Pajares, Cheong, & Oberman, 2004; Ryan & Pintrinch, 1997). Even though in general, students' threat and avoidance help-seeking tendencies were lower than the middle point of the scale as the literature indicated, the mean of threat and avoidance help-seeking were 2.36 and 2.30 on a 5-point scale. In other words, they were not very reluctant to seek help compared to the students in other studies where the mean threat and avoidance scores 2.25 on a 8-point scale (Pajares et al, 2004) or 1.44 on a 5-point scale (Huang & Law, 2022). Their instrumental help-seeking was also stronger. These tendencies might be due to the fact that they were in a DGBL environment as students might not think too much about learning, but focus on playing the games. Our results showed that threats and avoidance are positively related. Students who viewed help-seeking as a threat also avoided help-seeking. One

interesting finding is that threats and avoidance are positively correlated to instrumental help-seeking instead of negatively correlated as found in previous studies (e.g., Pajares et al., 2004; Ryan & Printrich, 1997). This difference may be due to the nature of instrumental help that students would ask AI for help. In other words, asking AI for help probably is interpreted differently than asking people (either teachers or peers) for help. In addition, our study results found that formal/informal help-seeking is positively correlated with instrumental help-seeking. This misalignment can be due to the definition of the constructs.

#### 5. Discussion and conclusions

This study attempts to examine the relationship between students' help-seeking profiles and their involvements in seeking help from AI in the context of digital game-based learning environment. Future research can explore students' self-efficacy regarding their impact on seeking help from AI, and their potential moderating role in the relationship between help-seeking profiles and behaviors. The results of the study help advances our understanding of help seeking from AI as an informal source of help that has a positive impact in digital game-based learning.

- Huet, N., Moták, L., & Sakdavong, J.-C. (2016). Motivation to seek help and help efficiency in students who failed in an initial task. *Comput. Hum. Behav.*, *63*, 584-593.
- Jimenez, O., Acholonu, U., & Arena, D. (2014). Tug-of-war: Seeking help while playing an educational card game. In F. C. Blumberg (Ed.), *Learning by playing: Video gaming in education* (pp. 232-261). Oxford.
- Karabenick, S. A. (2003). Seeking help in large college classes: A Person-centered approach. *Contemporary Educational Psychology*, 28, 37-58.
- Karabenick, S. A., & Knapp, J. R. (1991). Relationship of academic help seeking to the use of learning strategies and other instrumental achievement behavior in college students. *Journal of Educational Psychology*, 83, 221–230.
- Kizilcec, R. F., Perez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Computers & Education*, *104*, 18-33.
- Huang, K., & Law, V. (2022). Help seeking from peers in an online class: Roles of students' helpseeking profiles and epistemic beliefs. *Educational Technology & Society*, 25(3), 1-14. https://www.jstor.org/stable/48673720
- Koc, S., & Liu, X. (2016). Turkish Online Journal of Educational Technology. An investigation of graduate students' help-seeking experiences, preferences and attitudes in online learning, 15(3), 27-38.
- Pajares, F., Cheong, Y. F., & Oberman, P. (2004). Psychometric analysis of computer science helpseeking scales. *Educational and Psychological Measurement*, *64*(3), 496-513.
- Pintrich, P., Smith, D. A., Garcia, T., & McKeachie, W. J. (Eds.). (1991). A manual for the use of the motivated strategies for learning questionnaire (MSLQ). National Center for Research to Improve Postsecondary Teaching and Learning.
- Ryan, A. M., & Pintrich, P. R. (1997). 'Should I ask for help?' The role of motivation and attitudes in adolescents' help seeking in. *Journal of Educational Psychology*, *89*(2), 329. <u>http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=9707243861&lang=zh-tw&site=ehost-live</u>
- Schunk, D. H., & Ertmer, P. A. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. Academic Press.
- Xu, Z., Zdravkovic, A., Moreno, M., & Woodruff, E. (2022). Understanding optimal problem-solving in a digital game: The interplay of learner attributes and learning behavior. *Computers and Education Open*, 3, 100117. <u>https://doi.org/https://doi.org/10.1016/j.caeo.2022.100117</u>

# Question-Driven Design Process for XAI in Active Video Watching

# Raul Vincent LUMAPAS<sup>a\*</sup>, Antonija MITROVIC<sup>a</sup>, Matthias GALSTER<sup>a</sup>, Sanna MALINEN<sup>a</sup>,

Pasan PEIRIS<sup>a</sup> & Jay HOLLAND<sup>a</sup> <sup>a</sup>University of Canterbury, New Zealand \*raulvincent.lumapas@pg.canterbury.ac.nz

**Abstract:** Designing explanations for Artificial Intelligence (AI) systems continues to be a challenge due to AI's black-box nature. Among the solutions developed to help in designing explanations in AI technologies is the Question-Driven Design Process for Explainable Artificial Intelligence (XAI) User Experience. In this paper, we report on our experiences using the question-driven design process for XAI in active video watching. We used Active Video Watching (AVW)-Space, an AVW platform developed at the University of Canterbury, as the context for AVW. We conducted a survey to elicit questions from users on the AI features of the system. In the question analysis process, three human raters categorized the user questions into the different XAI bank categories. Results show that most users tend to ask "how" and "why" questions about the AI-enabled features in the platform. The results of the question analysis will be used in mapping the determined question categories to potential XAI techniques. This can help in deciding the types of explanations to provide to users of AVW in future works on XAI in active video watching.

Keywords: Explainable Artificial Intelligence, User Experience, Active Video Watching

#### 1. Introduction

Providing explanations about how Artificial Intelligence (AI) systems make decisions continues to be a critical and pervasive design issue in AI. Works related to XAI show a more algorithmcentered view, focusing on the researcher's intuition of what a good explanation is, based on the algorithms used in the AI system (Miller, 2019). However, explanations are often demanded by users who might not have a technical understanding of AI (Liao et al.,2020). This is also true for most end-users of data-intensive educational systems and innovations (Khosravi et al., 2022). Khosravi et al. (2022) present different design approaches for providing explanations in Artificial Intelligence in Education (AIED) systems. Among these are *Participatory design and co-design*, which are aimed at giving an active voice to the end users of the AIED systems.

An example of a participatory design and co-design is the Question-Driven Design Process for XAI User Experience (UX) by Liao et al. (2021). Liao et al. developed a "novel design process to enable designers and product teams to work with a toolbox of AI algorithms, focusing on the design issue of AI explainability." This process helps AI developers and designers to map user questions to current XAI tools. Given the participatory and co-design nature of this process, it also gives an active voice to end-users in shaping the explanations in the system. The process by asking users to provide questions they have about the decisions made by the AI system. This helps in identifying user needs and requirements and, subsequently, the choices of XAI techniques, design, and evaluation of XAI solutions. This process helps users who do not have data analysis experience to effectively convey and communicate their needs when using AI systems and how explanations can address these needs. This process involves four steps: (1) Question Elicitation, (2) Question Analysis, (3) Mapping Questions to Modelling Solutions, and (4) Iterative Design and Evaluation.

In this study, we explore the use of the question-driven design process for XAI in active video watching, with Active Video Watching (AVW)-Space as a specific case. AVW-Space is an active video watching platform developed at the University of Canterbury. AVW-Space

utilizes AI in its personalized prompts (called Nudges) and in assessing the quality of comments users make on the provided videos. Machine learning (ML) models were used to assess real-time the quality of user comments and show different prompts to support user engagement (Mohammadhassan et al., 2022). We discuss only the first two steps in the question-driven design process, Question Elicitation and Question Analysis, in this paper. The results show the potential of this design process in identifying user needs and requirements in active video watching. We have developed a streamlined version of the design process, employing an efficient way of collecting user questions, categorizing these questions, and identifying potential key user requirements for XAI in active video watching.

### 2. Methods and Results

Liao et al. (2021) start the Question-Driven Design Process for XAI UX with the Question Elicitation step, which aims to collect empirical data to help in understanding user requirements for XAI UX. The process is divided into two parts: (1) asking participants to come up with questions they have for the AI and (2) following up with participants to articulate the intentions behind asking a question, and their expectations for answers to the questions. The gathered guestions are then analyzed during the Question Analysis step. Liao et al. conducted the question elicitation process by interviewing nine users of an AI system. Participants were asked questions they have for the AI and the intentions behind the participants questions. Elicited guestions are categorized into categories indicated in the XAI Question Bank. There are ten categories in the XAI Question Bank: (1) Data, (2) Output, (3) Performance, (4) How, (5) Why. (6) Why Not. (7) How to be that, (8) How to still be this. (9) What If. and (10) Others. The descriptions and sample questions described in the XAI Question bank served as a basis for the human raters in categorizing the questions. Based on the categories, the priorities for which XAI design techniques and key user requirements for XAI UX can be identified and help in the design of explanations by employing the mapped solution prescribed by the Question-Driven Design Process.

We conducted a survey in lieu of the usual interviews done for the process, given that we had many participants, and it was not feasible to interview each one separately. Question elicitation focused on the two AI features of AVW-Space: nudges and comment quality assessment. The survey contained the following two questions: (1) What questions would you be interested in asking whenever a nudge/comment quality is shown?, and (2) Why would you be interested in asking the question/s above?

We had two groups of participants. The first group were students who have previously used AVW-Space, to learn face-to-face communication skills in SENG202, a second-year software engineering course. In the study, these participants completed the survey, and could also log into AVW-Space to remind themselves of the AI features. The second group were participants who were new to active video watching. They used AVW-Space for 30 minutes before completing the survey. The goal of the survey was to elicit the question related to XAI. Figure 1 provides an overview of the process.



Figure 1. Question Elicitation and Analysis process for Active Video Watching study

There were 16 and 87 participants in the first and second group of participants, respectively. After removing blank responses and completely irrelevant questions (which were not related to the AI features and AVW-Space), 135 questions were elicited for the nudge

feature and 138 questions for the comment quality feature. In the Question Analysis step, three human raters categorized the questions. Out of the 135 questions for the nudges, only 72 questions (53.3%) were classified by all the raters in the same category. The three raters agreed on the classification for 89 questions (64.5%) out of 138 questions for the comment quality. Table 1 shows the top four categories where all three human raters agreed on the classifications.

Categories	Nudges	Comment Quality
How	1	19
Why	4	17
How to be that		9
Others	15	12

Table 1. Top 4 Categories for Questions on AI Features in AVW-Space

As there were three raters, the inter-rater agreement was determined using Krippendorff's Alpha. The raters agreed in categorizing questions related to comment quality, having a value of 0.725 which is beyond the acceptable minimum Krippendorff's Alpha value ( $\alpha$ >0.66) (Krippendorff, 2010). However, there was low agreement when categorizing questions related to nudges ( $\alpha$ = 0.523). The raters agreed on more questions related to comment quality in comparison to questions related to nudges.

#### 3. Conclusions and Future Work

In this study, we utilized the question-driven design process to identify user needs and potential choices for XAI techniques in designing explanations in active video watching integrated with AI features. Given the high number of participants in our study, it was not feasible to interview them individually. Instead, we conducted a survey to collect user questions on the AI features. Based on the results from the question analysis step, it can be observed that there were more questions categorized into the XAI question categories on questions for comment quality as compared to nudges. Results show that there were more questions categorized. This can possibly show that the use of AI in assessing comment quality is clearer than nudges. Therefore, "how" and "why" questions are priorities, and techniques for designing explanations to these questions would benefit most users.

As this study only focused on the first two steps of the design process, future work will explore the next two steps. The results of the question analysis phase will be used to map the question categories to potential XAI techniques in future works. Given that there are multitudes of XAI techniques and tools available, the mapping will help in determining specific XAI techniques and tools appropriate for integrating explanations in Active Video Watching. This will then also help in deciding the types of explanations to provide for users of Active Video Watching in future studies.

#### References

- Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable Artificial Intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074. https://doi.org/10.1016/j.caeai.2022.100074
- Krippendorff's Alpha. (2010). In N. Salkind, Encyclopedia of Research Design (p. 670). SAGE
- Publications, Inc. https://doi.org/10.4135/9781412961288.n206
- Liao, Q. V., Gruen, D., & Miller, S. (2020, April). Questioning the AI: informing design practices for explainable AI user experiences. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-15).
- Liao, Q. V., Pribić, M., Han, J., Miller, S., & Sow, D. (2021). Question-Driven Design Process for Explainable AI User Experiences (arXiv:2104.03483). arXiv. http://arxiv.org/abs/2104.03483
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. Artificial intelligence, 267, 1-38.

Mohammadhassan, N., Mitrovic, A., & Neshatian, K. (2022). Investigating the effect of nudges for improving comment quality in active video watching. Computers & Education, 176, 104340.

# Enhancing Mobile Learning App for Revamped Blended Learning Class in Beginner's Chinese Course

Yuichi OHKAWA<sup>a\*</sup>, Xiumin ZHAO<sup>b</sup>, Takashi MITSUISHI<sup>c</sup>, Wen GUI<sup>b</sup> & Xuan WANG<sup>b</sup> <sup>a</sup>Graduate School of Education, Tohoku University, Japan <sup>b</sup>Institute for Excellence in Higher Education, Tohoku University, Japan <sup>c</sup>Center for Data-driven Science and Artificial Intelligence, Tohoku University, Japan <sup>\*</sup>yuichi.ohkawa.e1@tohoku.ac.jp

**Abstract:** The authors have been conducting Chinese language classes for first-year university students that combine face-to-face instruction and e-learning for review. They have also been developing a mobile application for review. Beginning this year, the curriculum has been modified so that one of the two linked weekly classes will be conducted as an on-demand class, and the app will be utilized to consolidate the learning content. The mobile application was also substantially modified accordingly. This paper describes the design.

Keywords: Smartphone learning app, blended learning, Chinese, speech recognition

### 1. Introduction

The authors have been implementing blended-learning for beginner Chinese language classes that combine face-to-face teaching and e-learning review. Since 2018, we have been using KoToToMo Plus, a mobile app with microlearning that allows students to review the content of those classes (Ohkawa et al., 2018). In addition, to encourage students to use the app, a My Page to track their learning status was introduced in 2022 (Nakagawa et al., 2022).

New students at our university are required to take a second foreign language class twice per week. In the previous curriculum, the two weekly classes were taught by different instructors and used different textbooks. However, there are issues with learning effectiveness and quality assurance when conducting parallel classes that are not coordinated. Hence, beginning in 2023, the two classes were converted into a cooperative class using the same textbook. The video-based on-demand class is one of the two weekly classes. After the class, students use a mobile app to learn vocabulary, grammar, and conversation. The other class is an in-person meeting.

Due to the change in curriculum, the mobile app's purpose has shifted from reviewing face-to-face classes to practicing for retention in on-demand classes, necessitating a modification. This paper describes the design of the resulting modified app and the issues for the speech recognition-based learning content included in it, as well as the methodology for addressing these issues. This paper constitutes part of a study to determine whether students can learn without difficulty after the curriculum change compared with the previous app and whether the learning motivation designed for the previous year's curriculum is effective.

# 2. Design of Learning App Modifications

# 2.1 Problems of the Existing App

The existing apps were insufficient for the new curriculum in terms of types of learning and the amount of learning per class (Gui et al., 2023). Since the total number of classes per year, including the on-demand type, would double, the number of learning chapters in the app was increased from 18 to 23 chapters. In addition, when students used the My Page function to check the chapters that needed to be studied based on the class's progress, they were required to locate the chapters with upcoming deadlines on the calendar screen, which was not fully utilized.

# 2.2 Design of the Modified App

The second, fourth, and fifth authors, who are also class instructors, proposed the addition of the following learning types.

- Vocabulary Practice 1: Students first listen to an audio example of the word, then read it aloud. Their voice is then converted to text using speech recognition, and their pronunciation is determined.
- Vocabulary Practice 2 and Grammar Practice 2: Students evaluate their comprehension of a word's four tones, pinyin, or grammar through multiple-choice questions.
- **Conversation Practice:** Role-playing conversations containing four to five sentences from the textbook allows students to practice listening and speaking with others.

In Vocabulary Practice 1, in acquiring word pronunciation using speech recognition, students must check the model voice and perform recognition numerous times. As a result, we designed it as depicted in Figure 1(a).

As shown in Figure 1(b), the Conversation Practice screen was designed to resemble a chat screen, alluding to the SNS screens that are popular among Japanese university students and with which they are familiar, so that users can practice while reviewing the conversation's content before and after the conversation. However, these forms using speech recognition were required to modify the questions to resolve issues with the speech recognition function of the operating system (OS), as described in the following section.

For the My Page, the design of the existing screen was modified so that the user can directly view which chapter is closest to the deadline and the number of unlearned questions, as shown in Figure 1(c).



# 3. Speech Recognition Issues and Solutions

The speech recognition capabilities of mobile OSs are designed for native speakers and do not account for the pronunciation errors of nonnative speakers. However, obtaining a corpus

of Chinese speech utterances by Japanese speakers is difficult, and preparing a specialized recognition system is incapable. Therefore, we decided to use a standard OS function for speech recognition to determine pronunciations, allowing for certain problems.

In speech recognition, results are obtained by taking into account not only the similarity of pronunciation but also the frequency with which a word occurs as a sentence. Therefore, even if a single word is attempted to be recognized, if it is likely to be used in a sentence, the recognition result may contain an unspoken subject word. Proper nouns that do not frequently appear in the conversation of native speakers are likely to be misidentified. Moreover, monosyllable-only short words and differences in tone of voice are also strongly affected by the probability of occurrence as a sentence, making correct recognition challenging for even native speakers.

Therefore, in the Vocabulary Practice 1 learning form, we asked native speakers to use a prototype of the modified app and replaced difficult-to-recognize words with words and expressions that could be recognized, even if they differed from those in the textbook. Most of the replaced words were short words consisting of single vowels and were made recognizable by adding Chinese adjectives equivalent to "very" or "one." Conversely, the example sentences in the Conversation Practice format contain a significant number of proper nouns. It is challenging to eliminate proper nouns and reconstruct conversational sentences. Therefore, even if the recognition result is incorrect, the system is designed to prompt recurrence up to four times, depending on the edit distance from the correct answer, before permitting the user to continue the conversation. This is because we placed greater emphasis on oral practice than on obtaining accurate pronunciation.

#### 4. Conclusions

This paper describes the design of a mobile learning application that has been modified to meet the requirements of the new blended Chinese classes that will be implemented in Tohoku University's beginner's Chinese courses beginning in the academic year 2023. It also described the changes made to the learning questions' content and the learning method's design due to the challenges of speech recognition. The app based on the above design has been in use since April 2023, with 27 classes and more than 900 new students.

Currently, we are preparing to analyze the log data obtained from the app's use, the learning effects compared to the previous years' curriculum, motivational effects, and other learning factors.

#### Acknowledgments

A part of this work was supported by JSPS KAKENHI Grant Numbers 19H04223, 19K00875, and 20K03119.

#### References

Ohkawa, Y., Kodama, M., Konno, Y., Zhao, X., & Mitsuishi, T. (2018). A Study on UI Design of Smartphone App for Continuous Blended Language Learning, *Proc. of ICBIR 2018*, 584-589.

- Nakagawa, R., Ohkawa, Y., Zhao, X., Takahashi, A., Ohyama, T., Mitsuishi, T., & Hayakawa, Y. (2022). Implementation of UX Design to Enhance Spontaneous and Continuous Study of a Mobile Application for Foreign Language Learning. *Proc. of EdMedia + Innovate Learning 2022*, 1119-1123.
- Gui, W., Zhao, X., Ohkawa, Y., Mitsuishi, T., & Wang, X. (2023). Renewal of Smartphone App for Blended Learning by Beginning Learners of Chinese. *Bulletin of Institute for Excellence in Higher Education Tohoku University,* 9, 141-147. (In Japanese)

# Investigating Players' Social Interactions in IOT Board Games

### Yueh-Chi Wang, Ju-Ling Shih, Shaun-Wen Chen

Department of Network Learning Technology, National Central University, Taiwan \*I0983560431@gmail.com

**Abstract:** This study aims to investigate the impact of integrating Internet of Things (IoT) technology with board games to enhance social interaction among players. Given the widespread application of IoT in various domains, incorporating it into board games can offer more intricate gameplay mechanisms and a broader array of interactions. Subsequent to gameplay, interviews will be conducted to gather players' reflections on the gaming process and their motivations for engaging with different players. This study anticipates that the interdisciplinary board game, combined with IoT technology integration, will further augment the diversity and richness of social interaction among players.

Keywords: Internet of Things, Game-Based Learning, Board Games, Social Interactions

# 1. Introduction

In recent years, Internet of Things (IoT) technology has found broad application across various domains such as IoT transportation, IoT healthcare, and IoT home automation. Nevertheless, the utilization of IoT in education remains relatively uncommon. Thus, this study aims to integrate IoT technology into educational board games, using Taiwan's history as a backdrop. This approach enables players to assume diverse roles and immerse themselves in historical events, effectively recreating past scenes. By means of this interdisciplinary board game, players can engage in more intricate social interactions with one another.

# 2. Related work

# 2.1 Internet of Things System

The Internet of Things (IoT) technology enables different objects to receive and transmit information through the internet, achieving information sharing and automated operations (Al-Fuqaha et al., 2015). IoT can be applied in various fields; sensors embedded in household appliances can remind users to refill water, while agricultural sensors can assist in determining the best planting time and improving yield (Chaudhary et al., 2019). This study employs IoT technology in board games, introducing novel experiences to the gameplay through object communication and calculations. This integration serves to amplify the educational and learning aspects of the game.

#### 2.2 Social interaction

Social interaction is an essential part of human society, and the foundation for building relationships among people. With the advancement of technology and the development of IoT, the way people interact socially is constantly changing. After introducing technology elements into teaching, technologically-rich classroom interactions are more complex and diverse comparing to traditional classrooms. This complexity is not only reflected in the increase of interaction frequency, variation of interaction targets, and amount of interaction contents, but also has a significant improvement in interaction quality (Wang, Kong, & Huang, 2016). This paper investigates the incorporation of IoT in board games, accomplished through the

continuous display of varying values on students' mobile devices via background computation. Such an application has the potential to encourage social interactions among students, enrich their interest in social learning, and enhance their participation in the game.

# 3. IOT Board Game

This board game is based on Taiwanese history. The app was developed using APP Inventor, as shown in Figure 1, and game objects composed of Mbot and RFID sensors, as shown in Figure 2. Each group of players has a mobile devices that displays the basic values of the Mbot they control. The buildings on the map contain RFID chips that affect the values of each player when Mbot senses them, and players can view the value changes displayed on the mobile devices. During the game, players need to use the mobile devices to control the movement of the Mbot on the map. Each game round triggers various significant historical events from Taiwan's history. Players assume the roles of different historical figures and collaborate to solve these events. Additionally, each team is assigned unique tasks, and players must collect various items during the game to enhance the capabilities of their respective teams.

This interdisciplinary board game fuses technology and social interaction, delivering a more immersive gameplay experience compared to traditional board games. While playing, player-to-player communication and interaction, as well as tactile engagement with game objects, are processed in the background and reflected on mobile devices. This enables players to visually observe value changes and attain data visualization. These values, stemming from these interactions, can directly or indirectly influence every move a player makes. Beyond in-person social engagement during gameplay, the integration of the Internet of Things also ensures that each player's decision yields distinct outcomes, consequently impacting social interaction and elevating the overall gaming experience.



Figure 1. APP introduction design



Figure 2. Mbot and RFID sensor

# 4. Research design

This study investigates the integration of IoT into a board game to augment social interactions. Students will be randomly divided into four groups, each assuming an ethnic role within the game. The game is expected to span approximately two hours. Following the game, individual players will be interviewed to ascertain whether fluctuations in game values influenced their strategies, perceptions, or inclination to interact with particular ethnic groups. Additionally, players will receive guidance to reflect on their experiences and complete a feedback questionnaire.

# 5. Expected Result

Currently, games are primarily categorized into two groups: digital games and physical board games, each carrying its own set of advantages and drawbacks. Digital games tend to emphasize audio-visual effects but lack social interaction, whereas traditional board games offer social engagement while having somewhat restricted game mechanics.

Thus, this study seeks to merge physical board games with Internet of Things (IoT) technology to enrich the depth of social interaction among players. The study incorporates IoT technology into physical board games, preserving the intricacy of digital games along with the interpersonal social interaction characteristic of physical board games. This endeavor not only enriches the game's diversity and entertainment quotient but also utilizes the data gathered from the game to analyze players' interaction patterns and collaborative behaviors. This aids in comprehending the extent of their social interaction during gameplay and allows for the assessment of players' learning performance within the game. The study uncovers their approaches to problem-solving and sheds light on their problem-solving abilities. By scrutinizing learning data, strengths and weaknesses in learning can be pinpointed, offering valuable insights for instructional design.

### Acknowledgements

This study is supported in part by the National Science and Technology Council (previously known as Ministry of Science and Technology) of Taiwan, under MOST 108-2511-H-008 -016 -MY4 and NSTC 112-2410-H-008 -010 -003.

- Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of Things: A survey on enabling technologies, protocols, and applications. *IEEE Communications Surveys & Tutorials*, *17*(4), 2347-2376.
- Chaudhary, S., Johari, R., Bhatia, R., Gupta, K., & Bhatnagar, A. (2019). CRAIoT: Concept, Review and Application(s) of IoT. In 2019 *4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU)* (pp. 1-4). Ghaziabad, India: IEEE. doi: 10.1109/IoT-SIU.2019.8777467.
- Wang, X. C., Kong, S. C., & Huang, R. H. (2016). Influence of digital equipment on interaction quality in technology-rich classroom. In 2016 *IEEE 16th International Conference on Advanced Learning Technologies (ICALT)* (pp. 455-459). doi: 10.1109/ICALT.2016.82

# Method to Promote Social Facilitation of Learners by Presenting Writing Sounds

### Tatsuya UEDA<sup>a\*</sup>, Tatsuya HAMADA<sup>a</sup>, & Hironori EGI<sup>a</sup>

<sup>a</sup> Department of Informatics, The University of Electro-Communications, Japan \*t-ueda@uec.ac.jp

**Abstract:** We propose a method to induce the presence of other learners using writing sounds without interfering with self-guided study. For learners who are more likely to progress when they feel the presence of others, it can be difficult to continue independent self-guided study alone at home. In addition, in certain environments, e.g., libraries and learning commons, where other learners are present, the presence of others may positively influence the continuation of learning. Thus, the proposed system attempts to induce the presence of others according to the learner's level of concentration, which is evaluated by the learner's leg movements. The proposed system was introduced experimentally to determine whether it can induce the presence of others in an environment with no other learners engaged in self-study activities.

Keywords: Learning support, social facilitation, self-study

### 1. Introduction

The ability to maintain motivation in autonomous learning is important in self-guided study. However, distractions frequently interrupt learning or prevent the learner from returning to study after taking a break. In addition, when working on tasks, e.g., preparing for exams or writing reports, some learners prefer to learn in a library or learning commons rather than at home. In other words, the presence of others may contribute to a learner's motivation to continue self-guided study. Thus, this paper proposes a method to induce the presence of others by presenting writing sounds according to the target learner's level of concentration.

A previous study proposed the "Enlight-Pen" device (Yoshihara, et al., 2003), which was designed to support learning by presenting information according to the learning status of others. The "Enlight-Pen" device supports the continuation of autonomous learning by illuminating LEDs embedded in a pen. However, flickering LED may interfere with the learner's concentration. Thus, in this study, we designed a system that plays a fluctuating writing sound to induce the presence of others. Writing sounds occur frequently during learning tasks; thus, we considered that such sounds could naturally communicate an environment in which others are learning.

In the proposed method, leg movements are measured (Aikawa et al., 2019) to estimate the target learner's level of concentration because a correlation has been demonstrated between leg movements and a learner's subjective fatigue.

# 2. Design

The proposed system comprises a single-board computer (Raspberry Pi Model 3 B+), a leg movement measurement module, a speaker (Creative Metallix), and a server PC. The leg movement measurement module was designed based on the findings of a previous study (Hamada et al., 2022). The leg movement measurement module comprises a passive infrared ray sensor (EKMC2609112K) and a single-board microcontroller (Arduino Nano Every). Figure 1 shows the installation position of the leg movement measurement module. The voltage values of the passive infrared sensor are acquired at 10 Hz and transmitted to the

server. The server determines whether the learner is concentrating or not based on the amount of leg movement. Then, the server sends the results to the Raspberry Pi, which in turn varies the volume of the writing sound according to the user's state, i.e., concentrating or not concentrating.



Figure 1. Leg movement measurement module (left) and installation position (right).

# 3. Experiment

An experiment was conducted to investigate the psychological influence and the effect on learning activities when using the proposed system during self-study. The experimental environment was set up assuming a self-study environment in which the proposed system was installed. This experiment was conducted with one group of participants using the system as the experimental group and a control group that did not use the proposed system. The subjects were science and engineering university students. Figure 2 shows the experimental environment and the flow of the experiment.

In this experiment, the participants took a test comprising 50 true or false questions before and after performing self-study. The content of the self-study was divided into three parts: Textbooks A, B, and C. Each part of the study was 30 minutes long and involved taking notes. At the end of the experiment, a questionnaire about the learning environment was administered. The questionnaire consisted of questions on a five-point scale and free comments regarding the experimental environment. In addition, a textbook related to a certification examination related to food advisors was used as the learning material.



Figure 2. Experimental environment (left) and experiment flow (right).

# 4. Results

The results of the pre-tests and post-tests are shown in Table 1. The results of t-tests on the pre-test results for the experimental and control groups exhibited no significant differences.

Similarly, no significant differences were observed in the results of the post-test. Thus, the proposed system does not necessarily contribute to learning outcomes.

	Experime	ntal group	Control group	
	Avg.	S.D.	Avg.	S.D.
Pre-test	30.20	3.35	28.63	3.78
Post-test	42.40	4.16	38.38	4.69
Diff.	12.20	3.35	9.75	4.17

Table 1. Pre-test and post-test results.

The mean value of the questionnaire for the experimental group, i.e., "The presentation of sound made me feel the activity of the person I was learning with," was 3.40. This result suggests that the proposed system may have induced a certain degree of the presence of others. In addition, the mean value of a question, i.e., "I could not concentrate due to the sound presentation," was 3.40. This result indicates that the presentation of sound may have interfered with the learner's level of concentration.

Comments from the experimental group's questionnaires stated that the presented sound was loud and very distracting, and the recorded writing sound was not similar to the sound of a pen being used during study. This suggests that the presented writing sounds were louder than the expected environmental sounds in a normal learning environment and that they did not adequately represent natural writing sounds.

# 5. Conclusion

In this study, we developed a system that presents recorded writing sounds according to the target learner's concentration. The proposed system was evaluated experimentally, and we found that the system was able to induce a certain degree of presence of others; however, it may have interfered with the learner's concentration. Thus, in the future, it will be necessary to consider an appropriate volume level that does not interfere with learning and to present sounds that accurately represent natural writing sounds. If the presentation of writing sounds can be improved to reduce the interference with the learner's concentration, we believe that it will be possible to induce the presence of others in a manner that does not interfere with concentration, which is expected to contribute positively to learning outcomes.

#### Acknowledgements

This work has been partly supported by the Grants-in-Aid for Scientific Research (Nos. 21K18496 and 21K02839) by MEXT (Ministry of Education, Culture, Sports, Science and Technology) in Japan.

- Aikawa, D., Asai, Y., & Egi, H. (2019). Proposing an Estimation Method of Mental Fatigue by Measuring Learner's Leg Movement. In Learning and Collaboration Technologies. Designing Learning Experiences Proceedings, Part I 21: 6th International Conference, LCT 2019, Held as Part of the 21st HCI International Conference. HCII, Orlando, FL, USA, July 26-31, 2019. Springer International Publishing, pp. 227-236.
- Hamada, T., Terui, Y., & Egi, H. (2022). Estimating Activity Conditions of Students in Class by Measuring Leg Movement. APSCE the 30th International Conference on Computers in Education (ICCE2022), Vol. 2, pp. 656-658.
- Yoshihara, S., Tsukada, K., & Yasumura, M. Enlight-Pen: The Pen for Helping Learners to Carry on the Self-Paced Learning (in Japanese). IPSJ interaction 2003, pp.193-194.

# **Designing Learning Companions for Enhancing Students' Writing Habits**

#### **Chang-Yen LIAO**

Department of Hakka Language and Social Sciences, National Central University, Taiwan \*CalvinCYLiao@gmail.com

**Abstract:** This study confronts the issue of insufficient writing practice among elementary students, a concern prevalent in Taiwan. Despite the integral role of writing in learning, students frequently need more opportunities for consistent writing practice, affecting their reading and writing abilities. This research proposes a novel approach through an artificial intelligence-based "Writing Companion," designed to supplement the writing practice gap in current education. Three design mechanisms of the Writing Companion include facilitating autonomous writing process. By encouraging consistent writing habits and promoting a balance of writing practice and strategy learning, this study aims to enhance students' writing literacy and motivation. The findings may provide critical insights into harnessing artificial intelligence to improve writing education.

Keywords: Writing Companion, Student Autonomy, Writing Habits

#### 1. Introduction

Learning to write is one of the best ways to cultivate students' self-expression, and writing plays a crucial role in learning. Writing is a tool for students to gather and organize information, develop their knowledge, and use language to communicate.

Some studies (Cutler & Graham, 2008) have suggested reforms for primary school writing education: (a) increase students' writing time; (b) increase time for writing expository text; (c) provide a better balance between writing time, learning writing strategies, and teaching writing skills; (d) focus more on fostering students' writing motivation (Peng & Liu, 2020). Taiwanese researchers (Chen, Xie & Huang, 2007) also pointed out that to become an excellent writer, we must help students broaden and deepen their life and reading experiences. These experiences become long-term memories that are easy to retrieve and provide continuous and intensive writing experiences. Students need opportunities to discuss and share excellent works with their peers. In the long run, students cannot cultivate writing literacy through a small amount of writing practice. In other words, there is only good writing literacy with sufficient writing experience (i.e., sufficient writing volume). Continuous writing experience can help students form writing habits, establish interest, and acquire writing skills.

Therefore, this study attempts to develop an artificial intelligence-based learning companion, the "Writing Companion," to allow elementary students to practice writing autonomously with three mechanisms to build students' writing habits. This study aims to fill the gap in current school education that cannot arrange enough teaching for writing, leading to a general lack of writing practice among students (Fernsten & Reda, 2011), a lack of writing willingness, and a lack of opportunities to practice writing other types of text.

# 2. Building Students' Writing Habits: Three Design Mechanisms

#### 2.1 Supporting Autonomous Writing Opportunities

"Writing Companion" provides a unique platform for fostering autonomous writing habits among students. Its 24/7 availability allows students to engage in writing exercises at their

convenience, thereby facilitating the development of regular writing habits outside the traditional time constraints of classroom instruction (Sun, 2010). "Writing Companion" encourages a personalized approach to writing practice by catering to individual learning needs. This targeted approach addresses specific learning needs and contributes to building consistent practice routines. For instance, if a student frequently struggles with grammatical errors, "Writing Companion" can suggest additional grammar exercises.

A distinct advantage of "Writing Companion" is its capability to provide real-time feedback. Unlike traditional writing exercises, where students often have to wait for the teacher's feedback, an immediate response from "Writing Companion" allows students to instantly recognize their areas of improvement. This prompt feedback mechanism can facilitate more active engagement in writing exercises, contributing to the establishment of regular writing noutines. In summary, "Writing Companion" is a valuable tool for promoting regular writing habits among students. It enhances the writing learning process through immediate feedback and personalized practice, making it a beneficial resource for cultivating autonomous writing habits and improving writing skills.

#### 2.2 Providing Prompts for Writing Practice

"Writing Companion" can be crucial in providing prompts for independent writing practice. Using natural language processing and machine learning technologies, "Writing Companion" can generate diverse writing prompts tailored to a student's current learning phase, personal interests, and areas of development (Proske, Narciss, & McNamara, 2012). This approach keeps students engaged and encourages them to integrate writing into their learning routine. This personalized interaction can help students develop the habit of structuring their narratives effectively. For example, if a student is learning narrative writing, "Writing Companion" might suggest prompts that guide the student to create a story with a clear beginning, middle, and end. Similarly, for an advanced student working on argumentative essays, "Writing Companion" can offer prompts that require crafting a persuasive argument on a complex topic. This regular practice can help the student establish a habit of building coherent and compelling arguments.

Beyond generating prompts, "Writing Companion" provides ongoing support throughout the writing process, making it a part of the student's habitual writing practice. If students struggle with developing a thesis statement, "Writing Companion" can offer constructive tips and examples. If a student faces difficulties structuring an argument, "Writing Companion" can ask guiding questions, making it a habit for the student to think in a structured manner while arguing a point. In summary, "Writing Companion" can be an intelligent and responsive partner, offering many writing prompts. By personalizing these prompts based on each student's unique learning journey, "Writing Companion" can significantly enhance the writing learning process. This approach can encourage students to adopt regular writing practice, thereby fostering effective writing habits.

# 2.3 Helping Students Monitor their Writing Process

"Writing Companion" is instrumental in fostering a regular writing habit among students. This tool allows students to manage their writing practice, creating a continuous record of their progress that can be revisited for reflection and further learning (Xiang, 2004).

For instance, "Writing Companion" can store and display a history of a student's writing submissions and the provided feedback. Reviewing their past work and comparing it with their current efforts can encourage students to maintain a consistent writing practice. Self-monitoring and reflection are essential for developing metacognitive skills crucial for effective learning. Moreover, "Writing Companion" can provide real-time analytics, offering students insight into their writing practice. These could highlight the frequency of writing exercises, recurring grammatical errors, or areas of improvement over time. Such immediate and personalized feedback can inspire students to stick to a regular writing schedule by demonstrating their progress and making their learning goals more attainable.

"Writing Companion" can also facilitate the setting and tracking personalized goals, motivating students to maintain a habitual writing practice. For instance, a student might aim to decrease the frequency of a specific grammatical mistake or increase the complexity of their sentence structures. "Writing Companion" can monitor the student's progress towards these goals, providing regular updates and encouragement, thereby establishing a routine of goal-oriented writing.

In summary, "Writing Companion" equips students with a powerful tool for fostering a regular writing habit, promoting self-reflection, goal setting, and a deeper understanding of their learning process. This innovative approach holds promise in supporting independent writing practice and enhancing writing skills.

# 3. Upcoming Work

This study engages approximately 60 third and fourth-grade primary school students in an experimental evaluation lasting at least one semester to cultivate robust writing habits. Initial data collection includes the assessment of student's current writing habits, interests, and skills, which inform the development of tailored writing activities. Students interact with the "Writing Companion" system during allocated writing sessions. This platform guides students through various writing tasks, assisting in goal setting, task selection, and execution via interactive dialogue. Post-session, students can evaluate their progress by reviewing their personal knowledge graph and writing trajectory within the system, updating their writing goals, and consulting the Writing Companion for different writing tasks. Post-semester data collection measures the evolution of students' writing habits. Concurrently, the system's backend is monitored to capture changes in students' writing behaviors over the semester. Thus, this study aims to foster sustainable student writing habits through consistent engagement with the Writing Companion system.

### Acknowledgments

The authors thank the National Science and Technology Council of the Republic of China, Taiwan, for financial support (NSTC 111-2410-H-008 -067 -MY3).

- Chan, T. W., Looi, C. K., Chen, W., Wong, L. H., Chang, B., Liao, C. C. Y., Cheng, H. N. H., Chen, Z. H., Liu, C. C., Kong, S. C., Jeong, H., Mason, J., So, H. J., Murthy, S., Yu, F. Y., Wong, S. L., King, R. B., Gu, X., Wang, M., Wu, L., Huang, R., Rachel Lam, R., & Ogata, H. (2018). Interest-Driven Creator Theory: Towards a Theory of Learning Design for Asia in the Twenty-First Century. *Journal of Computers in Education*, 5(4), 435–461, <u>https://doi.org/10.1007/s40692-018-0122-0</u>
- Chen, M. F., Xie, J. N., & Huang, K. R. (2007). An Analysis of Factors Affecting the Writing Performance of High School Students. Journal of Special Education Research, 32(3), 63–86.
- Chen, W. L., Chan, T. W., Wong, L. H., Looi, C. K., Liao, C. C. Y., Cheng, H. N. H., Wong, S. L., Mason, J., So, H. J., Murthy, S., Gu, X., & Pi, Z. (2020). IDC theory: habit and the habit loop. *Research and Practice in Technology Enhanced Learning*. *15*(10).
- Fernsten, L. A., & Reda, M. (2011). Helping students meet the challenges of academic writing. *Teaching in Higher Education*, *16*(2), 171-182.
- Peng, B. W., & Liu, Y. J. (2020). A Discussion on the Dilemma of Mandarin Chinese Writing Instruction in Elementary Schools. Taiwan Education Review Monthly, 9(7), 123-128.
- Proske, A., Narciss, S., & McNamara, D. S. (2012). Computer-based scaffolding to facilitate students' development of expertise in academic writing. Journal of Research in Reading, 35(2), 136–152.
- Sun, Y. C. (2010). Extensive writing in foreign-language classrooms: A blogging approach. *Innovations in Education and Teaching International*, 47(3), 327-339.
- Xiang, W. (2004). Encouraging self-monitoring in writing by Chinese students. ELT Journal, 58(3), 238– 246.

# Practices of ARCS Chinese language instructional design with MR application participation

Zhenni SHI <sup>a\*</sup>, Yuto NAGATA <sup>b</sup> & Yusuke MORITA <sup>c</sup>

<sup>a</sup>Graduate School of Human Sciences, Waseda University, Japan <sup>b</sup>School of Human Sciences, Waseda University, Japan <sup>c</sup>Faculty of Human Sciences, Waseda University, Japan \*shi\_zhenni0414@ruri.waseda.jp

**Abstract:** This inquiry employed the ARCS instructional design theory and an authordeveloped Mixed Reality (MR) application. To ascertain its efficacy, seven Chinese language learners who are Japanese native speakers, consisting of three males and four females (average age = 44.43 years, SD = 16.59), participated in a validation experiment. They were divided into two groups and underwent three times 100-minute Chinese language classes. The MR application, rooted in TPR, gamification, and engagement theories, emphasizes individual learning and integrates shared experiences in mixed reality, fostering collaborative learning. Evaluation encompassed engagement, motivation, learning strategies, achievement scores, and affective scales. The outcome reveals enhanced intrinsic motivation and collaborative engagement after three times classes.

**Keywords:** Mixed Reality, ARCS instructional design, Chinese language learning, engagement

#### 1. Introduction

Since the 1990s, the globalization of the economy has led to a rising number of Japanese individuals studying Chinese (Hirai, 2016). Japanese universities are witnessing an increasing trend in students selecting Chinese as their second foreign language. This surge is reflected in the growing implementation of Chinese language courses across Japanese universities, with over 15,000 classes and approximately 500,000 students annually (Sunaoka, 2017). Notably, most Chinese language learners, as indicated by the 'Six Language Survey,' are beginners, constituting more than 70%. Despite generally having high motivation, these learners often need a strong intrinsic drive compared to learners of other languages. Pronunciation practice dominates class time, hampering the development of practical skills (Sunaoka & Yamaguchi, 2014, Sunaoka *et al.*, 2016). In light of these findings, it is evident that research into effective design for Chinese language courses targeting Japanese learners is imperative.

Concurrently, addressing the decline in academic motivation and achievement has led to increased interest in student engagement as a potential solution. Engagement is believed to be adaptable, influenced by contextual factors, and responsive to environmental modifications. Student engagement encompasses three primary dimensions: behavioral, emotional, and cognitive, with collaborative engagement also considered (Kuh *et al.*, 2001). Simultaneously, the emergence of Mixed Reality (MR) technology, characterized by interactivity, the fusion of virtual and real elements, immersion, and unique physical attributes, offers opportunities to create captivating and interactive learning tools. The ARCS Motivation Model was created to enhance our grasp of motivation for learning and to address motivation-related issues systematically. It outlines four key conditions (Attention, Relevance, Confidence, and Satisfaction) for initiating and sustaining motivation (Keller, 1987).

In this study, an MR app designed to harness the interactive and immersive qualities of MR (Holz *et al.*, 2011) is utilized to boost classroom interest and elevate student

engagement levels. This study aims to formulate an instructional design using the ARCS model and incorporate the MR app into the textbook, thereby enhancing Chinese learners' engagement, internal motivation, improvement in learning strategies, and achievement of higher academic performance.

# 2. Design

## 2.1 MR Application

The MR application, designed for this research, incorporates gamification, Total Physical Response (TRP) theory, engagement theory, and pedagogical agents. It features six segments: S0 for introducing objectives, S1 for speech-to-text word learning, S2 for interactive word games, S3 for sentence practice, S4 for verb-object pronunciation, and S5 for grammar exercises. S2 and S5 promote collaborative learning with shared virtual objects, enhancing motivation and engagement.

### 2.2 ARCS instructional design

Table 1 shows a Chinese language class that was structured using the ARCS model and MR application. Each class spanned 100 minutes, with participants attending three sessions to emphasize collaborative activities and creative learning. The approach aimed to move beyond conventional mechanical exercises, focusing on meaningful language acquisition and application.

The introductory session employed real-life examples to familiarize learners with course content. The MR application's S0 section engaged learners, leveraging the MR headset's (Microsoft Hololens 2) immersive visual experience to stimulate interest. This session employed the Attention and Relevance strategies of the ARCS model. During the personal learning with MR session, learners utilized the MR application on Hololens2 to master new vocabulary and sentences. The app's voice recognition and feedback features allowed for personalized learning. Positive reinforcement and timely feedback further motivated learners, utilizing the Relevance and Confidence strategies of the ARCS model. In the listen to the lecture session, teachers addressed learners' challenges from the personal learning session. Pronunciation, sentence composition, and meaning were clarified. Grammar was emphasized for subsequent group tasks, employing the Attention and Relevance strategies. Group tasks prompted learners to apply knowledge gained, crafting dialogues and diary entries. Teacher feedback enhanced learners' confidence and satisfaction, embracing the Confidence and Satisfaction strategies. Group tasks with the MR involved pairs or trios completing MR application sections to reinforce vocabulary and grammar. Interactive features aimed to boost concentration, confidence, satisfaction, and collaborative engagement, combining Attention, Confidence, and Satisfaction strategies. The summary session saw teachers recap lesson content, address issues, and recognize group achievements. The ARCS model's Attention and Satisfaction strategies underpinned this session.

Sessions	ARCS model	Time(min)
1 Introduction	Attention, Relevance	5
2 Personal learning with MF	R Relevance, Confidence	20
3 Listen to lecture	Attention, Relevance	20
4 Group task	Confidence, Satisfaction	30
5 Group task with MR	Attention, Confidence, Satisfaction	20
6 Summary	Attention, Satisfaction	5

Table 1. ARCS instructional for Chinese Class Design

#### 3. Evaluation

To verify the effectiveness of the instructional design, seven participants (3 males and 4 females, average age = 44.43 years, *SD*= 16.59) were invited to participate in the pilot experiment. They were Japanese native speakers and beginners in learning Chinese, had



Figure 1. Average points of the intrinsic motivation scale

mastered Pinyin pronunciation. and possessed some rudimentary knowledge of Chinese. The questionnaire on the intrinsic motivation scale (6 items) was made based on the research of Okada (2019). The Behavioral engagement scale (4 items) and Emotional engagement scales (5 items) were taken from the Student Engagement Scale (Skinner et al., 2009), and the Collaborative engagement scale was taken from the research survey by Kuh et al. (2001). Figure 1 shows the result of intrinsic motivation. A withinsubjects ANOVA analysis on participants' intrinsic motivation scale responses before and after each session showed а

significant increase overall (F(3,18) = 5.32, p < .01). Emotional and behavioral engagement remained consistent across sessions. Collaborative engagement significantly increased in the second and third sessions (second-time class F(3,18) = 5.40, p < .01; third-time class F(3,18) = 6.02, p < .01), but not in the first session. Comparisons revealed higher collaborative engagement during group tasks with MR and group task sessions compared to personal learning with MR sessions.

The future objective involves optimizing the teaching model for better utilization of the MR application to achieve more effective learning outcomes.

#### Acknowledgments

This work was supported by JST SPRING, Grant Number JPMJSP2128.

- Hirai, K. (2016). Chinese Language Education in Japan and the World, and the Chinese Language Education Society. *Current Status and Challenges of Multilingual International Information Dissemination from Japan*. Asahi Publishing Company.
- Holz, T., Campbell, A. G., O'Hare, G. M. P., Stafford, J. W., Martin, A., & Dragone, M. (2011). MiRA Mixed reality agents. *International Journal of Human-Computer Studies*, 69(4), 251–268.
- Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of Instructional Development*, 10, 2–10.
- Kuh, G. D. et al., (2001). National Survey of Student Engagement: The College Student Report: NSSE Technical and Norms Report. Bloomington, IN: Indiana University Center for Postsecondary Research and Planning.
- Okada, R. (2019). The Influence of Interaction Activities in Classroom Situations on Intrinsic Motivation. *Bulletin of Educational Research and Teaching Development*, Kagawa University (Bull. Educ. Res. Teach. Develop. Kagawa Univ.), 39, 35-43.
- Sunaoka, K. (2017). Preface to the Series "Systematization of Foreign Language Education and Classroom Management." *Waseda University Graduate School of Political Science and Economics, Working Paper Series* No.J1608.
- Sunaoka, K., & Yamaguchi, T. (2014). Classroom Practices in Teaching French, German, Chinese, Spanish, and Russian in Japan: A Quantitative Analysis of the 2012 University Instructor Survey based on the New Language Education Perspective. In Comprehensive Research Report on a Common Language Education Framework for Multiple Foreign Languages Based on a New Perspective on Language Education (pp. 64-79).
- Sunaoka, K., Yamaguchi, T., & Hori, S. (2016). Current State of Foreign Language Education in Japanese Universities: Analysis of Learning Motivation of Second Foreign Language Learners and Teaching Activities of Instructors. In K. Sunaoka & S. Muroi (Eds.), *Current Status and Challenges of Multilingual International Information Dissemination from Japan: Prospects for Human Resources and Global Communication* (pp. 121-132). Asahi Publishing Company.

# The Impact of Metaverse Worlds on International Collaborative Learning for Cross-Cultural Understanding

#### Masako HAYASHI<sup>a\*</sup>, Takehiro SUZUKI<sup>a</sup>, Yuki KAWATA<sup>a</sup> & Keisuke GOTO<sup>a</sup> <sup>a</sup>Tohoku University, JAPAN \*masako.hayashi.c5@tohoku.ac.jp

**Abstract:** We examine factors that influence the students' memorability of the presentations given in the Metaverse International collaborative learning class, where students work in groups to give presentations to promote cross-cultural understanding. The results show that the students chose the presentation of the group using Metaverse World as the most memorable, with the "Metaverse World" being the reason.

**Keywords:** Metaverse Worlds, International Collaborative Learning, Cross-Cultural Understanding, Collaborative Presentation

# 1. Introduction

In the 'Metaverse International Collaborative Learning with Overseas Partner Universities' class, a collaborative presentation was given for the purpose of cross-cultural understanding. In this class, students from partner universities abroad, domestic and international students in Japan work together in a mixed environment of students from different cultural backgrounds. The introduction of the Metaverse in the HyFlex class format has improved various issues (Hayashi et al., 2022).

Through observing students in their cross-cultural understanding studies using the Metaverse, we thought that the worlds in the Metaverse might be useful to foster a deeper cultural understanding. These worlds can be freely created and edited, and they allow for a variety of experiences through the objects, functions, and sounds in the world, compared to simply viewing slides.

Therefore, we set a research question, "What influences the memorability of the presentations for the students who listen?" To answer this research question, we compared monomedia and multimedia such as metaverse worlds, VR videos, 2D videos, and images to determine what learning content was memorable for the students. This allowed us to understand how differences in the learning environment affected the students' learning.

# 2. Comparison of this study to previous studies

There is a study about cultural understanding in which students from three countries learn about Japanese culture in a metaverse (Inaba et al., 2023). Another example includes Shadiev et al. (2021), which is a study of cross-cultural understanding using VR.

However, to the best of our knowledge, there are no previous studies in which students from many different countries created their own worlds for the purpose of cross-cultural understanding, as in this study. Students used these worlds and VR movies to introduce their own culture and other cultures in an international collaborative presentation.

# 3. Research Methods

In the first semester of 2023, about 40 students (17 Domestic students, 13 international students, 10 overseas students) from 15 countries participated in the Intercultural

Collaborative Learning class. Overseas students are students from cooperating schools in Kenya, Indonesia, Belgium, etc., who participated mainly through Metaverse while still living in their own countries. In mixed groups of students from Japan and abroad, students give international collaborative presentations. In addition to explanations on slides, students introduce aspects of their culture by creating a world in the Metaverse, or by projecting a VR video of their own creation. The Metaverse platform used in this study is VRChat. At the beginning of the class, most students had no experience with the metaverse. In the first half of the class, the learners were divided into six groups according to their choice of the theme of the presentations, which introduce Japanese culture in comparison with other cultures around the world.

After the presentation, a report assignment was given, stating "Please list the three most memorable presentations. Write what was memorable and explain why." Of the 40 students, 24 submitted their reports by the due date, making this a valid response. First, in response to the question "Please list the three most memorable presentations," the votes received by each group were divided into two groups: the most memorable group and the first to third most memorable groups. Next, the comments were analyzed quantitatively for the question "Write what was memorable and explain why." In doing so, we divided the comments into three categories: monomedia represented by textual information, multimedia other than textual information, and others not belonging to these two categories. Furthermore, multimedia was divided into four categories: images, 2D videos, VR videos, and Metaverse worlds. The number of classmates who cited each media as a reason for their memorability was tabulated. If more than one media was mentioned in one person's memorability, scores were given to each of those media. Three of the four authors tabulated the results individually, and the four of us discussed the different results at a later date.

### 4. Result

Figure 1 shows the results of the ranking of the most memorable groups after the class. Group 1 and Group 6 created and used Metaverse worlds for their presentations, while the remaining groups used VR videos. Table 1 summarizes the reasons why the presentations were memorable, given in the comments, per attribute.



Figure 1. Ranking of the most memorable groups (number of people)

Table1. The reasons why the presentations were memorable

			Others			
	monomedia	Images	2D videos	VR videos	Metaverse World	Others
1st	8	1	6	0	18	3
2nd	13	1	0	3	12	1
3rd	14	2	0	7	2	7
total 35	25	4	6	10	32	11
	30		5			

As illustrated in Figure 1, the two groups that created and used Metaverse worlds for their presentations (Group 1 and Group 6) received 12 and 8 votes, respectively, in the most memorable group category. The average number of votes received by the other four groups, that did not create Metaverse worlds, was one vote. Of these groups, even the most memorable group, Group 5, received only 2 votes. This shows there is a significant difference between groups which did create Metaverse worlds and those which did not. In addition, for the combined results of the three most memorable groups, Group 1 and Group 6 were again the highest, with 19 votes and 18 votes, respectively. The other four groups that did not create a Metaverse received an average of 8.75 votes.

As shown in Table 1, from the comments describing the reasons why each group was memorable, many of the comments cited the presenting group's own Metaverse world

as the reason. 18 comments, the largest number for the most memorable group, cited the Metaverse world as the reason the presentation was memorable, while 8 comments cited monomedia as the reason it was memorable.

# 5. Discussion

The most common reason given for the top three most memorable groups was multimedia (52 comments), followed by monomedia (35 comments). In the four categories of multimedia, Metaverse world received the highest number of comments (32 comments), followed by VR videos (10 comments).

Thus, it was found that the groups that used the Metaverse worlds for their presentations were more likely to leave a lasting memorability on their classmates. In addition, many of the comments below indicate that the creation and use of Metaverse worlds in accordance with the content of the presentation deepened the cross-cultural understanding. The use of multimedia was more memorable than monomedia. In international collaborative learning where there is a language barrier, multimedia may be more helpful for learners to understand than text alone.

A degree of active experience was consistent with higher learners' ratings of memorability. Among the types of multimedia, the metaverse world allows learners the most active experience. While VR videos are superior to 2D videos in that the viewpoint can be selected independently, they are still more passive than the metaverse world. Furthermore, images, which offer a passive experience, were rated lowest.

# 6. Conclusion

In this study, where students gave international collaborative presentations, the groups that created and used Metaverse worlds occupied the first and second most memorable positions. In addition, of the reasons for being most memorable, the number one commented reason was the Metaverse world.

The Ministry of Education, Culture, Sports, Science and Technology (MEXT) issued "Guidelines for the Implementation of Distance Education at Universities and Colleges of Technology," in which it recommends the "Introduction of Metaverse." The method used in this study, which has a major development compared to Hayashi et al. (2022), utilizes Metaverse not only as a tool for communication, but also as a tool for students to learn about cultural understanding. This study demonstrates that the use of Metaverse worlds in presentation has an impact on student's memorability. The results obtained in this study contribute to the development of research studies on distance education using Metaverse.

# Acknowledgements

This research was supported in part by a grant from the Takahashi Foundation for Industrial and Economic Research, the Center for Knowledge Creation, Tohoku University Research Promotion and Support Organization, and by a joint project research by the Research Institute of Electrical Communication, Tohoku University.

- Hayashi, M., Yoshida, H., Maruyama, N., & Suzuki, T. (2022). Advantages and Challenges of Intercultural Collaborative Learning Using Metaverse. Annual Report of Center for Language and Culture Education. Tohoku University, 78-86.
- Inaba, M., Hosoi, K., Thawonmas, R., Nakamura, A., Uemura, M. (2023). Research on Supporting Japanese Culture Learning Mediated by "Play" on Metaverse. Replaying Japan, 5, 27-33.
- Shadiev, R., Wang, X., & Huang, Y. (2021). Cross-cultural learning in virtual reality environment: facilitating cross-cultural understanding, trait emotional intelligence, and sense of presence. Educational Technology Research and Development 69(5), 2917-2936.

# Designing an Online Course on Learning Analytics for Educators: Preliminary Insights from a Scoping Review

# \*Nurbiha A SHUKOR<sup>a</sup>, Norah MD NOOR<sup>a</sup>, Aini Marina MA'ROF<sup>b</sup>, Noor Dayana ABD HALIM<sup>a</sup>, Matthew MCLAIN<sup>c</sup>, Phillip ROTHWELL<sup>c</sup>, Graham DOWNES<sup>c</sup>, Frances TRACY<sup>c</sup>

<sup>a</sup>Universiti Teknologi Malaysia, Malaysia <sup>b</sup>Universiti Putra Malaysia, Malaysia <sup>c</sup>Liverpool John Moores University, UK

**Abstract:** This paper presents a project to develop graduate digital future work skills in data analytics. We describe a review activity carried out to guide the collaborative development of content for an online Data Analytics Course for Educators in Malaysia and the United Kingdom. Our methodology for systematically searching online learning platforms for current courses is described. We also give an overview of our preliminary findings in relation to the duration of learning, the types of online materials used and the range of topics included. Finally we describe our future plans to review the quality of the courses in relation to the depth, scope and complexity of the learning they promote.

**Keywords:** Data analytics, Learning Analytics, Content development, Online learning, Micro credential

# 1. Introduction

A new partnership called "i-Digitics" has been formed between Universiti Teknologi Malaysia (UTM), Universiti Putra Malaysia (UPM) and Liverpool John Moores University (LJMU) to develop future employability skills for a workplace driven by technological change. We are developing a new international curriculum on data analytics and offering it to education students online. Data analytics increasingly impact on educators at all levels and understanding the structure, uses and manipulation of data will help to empower those professionals to use them for enhancement of teaching and student support (Ndukwe and Daniel, 2020). We describe our initial review of currently available online courses on Data Analytics in Education. The content review represents a key task for our research group to identify the strengths and gaps in current provision to guide our own development of curriculum content to meet the needs of our target audience. The methodology describes the approach we have taken, future plans for analysis and our preliminary findings and progress. This will be of interest to other researchers and course developers in this field.

# 2. Methodology

There are many learning platforms that offer online courses on learning analytics and related content coverage such as data science, big data, and data analytics. This study applies a three-stage scoping review methodology to carry out the needs analysis for an educators' learning analytics course. The three stages include reviewing the learning platforms, selection criteria and online courses selection.

In the first stage, the available online learning platforms were analyzed and it was found that there are eighty-three 83 online learning platforms offering thousands of online courses as reported by Class Central (Shah, Pickard, Ma, 2023). From these online learning platforms, the top online learning platforms hosting the largest number of courses are: Udemy (28,216 courses), Coursera (14,638 courses), Linkedin Learning (11,802 courses), Skillshare (9455

courses) and edX (5888 courses). From this selection, the following research question is formulated for scoping review:

- 1. What is the extent and range of the current courses offered online related to learning analytics?
- 2. What is the characteristic of the online courses related to learning analytics in terms of quality (depth of learning topics) and core content (duration of learning, range of learning topics, learning materials)?

The next stage is to identify the selection criteria for reviewing the online courses in the respective learning platforms based on English Language as the language of learning, duration of learning that is minimum at 8 hours and study focus that only covers analytics in education, data literacy and learning analytics.

Finally, based on the selection criteria, specific keywords were used to search for courses that relate to learning analytics. Each learning platform was queried three times, each time with one of the following search keywords. The keywords used are 'learning analytics', 'analytics and education', 'learning analytics and education'. Each query was then filtered to include courses in English Language and learning duration of minimum eight hours.

# 3. **Preliminary Findings**

Based on the selection criteria of language and duration of learning, there were a total of 395 courses selected for further investigation consisting of 63 courses in Udemy, 169 courses in Coursera, 74 courses in Linkedin Learning, 20 courses in Skillshare and 69 courses in edX.

Upon review based on the study focus, there were ten (10) courses found with 3 courses in Udemy, 3 courses in Coursera, 2 courses in Linkedin Learning, 0 courses in Skillshare and 2 courses in edX. A description of each of these courses is provided in Table 1 of our appendix handout at the conference. The key features of the courses are described as follows:

# 3.1 Duration of learning

The preliminary findings show that an online course that lasts around 8 hours is best for courses with quick overviews. The duration is suitable for introducing learners to a specific topic or providing a basic understanding of a subject. A course within 8 to 16 hours allows for a more in-depth exploration of a topic, appropriate for intermediate-level content where learners can gain a solid grasp of foundational concepts and even apply them in practical exercises. In longer duration courses exceeding 16 hours, these courses have complex content, teach advanced skills, or provide comprehensive learning experiences.

#### 3.2 Learning materials

Online learning materials refers to the educational assets shared on digital platforms. It can be reading material, videos, audios, interactive things such as simulations, pictures, quizzes, and assignments. The preliminary findings suggest that all the online courses provide learning materials in the form of video recordings, reading materials, and assessment in the form of self-paced short quizzes.

# 3.3 Range of Learning Topics

In an online learning course, a "topic" refers to a specific subject, theme, or area of knowledge that is being taught and studied. Learning topics in online learning courses are divided into different modules or units, and each module may cover one or more topics. Topics serve as the building blocks of the course's curriculum, guiding the progression of learning from basic to more advanced concepts. The preliminary findings show that the learning topics related to learning analytics range from fundamental concepts to technical content. Thematic analysis will be conducted to explore the depth of the range of the learning topics in the 10 courses.

## 3.4 Quality of Courses

In this study, we posit that the quality of the courses will be determined by the depth of the learning content. Following this developing scoping review, our next focus will be a qualitative evaluation of these identified 10 courses. Using Bloom's taxonomy (Bloom, 1956) in its most recent form as updated by Anderson and Krathwol (2001), we aim to assess the depth, scope and complexity of the learning outcomes in these courses, providing insight into their educational efficacy and quality.

#### 4. Conclusion

Learning analytics plays a crucial role in today's teaching practices. As the educational sector continues to produce and utilise vast amounts of data, teachers need skills to interpret and apply this data effectively in their classrooms. This ability enables them to make more informed decisions about student learning, adapt teaching strategies, and evaluate educational outcomes with greater precision.

Our scoping review provides an initial overview of online courses that may be relevant for teachers to understand learning analytics. Preliminary findings from the 10 courses indicate that courses on analytics are generally designed into three learning durations: courses that require less than 8 hours of completion time, 8-16 hours, and more than 16 hours of learning time, ranging from fundamental topics in data science to highly technical aspects of data analytics. However, the learning materials available are still limited to mainly recorded videos, and a sparse number of downloadable reading materials and one-way exercises.

By identifying and analysing these courses, our review aims to identify gaps that do not fully cater to the specific needs of educators and highlight the potential areas for course development and refinement. Ensuring that teachers have access to pertinent and effective training in learning analytics is crucial for continued advancement of educational practices. While many online courses exist, their relevance and alignment with the educational context warrants further exploration. This work, therefore serves as an initial step, aiming to bridge the existing knowledge gap and support educators in their professional growth.

#### Acknowledgements

i-Digitics project is supported by a UK-Malaysia University Partnerships Catalyst Grant from the British Council's Going Global Partnerships programme. The programme builds stronger, more inclusive, internationally connected higher education and TVET systems.

# References

Anderson, L. W., Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing: A Revision of Bloom's Taxonomy of Educational Objectives. New York: Longman.

Bloom, B. S. (1956). Taxonomy of educational objectives. Vol. 1: Cognitive domain. New York: McKay, 20, 24.

Chucherd, O., Vallibhakara, S. A. O., Paiwattananupant, K., Puranitee, P., Wattanayingcharoenchai, R., & Vallibhakara, O. (2023). The effect of online video-assisted teaching program on medical students learning procedure of fractional curettage. *BMC Medical Education*, 23(1), 1-8.

Malakul, S., & Park, I. (2023). The effects of using an auto-subtitle system in educational videos to facilitate learning for secondary school students: learning comprehension, cognitive load, and satisfaction. *Smart Learning Environments*, *10*(1), 4.

Ndukwe, I.G., & Daniel, B.K. (2020) Teaching analytics, value and tools for teacher data literacy: a systematic and tripartite approach. International Journal of Educational Technology in Higher Education, 17 (22)

Shah, D., Pickard, L., & Ma, R. (2023). Massive List of MOOC Platforms Around the World in 2023. https://www.classcentral.com/report/mooc-platforms/

# Development of a Music Educational Board Game with Mobile Device: Learning Musical Theory and Emotional perception through Gameplay

## Song-Zhu XIAO, Chih-Chen KUO, Huei-Tse HOU\*

Mini Educational Game Development Group, Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taiwan <u>\*hthou@mail.ntust.edu.tw</u>

**Abstract:** Music appreciation is one of the most important courses in arts education. In this study, we designed a music education board game "*Soul Mate*" with mobile devices, which provided learners with multiple representations learning context. In the game, learners scan the QR codes on the cards with their mobile devices to listen to the music. Through the process of listening experience, learning experience and observing interaction, learners can make judgment, understand their own emotional perception of listening to music, and learn about music theory. Preliminary results indicate that gamified learning is effective in enhancing learners' learning outcomes and is able to maintain learners' flow and keep them engaged in learning.

**Keywords:** mobile device, music educational board game, music theory, emotional perception

#### 1. Introduction

Music appreciation is one of the most important courses in the arts field of higher education, and it can cultivate the art appreciation ability of university students (Lv, 2022). Thompson et al. (2022) refer to the fact that art appreciation education often treats the learner's level of enjoyment as a response to a work of art and neglects the importance of intellectual understanding. Over the years educational research has explored the effectiveness of many applications of technology in the teaching and learning of subjects such as language, science, and mathematics, but few studies have been conducted on subjects in the field of music (Wang, 2021). Research has found that the use of technology in music education can be very effective and the use of games learning can help music skills and knowledge acquisition (Gower & McDowall, 2012). After the epidemic, the interactive use of mobile devices in arts education has gained wider attention (Vorontsov, 2021), because mobile devices can offer the possibility of presenting musical works (Ouyang, 2022). In this study, a music education board game with a mobile device was developed to allow learners to learn music theory, to be aware of their own emotional perceptions when listening to music, and to promote the flow of the learners' mind through the game.

# 2. Method

This board game contains four types of cards: Music, Music Theory, Emotion, and Event (Figure 1). The Music card is presented by QR code, using Chopin's 24 preludes as the source material (including 24 major and minor keys). The music theory cards, divided into two categories: tone and tempo, e.g., listening to the music and identifying the tempo of the piece as Allegro Major, Allegro Major, or Allegro Minor, can help learners to acquire knowledge of music theory. The emotion cards are designed with reference to Russell's (1980) model of the dimensions of emotion, and they contain a total of 12 emotions, including excitement, relaxation, sadness, and fear, etc. The content of the event card is about daily topic, which can be used to stimulate interactions and discussions among the learners. This board game is

divided into two stages. In the first stage, after listening to a piece of music, learners choose the corresponding music theory and emotion cards (Figure 2), learn music theory from checking the correct answers, and at the same time observe their own emotional perceptions of the music. In the second stage, learners take turns to act as a guide, and the other three learners guess the guide's emotional feelings towards the music based on the observations made in the previous stage to achieve the learning interactions and to confirm the knowledge of the emotional feelings. The participants in this study were 20 students (12 males and 8 females) from a university in Taiwan. After the learners enrolled online, the researcher assigned four people into a group to play the game according to the participation time. The pre-test and post-test of this study were based on the same assessment of learning effectiveness, which was divided into two parts, namely, music theory and emotion perception. The questions were developed by music teachers with more than 5 years of teaching experience in universities. The Kiili Flow Scale (2006) translated and revised by Hou & Li (2014) was referred to in this study. The flow scale includes two dimensions: Flow Antecedents and Flow Experience. All scales were scored according to the Likert scale, and the reliability of the flow questionnaire (Cronbach's alpha= 0.914) indicated a high degree of internal consistency. The learning process included game explanation (10 minutes), pre-test (20 minutes), game activities (50 minutes), post-test and flow guestionnaire (30 minutes).



Figure 1. Cards and Instructions for Games



Figure 2. Playing "Soul Mate"

# 3. Results and Discussions

The purpose of this study is to investigate the music theory, emotional perception, and flow performance of learners on the music education board game "*Soul Mate*". Table 1 showed the results of the learners' learning performance of music theory and emotion perception before and after participating in the game. The post-tests of music theory (z = -3.03, p = 0.002 < 0.05) and emotion perception (z = -3.73, p = 0.000 < 0.05) were significantly higher than the pre-tests, and it suggested that the use of game learning in the subject of music domain can effectively enhance learners' effectiveness and emotion perception. The results showed that the use of game-based learning in the music field can effectively enhance learners' learning. Table 2 shows learners' state of flow during the game and, the result revealed that the overall mean of the flow (M = 4.15, SD = 0.52) was significantly higher than the median 3 of the scale (z = 3.853, p < 0.001). Most of the Flow Antecedents and Flow Experience sub-dimensions were also significantly larger than the median of 3 on the scale, and many sub-dimensions were even higher than 4. This result suggested that the overall game mechanism enables learners to clearly understand the activity goals of the game and engage in the game to achieve a high flow experience.

-		•	•	-	
Dimension		М	SD	Z	р
Music Theory	pre-test	44.70	19.04	3 03**	0.002
	post-test	64.30	11.64	3.03	0.002
Emotion Percention	pre-test	8.75	5.83	3 73***	0.000
	post-test	16.90	3.32	5.75	0.000
**					

Table 1. Music Theory and Emotion Perception Descriptive Statistical Analysis

\*\**p* <0.01 <sup>, \*\*\*</sup>*p* <0.001

Dimension	M	SD	Z	р
Overall Flow	4.15	0.52	3.853***	0.000
Flow Antecedent	4.07	0.56	3.801***	0.000
Challenge-skill balance	4.05	0.71	3.706***	0.000
Goals of an activity	4.35	0.69	3.862***	0.000
Unambiguous Feedback	3.97	0.72	3.510***	0.000
Sense of Control	4.30	0.71	3.760***	0.000
Action-awareness Merging	3.68	0.75	2.995**	0.003
Flow Experience	4.09	0.49	3.873***	0.000
Concentration	4.30	0.75	3.681***	0.000
Time distortion	4.10	0.66	3.750***	0.000
Autotelic experience	4.60	0.68	3.914***	0.000
Loss of self-consciousness	3.35	0.86	1.576	0.115
Overall Acceptance	4.15	0.52	3.853***	0.000
Perceived Usefulness	4.07	0.56	3.801***	0.000
Perceived Ease-of-use	4.05	0.71	3.706***	0.000

Table 2. Flow and Game Acceptance Descriptive Statistical Analysis

\*\**p* <0.01 · \*\*\**p* <0.001

### 4. Conclusion

In this study, we developed a music education board game "*Soul Mate*" with a mobile device. The game combines emotion perception, contextual interaction, and cognitive design through the medium of music, allowing learners to learn in a multi-character board game. This board game is for 4 players to play together, use the mobile device to scan the QR code on the cards to listen to the music, and then realize their own emotional perception and music theory through self-feeling. According to the results of the study, *Soul Mate* had a positive effect on the learners' knowledge of music theory and emotional perception, and the learners were highly engaged in the process. For the future study, we will continue to explore the effect of the study on the learners compared to the control group (traditional teaching) and collect more qualitative data such as interview and behavior pattern analysis.

# Acknowledgements

This research was supported by the projects from the Ministry of Science and Technology, Republic of China, under contract number 110-2511-H-011 -004 -MY3, MOST-111-2410-H-011 -004 -MY3.

- Gower, L., & McDowall, J. (2012). Interactive music video games and children's musical development. *British Journal of Music Education, 29*(1), 91–105. https://doi.org/10.1017/S0265051711000398
- Hou, H. T. & Li, M. C. (2014). Evaluating Multiple Aspects of a Digital Educational
- Problem-solving-based Adventure Game. *Computers in Human Behavior, 30*, 29-38.
- Kiili, K. (2006). *Evaluations of an experiential gaming model*. Human Technology: An Interdisciplinary Journal on Humans in ICT Environments.
- Lv, H. (2022). Analysis of the Basic Characteristics and Teaching Environment and Mode of Music Appreciation Course Based on Core Literacy. *Journal of Environmental and Public Health*, 2022.
- Ouyang, M. (2022). Employing Mobile Learning in Music Education. *Education and Information Technologies*, 1-17.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology,* 39(6), 1161.
- Thompson, W. F., Bullot, N. J., & Margulis, E. H. (2022). The psychological basis of music appreciation: Structure, self, source. *Psychological Review*.
- Vorontsov, C. O. (2021). Mobile devices in teaching information technologies at institutions of higher art education. *Information Technologies and Learning Tools, 83*(3), 100–112.
- Wang, Y. H. (2021). Exploring the effects of using various designs of game-based materials on music learning. *Interactive Learning Environments*, 1-15.

# Designing a Virtual Reality Game for Religious Culture Guided Tour by Combining Voice Guided Scaffolding and Situated Learning Mechanism

Jui-Jong WANG, Chih-Chen KUO, Huei-Tse HOU\*

Mini Educational Game Development Group, Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taiwan <u>\*hthou@mail.ntust.edu.tw</u>

**Abstract:** The study developed an online virtual reality educational game, "*Temple Keeper Training Day Camp*", integrating an online virtual reality game, voice guiding and situated learning mechanism. The learners play the role of the temple keeper trainee and have to utilize the knowledge of the religious cultural customs and also the history of the city where the temple located to help and fulfill all the tasks from the pilgrims. We investigated learner's learning performance, flow status and game acceptance. The results revealed that their learning performance had been significantly improved, indicating that the game did effectively improve the learner's knowledge about the temple customs and history of the city. Also, the high flow status and game acceptance both showed the learners had an enjoyable and fruitful learning experience.

**Keywords:** Virtual Reality Game, religious culture, voice guided tour, scaffolding, situated learning

#### 1. Introduction

Culture heritage preservation not only has gain worldwide attention but also faces great challenge, especially intangible cultural heritage (UNESCO 2003). Assisting different generations to understand conventional customs and traditions deeply and to avoid these important cultures being simplified and disappeared is necessary (Chang, 2008). Virtual reality technologies with user-friendly interface and high level of immersion for the participants increase and stimulate the interests of the users (Schouten, van den Hooff, & Feldberg, 2010), and convey context and the experience of the intangible cultural heritage (Selmanović et al., 2020). Digital game-based learning is considered to effectively arouse learner's interest and learning motivation (Clark et al., 2016), and flow (Sun et al., 2017). Therefore, this study expects a learning method that combines a virtual reality game, voice guiding and situated learning mechanism to improve learner's interest and motivation and gradually enhance the learners' knowledge of religious culture of the temples in Taiwan. Thus, our research team developed a virtual reality game, "Temple Keeper Training Day Camp" (Figure 1 and 2) by ThingLink, a spherical video-based virtual reality platform. A story was applied in the game as situated learning to facilitate the learners to participate in this activity. An old Temple Keeper was the trainer representing audio scaffolding to pass on the knowledge in the temple. The learners as Temple Keeper Trainee and had to utilize the knowledge they learned during the game and fulfill all the tasks from the pilgrims. Once the learners entered a scene, they heard the audio guiding from the trainer telling the fun facts of the temple as guided scaffolding to simulate the situated learning context. The learners had to explore by clicking the interactive buttons and fulfill the tasks from the pilgrims. The tasks were about the religious cultural customs of the temple and also the history of the city where the temple located. The learners will receive a 3-level certificate in the end of the game according to their performance: The Rookie, the Potential and the Qualified Keeper.





Figure 1. Interactive buttons for exploring

Figure 2. Task to be solved from the pilgrims

# 2. Method

The purpose of this study is to evaluate students' learning performance, flow and game acceptance toward this educational game. Participants in this study were 18 adults, age from 20 to 59, in Taiwan (7 males, 11 females). Half of them had never been to this temple or in the same religion. All the participants joined the game with a personal computer or tablet at home. The researchers had the same questions of the pre-test and post-test with different order of questions to avoid memory effect. There were 14 questions about the knowledge of religious culture of temples in Taiwan, and 6 questions about the history of the city where the temple located. We evaluated the learners' flow status by Kilis's flow scale (2006), which was translated and revised by Hou, & Li (2014). The flow scale includes 2 dimensions, namely the flow antecedent and flow experience. All scales were scored on a five-point Likert scale. The reliability of the flow questionnaire (Cronbach's alpha=0.953) showed high internal consistency. Also, we adapted Davis's (1989) technology acceptance model to explore the learners' perceived usefulness and perceived ease of use toward the game and the reliability of it was 0.886 (Cronbach's  $\alpha$ =0.886). The procedure of the study was as follow: Pre-test for 10 minutes, game for 60 minutes and post-test, flow and game acceptance questionnaires for 20 minutes.

# 3. Results and Discussions

For learning performance, a Wilcoxon signed rank test was used to compare the results of learning performance between the pre-test and post-test. The results showed that the score for the post-test was significantly higher than pre-test (Z=-3.523, p<0.001) (see Table 1), and it suggested that learners' knowledge of religious culture had been better improved through the game. For the flow status, a one-sample Wilcoxon Test was performed, the overall flow score (M=4.26), flow antecedent sub-dimension (M=4.40), and flow experience sub-dimension (M=4.30) were all significantly above the median of the scale (the median in a five-point scale =3) (Z=3.724, p<0.000) (see Table 2). The results indicated that learners were deeply involved in the game. For game acceptance, a one-sample Wilcoxon signed rank Test was conducted, the overall game acceptance score (M=4.58), perceived usefulness (M=4.75), and perceived ease of use (M=4.35) were significantly higher than median (the median in a five-point scale=3) (Z=3.750, p<0.000) (see table 2), and it suggested that learners have highly accepted the technology game design and they didn't have trouble accessing the game. In other words, this game was easy to be familiarized with and also helped the learners explore and learn by themselves.

		_					
	М	SD	Z				
pre-test	54.44	17.73	3 503***				
post-test	79.44	10.13	-3.525				
<i>p</i> <0.05 · <i>**p</i> <0.01 · <i>***p</i> <0.001							
Table 2. Flow and Game Acceptance Descriptive Statistical Analysis							
Dimension	М	SD	Z	р			

Table 1.	The Mean	and Standard	Deviation o	of Learning	Performance
----------	----------	--------------	-------------	-------------	-------------

Overall Flow	4.26	0.58	3.724***	0.000
Flow Antecedent	4.40	0.65	3.683***	0.000
Challenge-skill balance	4.28	0.77	3.455***	0.001
Goals of an activity	4.33	0.64	3.654***	0.000
Unambiguous Feedback	4.25	0.83	3.454***	0.001
Sense of Control	4.36	0.76	3.475***	0.001
Action-awareness Merging	3.83	0.86	2.981**	0.003
Flow Experience	4.30	0.54	3.725***	0.000
Concentration	4.40	0.65	3.652***	0.000
Time distortion	3.83	0.79	3.114**	0.002
Autotelic experience	4.49	0.50	3.748***	0.000
Loss of self-consciousness	4.19	0.75	3.530***	0.000
Overall Acceptance	4.58	0.51	3.750***	0.000
Perceived Usefulness	4.75	0.44	3.857***	0.000
Perceived Ease-of-use	4.35	0.75	3.579***	0.000

\**p* <0.05 · \*\**p* <0.01 · \*\*\**p* <0.001

#### 4. Conclusion

The study developed an online educational game, "*Temple Keeper Training Day Camp*", integrating online virtual reality, voice guiding and situated learning mechanism, to enhance learners' knowledge of religious culture of temples in Taiwan and the history of the city. The results showed the game did assist the knowledge of temples and the city. Also, the high flow status indicated high engagement and enjoyment during the game, and as for high game acceptance, it suggested this game was easy for learners to manipulate and did facilitate learning. It indicated a virtual reality game integrated with voice guiding scaffolding and situated learning mechanism can be an effective way in promoting learners' performance. Future study would compare the effectiveness among different ages, genders and religions.

#### Acknowledgements

This research was supported by the projects from the Ministry of Science and Technology, Republic of China, under contract number 110-2511-H-011 -004 -MY3, MOST-111-2410-H-011 -004 -MY3.

- Chang, C. L. (2008). *Religious ritual and social change in Taiwan* (Vol. 1). LEARNING TIDE PUBLICATION CO., LTD.
- Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of educational research*, *86*(1), 79-122.
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3), 319-340.
- Hou, H. T., & Li, M. C. (2014). Evaluating multiple aspects of a digital educational problem-solving-based adventure game. *Computers in Human Behavior*, *30*, 29-38.
- Kiili, K. (2006). Evaluations of an experiential gaming model. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments.*
- Ministry of the Interior ROC. (2023, July19). Statistical Yearbook of Interior, https://rb.gy/sqyvj
- Schouten, A.P., van den Hooff, B.J., & Feldberg, J.F.M. (2010). Real decisions in virtual worlds: Team collaboration and decision making in 3D virtual worlds. In *Proceedings of the 31st International Conference on Information Systems* (pp. 1-19). St. Louis: MO.
- Selmanović, E., Rizvic, S., Harvey, C., Boskovic, D., Hulusic, V., Chahin, M., & Sljivo, S. (2020). Improving accessibility to intangible cultural heritage preservation using virtual reality. *Journal on Computing and Cultural Heritage (JOCCH)*, *13*(2), 1-19.
- Sun, J. C. Y., Kuo, C. Y., Hou, H. T., & Lin, Y. Y. (2017). Exploring learners' sequential behavioral patterns, flow experience, and learning performance in an anti-phishing educational game. *Journal* of Educational Technology & Society, 20(1), 45-60.
- United Nations Educational, Scientific and Cultural Organisation. 2003. "Convention for the Safeguarding of the Intangible Cultural Heritage". https://ich.unesco.org/doc/src/15164-EN.pdf

# Design and Preliminary Evaluation of an Educational Board Game on Urban Culture and History with Mobile Conceptual Scaffolding

Hau-An YU, Chih-Chen KUO, Huei-Tse HOU\*

Mini Educational Game Development Group, Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taiwan <u>\*hthou@mail.ntust.edu.tw</u>

**Abstract:** With the development of urbanization, local towns are at risk of losing their local culture as a result of out-migration. Local literature and history education is limited by time and teachers' qualifications. In this study, we designed a board game "Walking on Puzi Street" to assist students' learn local history and culture by combining scaffolding on the cards with the mobile conceptual scaffolding provided from the mobile app. Learners take on the role of a tour guide, leading travelers on a tour of the town of Puzi in Taiwan. The preliminary study investigated participants' flow, acceptance and sense of place. The results showed that the game mechanism can engage learners in the game, facilitate local literature and history learning, and bring them sense of interest in the place. Also, high rates of learners used mobile scaffolding.

**Keyword :** local history, educational board games, mobile scaffolding, conceptual scaffolding

#### 1. Introduction

Communities or places are important starting points for understanding oneself and one's surroundings (McInerney, smyth, Down, 2011), and Taiwan is facing the effects of issues such as urban-rural development imbalance and policies that emphasise place-making (Executive Yuan, 108). However, a lack of relevant teaching aids and students' motivation to learn about such topics is critical. (Halvorsen, Harris, Doornbos, Missias, 2021). Therefore, this study designed a board game, "Walking on Puzi Street " for learners learning local knowledge and history. Study has found that game-based learning can significantly enhance learners' intrinsic motivation (Tüzün, 2009). "Walking on Puzi Street" is a local tourism game in which players as tour guides from a travel agency in Puzi lead tourists to destinations that match their ideal preferences. Through the tourist cards (Figure 1), players can learn about 1 to 3 features that the tourists wish to visit. The closer the destination cards align with the tourists' descriptions, the higher the favor score (points) awarded to the players. Lin et al. (2023) assert that learners benefit from appropriate support when facing challenges. The game incorporates two types of conceptual scaffolding: information provided on the cards and QR codes as scaffolds during the game (Figure 2).Players receive partial information about the attractions from the destination cards, serving as conceptual scaffolds on the cards, and use this information as clues for guiding tourists to their destinations (Figure 3) on the game board. Additionally, players can obtain extended information related to solving key puzzles by scanning QR codes, providing real-time action scaffolding with detailed content about the attractions. The mobile conceptual scaffolding also allows players to learn about the historical stories of the attractions. The primary objective of this research is to conduct a preliminary analysis of learners' flow experience, sense of place, learning performance, and acceptance within the game, while observing learner behaviors.


Figure 1. Tourist cards Figure 2. Destination cards

Figure 3. Actual Gameplay

# 2. Methodology

A total of 21 participants, aged 20 years old or above, were randomly divided into groups of 3-4 to play board games. Pre- and post-testing designs have been approved by local historians. The Flow Scale designed by Kiili (2006) translated and revised by Hou & Li (2014) was referred to in this study. The reliability of the flow questionnaire (Cronbach's alpha=0.876) indicated a high degree of internal consistency. In terms of acceptance, this study adapted the acceptance scale proposed by Davis (1989), which also had high internal consistency in terms of reliability (Cronbach's alpha = 0.857). The Sense of Place Scale was modified with the Tourism Purpose Scale designed by Abou-Shouk & Zoair et al. (2018), with a high degree of internal consistency in reliability (Cronbach's alpha = 0.943). The total game activity time is 120 minutes, including game activity flow description (10 minutes), pre-test (20 minutes), game teaching (10 minutes), game play (50 minutes), and post-test and questionnaire completion (30 minutes).

# 3. Results

The Wilcoxon signed-rank test was employed to analyze the difference between the pre and post-tests (Z=-3.734, p<0.001) shown as Table 1. The post-test was significantly higher than the pre-test, indicating that the game enhanced the learners' learning *knowledge of local history*. Table 2 shows the learners' *flow, acceptance, and sense of place*. Overall *flow* (M=3.86, SD=0.53) was significantly higher than 3 (the median of scale) (Z=3.98, p<0.001). The mean values of the *flow antecedent* and *experiences* for each sub-dimension were significantly higher than 3. Acceptance results showed that usefulness, ease of use, and game design elements were all significantly above 3. The sense of place results showed that only recreational activities of sense of place higher than 3. In addition, observations data revealed that 17 out of 21 participants (81%) used the mobile scaffolding during the game.

post-tes	st(N=21)	pre-test	(N=21)		
М	SD	М	SD	Ζ	р
53.1	13.74	34.29	14.6	-3.734***	.000

Table 1 Statistical Analysis of Learning Performance

\*\*\*p< 0.001

Table 2 Flow, Game Acceptance, Sense of Place Descriptive Statistical Analysis

Dimensions	М	SD	Z	р
Overall Flow	3.86	0.53	3.98***	.000
Flow Antecedent	3.86	0.55	3.9***	.000
Challenge-skill balance	3.86	0.71	3.46**	.001
Goals of an activity	4.29	0.58	3.96***	.000
Unambiguous Feedback	4	0.71	3.62***	.000
Sense of Control	4.07	0.97	3.4**	.001

Action-awareness Merging	3.07	0.95	0.19	.852
Flow Experience	3.86	0.65	3.65***	.000
Concentration	3.99	0.83	3.56***	.000
Time distortion	3.54	1.12	2.03*	.042
Autotelic experience	4.19	0.68	3.83***	.000
Loss of self-consciousness	3.26	1	1.27	.203
Overall Acceptance	3.96	0.61	3.95***	.000
Perceived Usefulness	4.05	0.77	3.61***	.000
Perceived Ease-of-use	4.06	0.8	3.66***	.000
Game Design Elements	3.84	0.73	3.56***	.000
Overall Sense of Place	3.25	0.71	1.58	.114
Place attachment	3.13	0.71	0.94	.349
Expressive attachment	3.32	0.87	1.72	.086
Recreational activities	3.52	0.87	2.3*	.022
Tourist satisfaction	3.13	0.89	0.77	.439
Intentional revisit	3.21	0.89	1.44	.15

\*p< 0.05, \*\*p< 0.01, \*\*\*p< 0.001

#### 4. Conclusion

In this study, an educational game "Walking on Puzi Street", which integrates cards and mobile conceptual scaffolds to help students learn local history and culture. The results revealed that the game is helpful to understand of local culture and history, easy to operate and entertaining in terms of sense of place. Finally, the game may help players to understand local literature and history. In the future, design the game to promote sense of place and comparing more learners' learning experiences between traditional tour group are needed.

#### Acknowledgements

This research was supported by the industry-academia cooperation project in NTUST, under Contract numbers 10492-00213.

- Abou-Shouk, M. A., Zoair, N., El-Barbary, M. N., & Hewedi, M. M. (2017). Sense of place relationship with tourist satisfaction and intentional revisit: Evidence from Egypt. *International Journal of Tourism Research*, 20(2), 172–181.
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3), 319–340.
- Executive Yuan. National Strategic Plan for Local Innovation (2019). *No.1070044997* https://reurl.cc/mD4Z9j
- Halvorsen, A.-L., McArthur Harris, L., Doornbos, L., & Missias, M. T. (2021). Lesson study in historical inquiry: Teachers working across rural communities. *Teaching and Teacher Education*, *97*, 103206.
- Hou, H. T.\* & Li, M. C. (2014). Evaluating Multiple Aspects of a Digital Educational Problem-solving-based Adventure Game, *Computers in Human Behavior*, 30, 29-38.
- Kiili, K. (2006). Evaluations of an Experiential Gaming Model. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, 2(2), 187–201.
- Lin, Y.-C., & Hou, H.-T. (2023). Designing a Strategic Analysis and Planning Skills Training Board Game Using Mobile Technology and a Dual-Scaffolding Mechanism.
- McInerney, P., Smyth, J., & Down, B. (2011). "Coming to a place near you?" The politics and possibilities of a critical pedagogy of place-based education. *Asia-Pacific Journal of Teacher Education*, *39*(1), 3–16.
- Tüzün, H., Yılmaz-Soylu, M., Karakuş, T., İnal, Y., & Kızılkaya, G. (2009). The effects of computer games on primary school students' achievement and motivation in geography learning. *Computers & Education*, 52(1), 68–77

# A Preliminary Evaluation of Using Realistic Virtual Space in Designing Gamification Training Activities for Museum Interpreters

Shen-Yang NI<sup>ab</sup>, Shu-Wei LIU<sup>a</sup>, Huei-Tse HOU<sup>a\*</sup>

<sup>a</sup>Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taiwan <sup>b</sup>Exhibition and Education Division, National Museum of Marine Science and Technology

\*hthou@mail.ntust.edu.tw

**Abstract:** In the digital age, the job roles and skills required of modern museum professionals have changed significantly, and many museums have begun to apply digital technology to the training of museum professionals. In this study, we sought to address the shortcomings of traditional training methods by designing a museum gamification training activity to enable learners to develop their museum skills in a highly simulated virtual space. Preliminary results indicate that the gamification of the interpreter training was able to provide learners with flow and learning motivation and that the learners were highly receptive to the training activities and highly appreciated the use of the simulated virtual space to help them learn museum interpretation.

**Keywords:** digital gamification learning, educational game, interpreter training, situated learning

# 1. Introduction

In the digital age, digital transformation is now the key to the sustainable and innovative operation of a museum. In 2020, International Council of Museums(ICOM) states that museums need to rethink their new digital policies and expand their digital exhibition and education services. In terms of the training of interpretation, there are also many shortcomings that need to be improved. González et al. (2016) indicated that the training method led by a single instructor may result in a lack of interaction during the training process. The traditional classroom training method is also prone to problems such as reading from the book, lack of training in content planning, and insufficient manpower (Tasdemir et al., 2014; Nespeca et al., 2023). Tang and Mo (2022) even indicated that the lack of interaction in traditional classrooms is difficult to arouse learners' motivation, whereas virtual learning spaces with gamification could give them more opportunities to repeat operations and experiences, thus achieving indepth and effective learning. *Gather Town* (https://www.gather.town) could be a potential digital platform. It could provide the virtual space designed by the instructor and allow learners to explore and learn independently (McClure & Williams, 2021).

As a consequence, we used *Gather Town* to simulate a marine science and technology museum in Keelung in Taiwan, to create a simulated museum virtual space (Figure 1), and designed as a museum gamification training activity where learners could move their characters around to explore and click on NPCs or exhibit display board to interact with them in the form of *Google Forms* (Figure 2). On entering the game, learners were first given a game mission by the trainer NPC. The learner has to follow the instructions and go to the virtual museum within the allotted time to look for the exhibit display board, use the clues on the boards to complete the explanation script, and at the same time, through the dialogue with the visitor NPC in the virtual museum, learn the details to pay attention to when explaining the interpretation. After completing a 300-word explanation script, learners will receive timely scoring and feedback from two professional museum docents to pass the interpreter training game levels and achieve the higher scores. In addition to increasing motivation, interactive game objects can be used repeatedly to improve interpreter skills, thus achieving the effect of learning transfer that is well-trained before the realistic training.





Figure 1. simulated museum virtual space

Figure 2. Interacting with Virtual NPCs

# 2. Method

The participants in this study were 11 college or graduate students (4 males and 7 females) in Taiwan. We applied the single group post-test design and the learning activity was conducted online. The Kiili Flow Scale (2006) translated and revised by Hou & Li (2014) was referred to in this study. The flow scale includes two dimensions: Flow Antecedents and Flow Experience. All scales were scored according to the Likert scale. The reliability of the flow questionnaire (Cronbach's alpha=0.836) indicated a high degree of internal consistency. In terms of the acceptance of the game by the learner, the technology acceptance scale proposed by Davis (1989) modified by this study was used, which included three dimensions: Perceived Usefulness. Perceived Ease-of-use, and Game Design Elements, and the Likert scale was used. The reliability of the game acceptance scale (Cronbach's alpha=0.723), which has credibility. In terms of the motivation of the learning by the learner, the learning motivation scale proposed by Pintrich (1991) modified by this study was used, which included three dimensions: Intrinsic Goal Orientation. Extrinsic Goal Orientation. and Self-Efficacy for Learning and Performance, and the Likert scale was used. The reliability of the learning motivation scale (Cronbach's alpha=0.910), which has credibility. Prior to the study, all participants were asked to sign an informed consent form. The learning activity had two phases, the first phase was to know the exhibition (120 minutes); the second phase was the museum gamification interpreter training, which included game explanation (10 minutes), game activities (120 minutes), and post-test guestionnaire (20 minutes).

# 3. Results

In this study, the Wilcoxon signed-rank test was used to analyze learners' flow, acceptance, and learning motivation, the results shown as Table 1. The result revealed that the overall flow (M = 4.31) was significantly higher than the median (the median in a five-point scale = 3). This indicates that the overall gamification activity design mechanism enables the learner to clearly understand the game's objective of the activity and is proactively engaged in the game to complete tasks, achieving a high level of flow experience. The result revealed that the overall acceptance (M = 4.64) was significantly higher than the median. It indicates that the learner had a highly accepted acceptance of the game, not only easy to use, but also promoted the learners' learning in museum interpretation. The result revealed that the overall learning motivation (M = 4.16) was significantly higher than the median. This result represents that the museum gamification interpreter training activities are attractive to the learners and can promote the learners' motivation to learn museum interpreting skills.

Dimension	М	SD	Z	р
Overall Flow	4.31	0.37	2.940**	.003
Flow Antecedents	4.40	0.34	2.941**	.003
Flow Experience	4.23	0.43	2.937**	.003
Overall Acceptance	4.64	0.28	2.937**	.003
Perceived Usefulness	4.86	0.32	3.125**	.002
Perceived Ease-of-use	4.58	0.42	2.969**	.003

Table 1. Flow and Acceptance and Learning Motivation Descriptive Statistical Analysis

Game Design Elements	4.49	0.48	2.947**	.003
Overall Learning Motivation	4.16	0.58	2.937**	.003
Intrinsic Goal Orientation	4.06	0.66	2.805**	.005
Extrinsic Goal Orientation	4.04	0.74	2.673**	.008
Self-Efficacy for Learning and Performance	4.52	0.59	2.988**	.003

\*\**p* < 0.01

#### 4. Conclusion

In this study, we used *Gather Town* to create a simulated virtual museum space from which we designed a museum gamification interpreter training activity. Providing learners with the freedom to explore the virtual museum space and complete the task of compiling interpreter script and enhancing their interpreter skills through interactions with training NPC, exhibit display board, and visitor NPC to improve the problem of low interactivity and lack of motivation in the existing interpreter training. Preliminary results indicate that the gamification of the interpreter training was able to provide learners with flow and learning motivation and that the learners were highly receptive to the training activities and highly appreciated the use of the simulated virtual space to help them learn museum interpretation. After the initial analysis of this study, this study will continue to explore the effect of the study on the learners compared to the control group (traditional training) by means of a quasi-experimental design.

#### Acknowledgements

This research was supported by the projects from the Ministry of Science and Technology, Republic of China, under contract number 110-2511-H-011 -004 -MY3, MOST-111-2410-H-011 -004 -MY3.

- Carey, E., Hill, F., Devine, A., & Szűcs, D. (2017). The modified abbreviated math anxiety scale: A valid and reliable instrument for use with children. *Frontiers in psychology*, *8*, 11.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340.
- González-Davies, M., & Enríquez-Raído, V. (2016). Situated learning in translator and interpreter training: bridging research and good practice. *The Interpreter and Translator Trainer, 10*(1), 1-11.
- Hou, H. T. & Li, M. C. (2014). Evaluating multiple aspects of a digital educational problem-solving-based adventure game, *Computers in human behavior, 30*, 29-38.
- ICOM (2020). Report: Museums, museum professionals and COVID-19: follow-up survey. Retrieved from https://icom.museum/wp-content/uploads/2020/11/FINAL-EN\_Follow-up-survey.pdf
- Kiili, K. (2006). Evaluations of an experiential gaming model. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments.*
- McClure, C. D., & Williams, P. N. (2021). Gather. town: An opportunity for self-paced learning in a synchronous, distance-learning environment. *Compass: Journal of Learning and Teaching, 14*(2), 1-19.
- Nespeca, R., Quattrini, R., Ferretti, U., Giotopoulos, K., & Giannoukou, I. (2023). Digital transition strategies and training programs for digital curation of museum. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48*, 1127-1134.
- Pintrich, P. R. (1991). A manual for the use of the motivated strategies for learning questionnaire (MSLQ).
- Tang, J. T., & Mo, D. (2022). The transactional distance in the space of the distance learning under post-pandemic: a case study of a middle school in Northern Taiwan using gather to build an online puzzle-solving activity. *Interactive Learning Environments*, 1-18.
- Tasdemir, A., Kartal, T., & Ozdemir, A. M. (2014). Using science centers and museums for teacher training in Turkey. *The Asia-Pacific Education Researcher*, 23, 61-72.
- Tokila, L. (2021). professional development in the digital era communication skills and competence training programmes for museum staff: the case of finland, paper presented in 2021 ICTOP annual conference. http://ictop.org/wp-content/uploads/2021/01/ICOM-ICTOP-Communication-Oriented-Museums-reduced-final.pdf

# Unmanned Robotic Online Laboratory with an Intelligent Cloud Teacher

Dongkun HAN<sup>a\*</sup>, and Martin Yun-Yee LEUNG<sup>a</sup>

<sup>a</sup>Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China \*dkhan@mae.cuhk.edu.hk

Abstract: One-on-one student-instructor communication is essential in many aspects for science and engineering education. Especially under and post COVID-19, individual student-instructor interactivity becomes increasingly precious yet growingly difficult. To address this issue, this poster paper develops a so-called "cloud teacher", which is an intelligent conversational agent for answering the questions from students in building and manipulating their robots online. The basic underlying idea is to train the agent by utilizing open-source artificial intelligence tools from Google's DeepMind, such that the agent can understand and answer the questions raised by students in making their robots online. Specifically, distinct features of cloud teacher include: 1) Provide instant textual answers to students' questions regarding robot assembly and computer programming (Arduino and SolidWorks: Two most common software in robot control and 3D printing). 2) Supply video demonstrations for helping students in building and manipulating their robots. 3) Automatic assessment of students' online lab performance on controlling their robots. 4) Personalized quizzes based on big data analytics of students' historical questions and performance. Based on the above features, an unmanned robotic online lab has been constructed with 11 sets of real robotic arms, controllers, touchable monitors and corresponding software. The approach of Alpowered conversational agent has demonstrated its effectiveness in 4 engineering courses with more than 300 students. Furthermore, we collaborated with an oversee university, and applied it in the distant robotics education with 50 students from U.K. This method has a high potential to be explored across different disciplines and countries for virtual experiential learning in the new normal.

**Keywords:** Artificial Intelligence in Education; Experiential training; Online teaching; Education in Robotics;

#### 1. Introduction

Experiential learning is indispensable in knowledge and skills acquisition in many subjects. It includes a laboratory hands-on procedure which enable students unique chance for personal and practical manipulation of materials and facilities to acquire and enhance the understanding on concepts, problem analysis-evaluation-solving skills (NSTA, 2007). Experiential learning could be found in many discipline and subjects, like physics, chemistry, medicine, mechanics, robotics, etc. Face-to-face laboratory has been developed along with the procedure of development of science and technology such that the concepts introduced in lectures and books could be enhanced and reinforced by hands-on training (Diwakar, S, 2023). However, one nonnegligible question of face-to-face laboratory is its physical and economic constraints, especially under and post the period of COVID19.

To overcome these constraints, online laboratory is proposed and displayed its vitality especially in last two decades. Online laboratory could be roughly classified into two categories: virtual laboratories and remote ones (Tzafestas, C. S., 2006). Each category has its own advantages: The former tends to supply with less realism but more scalable and easier in maintenance, while the latter is more practical and error-prone. In teaching robotics, we focus on using online laboratory to meet the strong demands of students in operating and manipulating the real robots in practice. The effectiveness of online laboratory

has been shown compared to traditional face-to-face laboratory by evaluating the students experiences in online laboratories in some science courses (Rowe, R. J., 2018). A remotely online accessible swarm robotics research testbed, called Robotarium is proposed for multi-robot research facility (Pickem, D., etc. 2017). A hybrid online laboratory model is provided by merging both categories of online labs and show how it can leverage the advantages of both type of labs (Rodriguez-Gil, L., 2016, and Andujar, J. M., 2010).

This poster paper displays how we developed an unmanned robotic lab (Section 2.2-2.3) and its impact on teaching and learning (Section 2.3).

# 2. The Intelligent Cloud Teacher

# 2.1 Main Development Procedure

The whole process of the project development can be generally divided into four parts: The first part focuses on training the intelligent agent with a database of fundamental knowledge of Mechanical Engineering related to the U.S. Mechanical Engineering Syllabus for Mechanical Engineering Graduates. The second part is concerned with the training of the intelligent agent regards to a database of Arduino programming. The third part aims to train the intelligent agent with a database of VEX hardware assembling. In the last part, a mobile APP will be developed and implemented in an online robotic laboratory (see Figure 1).



Figure 1: Development pipeline of the proposed smart agent

# 2.2 Construction of Unmanned Robotic Online Lab

The basic idea is that we combine the flipped lab with a real online lab, so students can get access and conduct experiments in an online mode. An online robotic lab can be constructed based on our robotic lab with real robots and computers. The conventional robotic lab can be transformed to an online lab thanks to the remote control technique, where students could use their personal computers in distance to control the computers in the lab, further tuning and controlling the robot in real time. On the other hand, this project combines with flipped lab, which allows students to learn from videos of lab instructions before each online lab.

As a result, more time and efforts could be devoted into problem solving and students-instructor interaction in synchronous online lab. students raise their hand by flashing their lights, it is a quite effective way in communication in the lab. So we have extra benefits by avoiding healthy problems and safety problems compared to conventional labs. To build the online lab, we developed different hardware and software components: First, we assembled the robotic arms, controller, drawing panels for students (See Figure 2). In addition, we developed corresponding software, like, Arduino-based programme to control the robotic arm. Most importantly, we installed remote control software, which makes the

programming and controlling of robots possible. Students can use remote control software to manipulate their robots in distance.



(a)

(b)

*Figure 2: (a) One set of remote controllable robotic equipment with sensors and a touchable drawing panel. (b) Ten sets of remote controllable robotic facilities.* 

# 2.3 Impact on Improving the Quality of Education

The monitoring data have been collected by the evaluation methods and illustrates the usefulness of this work:

1) Rate of positive feedback is 93.2% in the Survey on the experience of using the cloud teacher towards the end of the course UGEB2303 Robots in Action.

2) Rate of positive feedback is 82.9% in the survey on the unmanned robotic online lab towards the end of the course UGEB2303 Robots in Action.

3) Rate of positive feedback is 100% in focus group interview with a small group of volunteer students of the course UGEB2303 Robots in Action.

4) Rate of positive feedback is 84.6% on the course website and small group forum. Note that the most remarkable achievement of the proposed cloud teacher is to

provide a new way through which students can get hands-on training and achieve the desired outcomes of lab sessions online. In addition, it goes beyond the traditional face-to-face robotic laboratory and supplies with intelligent chatbot in Arduino programming, real-time automatic assessment of students lab performance and personalized quizzes based on big data analytics of students' historical questions and performance.

# References

Diwakar, S., Kolil, V. K., Francis, S. P., & Achuthan, K. (2023). Intrinsic and extrinsic motivation among students for laboratory courses-Assessing the impact of virtual laboratories. *Computers & Education*, *198*, 104758.

National Science Teachers Association. (2007). NSTA position statement: The integral role of laboratory investigations in science instruction. NSTA Handbook: 2010, 11, 201-204.

Tzafestas, C. S., Palaiologou, N., & Alifragis, M. (2006). Virtual and remote robotic laboratory: Comparative experimental evaluation. *IEEE Transactions on education*, *49*(3), 360-369.

Rowe, R. J., Koban, L., Davidoff, A. J., & Thompson, K. H. (2018). Efficacy of online laboratory science courses. Journal of Formative Design in Learning, 2, 56-67.

Pickem, D., Glotfelter, P., Wang, L., Mote, M., Ames, A., Feron, E., & Egerstedt, M. (2017). The robotarium: A remotely accessible swarm robotics research testbed. In *2017 IEEE International Conference on Robotics and Automation*, 1699-1706.

Rodriguez-Gil, L., Garcia-Zubia, J., Orduña, P., & Lopez-de-Ipina, D. (2016). Towards new multiplatform hybrid online laboratory models. *IEEE Transactions on Learning Technologies*, *10*(3), 318-330.

Andujar, J. M., Mejías, A., & Márquez, M. A. (2010). Augmented reality for the improvement of remote laboratories: an augmented remote laboratory. IEEE transactions on education, 54(3), 492-500.

# Learning Concentration on Virtual Reality Learning: Scale Development and a Pilot Study

Yi Hsuan Wang<sup>a</sup>, Chun-Ping WU<sup>b\*</sup>, Kuan You PAN<sup>b</sup> & Yu HUI CHEN<sup>a</sup>

<sup>a</sup>Department of Educational Technology, Tamkang University, Taiwan <sup>b</sup>Department of Education, National University of Tainan, Taiwan\* \*cpwu303@gmail.com

**Abstract:** The purpose of this study was to develop and validate a multidimensional instrument for elementary and junior high school students' learning concentration on learning using Virtual Reality Technology. Two studies were conducted. The VR Learning Scale (VRLS) was developed based on in-depth literature review and theories followed by a review of eight experts. Data from 259 students ranging from 4th grade to 7th grade with VR learning experience was collected and analyzed using the exploratory factor analysis, the VRLS was validated in four dimensions: immersion, behavior concentration, cognitive concentration, and emotional concentration. A pilot study of 37 elementary school students was conducted and the results indicated the significant correlation between the four factors of VR learning and, learning motivation, technology acceptance and satisfaction. Implication for future studies is provided.

Keywords: learning concentration, VR learning, Virtual Reality

#### 1. Introduction

Virtual Reality (VR), also known as Immersive Virtual Environments (IVE), creates the impression of being in a virtual world and produces sensory stimuli that surround the subject perceptually. The adoption of VR for learning has redefined learning environments. The advancement of VR technology, coupled with a drop in prices and increased affordability, has attracted educators' attention (Bower et al., 2020; Huang et al., 2019), enabling educational institutions to seamlessly integrate it into their teaching practices. The integration of VR devices into learning scenarios has provided students with diverse and engaging experiences. Three primary types of virtual reality (VR) used in educational settings can be delineated: (1) VR for observational learning, (2) VR for skill training, and (3) VR for creative endeavors. When effectively implemented, VR can provide a profound sense of presence and immersion, opening up new educational possibilities (Häfner et al., 2018).

The most recent studies have focused on the adoption of VR to facilitate learning or have discussed the development of VR content alongside learning strategies (Hamilton, et al., 2021; Won et al., 2023). However, there is a gap in the research when it comes to evaluating learner interaction and participation during VR-based learning experiences. Therefore, the primary objective of this study was to develop a VR Learning Scale (VRLS) that assesses the spectrum of learner involvement and engagement, drawing inspiration from Witmer and Singer's (1998) explication of presence, immersion and engagement. The resulting scale was validated through factor analysis. This paper provides an overview of the development process of the VRLS and its subsequent evaluation using Exploratory Factor Analysis (EFA). A pilot study was conducted to further investigate students' VR learning experience.

#### 2. Literature Review

Utilizing VR to support learning provides learners with immersive environments and educational experiences. These experiences have the potential to enhance learner engagement, ultimately leading to increased involvement and self-directed exploration over

the course of the learning process (Won et al., 2023). The effectiveness of virtual environments would be related to the sense of presence, and according to Witmer and Singer (1998), presence is defined as the subjective experience of being in one place or environment, and refers to experiencing the computer-generated environment rather than the actual physical locale. To measure presence, one should address factors that influence involvement and immersion. Involvement depends on the degree to which individuals attach themselves to stimuli, activities, or events. When individuals pay more attention to the stimuli within the virtual environment (VE), their engagement with the VE experience increases (Witmer & Singer, 1998). Immersion refers to interacting with an environment that provides a continuous stream of stimuli and experiences (Witmer & Singer, 1998). A VE that generates a greater sense of immersion will result in higher levels of presence. Both involvement and immersion are necessary for experiencing presence. Involvement

The developed VRLS in this study encompassed two primary factors: immersion and learning concentration. The items assessing immersion were drawn from a Presence Questionnaire (PQ) developed by Witmer et al. (2005), which gauges the extent to which individuals experience presence in a virtual environment. The PQ factors include Involvement, Adaptation/Immersion, Sensory Fidelity, and Interface Quality. Specifically, items from the Adaptation/Immersion factor were incorporated into our VRLS. For learning concentration, the items were designed to capture the flow experience, involvement, and learning engagement. Flow, as defined by Csikszentmihalyi (1975, p. 36), is a 'holistic sensation that people feel when they act with total involvement.' Involvement, according to Witmer and Singer (1998), represents a state resulting from focusing one's mental energy and attention on a coherent set of stimuli or meaningfully related activities or events. Fredericks, Blumenfeld, and Paris (2004) posit that engagement comprises behavioral, emotional, and cognitive dimensions, with engagement playing a pivotal role in successful learning (Henrie, Halverson, & Graham, 2015). A higher level of engagement holds the potential to yield favorable learning outcomes (Lee, 2014). Therefore, items from the behavioral, cognitive, and emotional scales were incorporated into the measure of learning concentration within the VRLS.

#### 3. Results and Conclusion

#### 3.1 Results: Instrument Validity and Reliabilities

The developed VRLS, encompassing two primary factors: immersion and learning concentration. To ensure the content validity, the VRLS was reviewed by eight experts, including three college professors specializing in educational technology, and five experts who are current elementary and junior high school teachers. 259 students ranging from 4th grade to 7th grade with learning experience with VR participated in this study. As to the construct validity, the item-analysis result indicated the CR value of each item was above 3.0 and reached the significant level. An EFA determined the factor structure with KMO values exceeding .90 for each factor. Principal component analysis with Varimax orthogonal rotation extracted the factors. Items with factor loadings below .40 were removed, and the scale structure was made up of four factors, immersion, behavioral concentration, cognitive concentration, and emotional concentration. The Cronbach's alpha ( $\alpha$ ) internal consistency reliability coefficient for each sub-scale was .94, .95, .97 and .96, respectively.

Additionally, this study tried to establish the validity of using the emotional graphics to present participants' emotion status. We invited 214 elementary students to match the six emotional graphics with corresponding text-based description. As shown in Table1, the emotional graphics mostly aligned with the intended emotional states, except the "Expected" item which accounted for 67.8%.

Table 1. Descriptive Results of Emotional Graphics for Presenting Participant Emotion Status

	Нарру	Expected	Curious	indifferent	depressed	scared
Agreed	191	145	193	199	194	206
%	89.3	67.8	90.2	93	90.7	96.3

#### 3.2 Pilot Study

37 students with the learning experience with VR participated in the study. Students' reported scores of the variables, including immersion, behavior concentration, cognitive concentration, emotional concentration, learning motivation, technology acceptance, and learning satisfaction, were listed in Table 2. As shown in Table 3, significant correlations between learning motivation and the four factors of VRLS were found. Similar results were obtained in the relationship between the four factors and technology acceptance. However, the relationship between immersion and learning satisfaction is not significant.

	immersion	behavior concentration	cognitive concentration	emotional concentration	motivation	technology acceptance	Satis- faction	
Μ	5.06	5.12	5.13	5.19	5.39	5.33	5.42	
SD	0.64	0.96	0.82	0.81	0.98	1.06	0.85	

|--|

Table 3. The Correlation among Examined Variables (N=37)

	motivation	technology acceptance	learning satisfaction
immersion	0.49**	0.43**	0.32
behavior concentration	0.40*	0.48**	0.36*
cognitive concentration	0.48**	0.50*	0.40*
emtional concentration	0.52**	0.63**	0.51**

#### 3.3 Conclusion

In this study, the VRLS was developed based on a comprehensive review of the literature and comprised two primary factors: immersion and learning concentration. The scale items were adapted from Witmer et al.'s PQ (2005) and Fredericks, Blumenfeld, and Paris (2004). The VRLS demonstrated content and construct validity, and a pilot study of using the VRLS in comprehending students' learning experiences was carried out. Upcoming findings of this research series will be shared in the coming times.

#### Acknowledgements

The research project is funded by the Ministry of Education in Taiwan of the 5G New Technology Learning Demonstration School Guidance Program- Evaluating the teaching and learning effectiveness of innovative technology integration. We would like to thank all the people who prepared and revised previous versions of this document.

- Bower, Matt & DeWitt, Dorothy & Lai, Jennifer. (2020). Reasons associated with preservice teachers' intention to use immersive virtual reality in education. *British Journal of Educational Technology. 51*. 10.1111/bjet.13009.
- Fredericks, J.A., Blumenfeld, P.C., & Paris, A.H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59-109.
- Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. Journal of Computers in Education, 8, 1–32. https://doi.org/10.1007/s40692-020-00169-2
- Huang KT, Ball C, Francis J, Ratan R, Boumis J, Fordham J (2019) Augmented versus virtual reality in education: an exploratory study examining science knowledge retention when using augmented reality/virtual reality mobile applications. *Cyberpsychol Behavior Social Network, 22*(2):105–110. https:// doi. org/ 10. 1089/ cyber. 2018. 0150
- Won, M., Ungu, D. A. K., Matovu, H., Treagust, D. F., Tsai, C. C., Park, J., Mocerino, M., & Tasker, R. (2023). Diverse approaches to learning with immersive Virtual Reality identified from a systematic review. *Computers and Education*, 195, [104701]. https://doi.org/10.1016/j.compedu.2022.104701

# Methodology for the Participatory Design of a Learner-Facing Analytics Dashboard

Marie-Luce BOURGUET

Queen Mary University of London, UK marie-luce.bourguet@qmul.ac.uk

**Abstract:** In this work-in-progress paper, we describe the participatory design approach we adopted for the inclusive design of a student-facing learning analytics dashboard. The objective is to involve learners throughout the design process to help prevent instructors' biases ending up in the design. The design participants developed learner personas, use scenarios and storyboards to capture the broad spectrum of students' abilities, skills, objectives, and situations. These guided the decisions made for the essential features, functions, and interactive behaviours of the dashboard and for its initial functional implementation.

Keywords: Participatory design, learning analytics, dashboard, personas

# 1. Introduction

Learning Analytics Dashboards (LAD) designed for students are tools capable of monitoring and communicating learning analytics data directly to the learners themselves. These systems can empower students, assisting them in measuring their learning progression, fostering their ability to reflect critically, and enhancing their self-regulated learning competencies (Molenaar, Horvers, Dijkstra, & Baker, 2019). However, despite the rapid pace of these dashboards' development, our comprehension of their effective design, how these tools facilitate or fail to support learning, and the reasons for their general lack of adoption among students remain unclear (Bodily, & Verbert, 2017). The purpose of our research is to create novel designs of LADs that support all learners in the most inclusive manner, and that foster engagement and trust, leading to regular use over time especially by struggling students. In this work-inprogress paper, we describe the participatory design approach we have adopted for the inclusive design of such a dashboard.

Participatory Design (Spinuzzi, 2005) is a user-centric methodology that involves all key stakeholders in the design process, aiming to better anticipate, understand, and fulfill their requirements. This approach encourages stakeholders (especially end-users) to collaborate with designers, researchers, and developers throughout the innovation journey, contributing not only to decision-making but also to the generation of ideas (van Oorschot, Snelders, Kleinsmann, & Buur, 2022). In the context of educational technology, participatory design entails learners, instructors and designers collaborating to develop prototypes and evaluate design ideas. The benefits of participatory design are numerous. Beside preventing personal biases ending up in the instructional design, instructors can gain a deeper understanding of learners' needs, preferences, and experiences. This generally leads to more effective and inclusive solutions. Involving learners in the design process can also help build trust and create a sense of ownership.

# 2. Methodology and Results

#### 2.1 Data

Our project uses learning traces generated by one cohort of around 240 telecommunications engineering undergraduate students enrolled on a course on Multimedia Fundamentals. The course presents a variety of digital learning resources and activities on the Virtual Learning

Environment (VLE), including reading materials, interactive videos (H5P technology), quizzes, and exercise worksheets. Students are expected to complete all the learning activities before attending a synchronous tutorial each week. In the 2022 iteration of the course, each student engaged in a total of 65 digital learning activities. Daily records were maintained of their digital footprints (Bourguet, 2022).

### 2.2 Design Process

The participatory design process we experimented with comprises three phases. The first phase involves an 'initial exploration of the work', aimed at understanding the context of learning. The second phase, referred to as the 'discovery process', is highly interactive and employs methods like brainstorming, persona creation, and storyboarding. The third and final phase involves 'prototyping' the educational resource or technology. Six students from the course were selected to join as uncompensated participants in the project. Alongside a final year undergraduate student, tasked with creating a functional prototype, and the course instructor, they formed the participatory design team. A total of eight participatory design sessions were held (a cumulative 20 hours), with some conducted online and others in person.

During the initial phase, covered in the first two sessions, the goals were to establish through guided discussions a consensus on the participatory design workflow; to foster a shared understanding of the learning context (learning objectives, course expectations, delivery mode, learning activities, and students' digital traces); to raise awareness of ethical concerns related to the adoption of learning analytics; and to discuss principles of inclusive design. Specifically, the data collection, and how digital traces could be used to personalise the learning experience were discussed. Ethical concerns were also addressed about the potential impact of using LADs for students, staff, and the university, as well as the privacy concerns that might arise. The team agreed on a workflow for every meeting, which included identifying and discussing problems, generating and summarising ideas, and creating a meeting summary.

The discovery process spanned the following three sessions with the objectives to generate ideas for the LAD's usage, content, and appearance; to develop learner personas that represent different student profiles; and to craft scenarios and storyboards for each persona. A brainstorming session, resulting in a word cloud of attributes allowed the participants to elicit the characteristics that should be used to describe as broadly as possible their fellow students. Referring to themselves, to their friends, classmates and acquaintances, the project team then endeavored to instantiate these variables, discussed how these traits affected the students' ability to learn, and what LAD features could best support them.

The final stage of the design process comprised the last three sessions aiming to draft various paper prototypes, and to discuss graphical and functional elements of the LAD, like interactivity, customisation, and adaptation (see example in Figure 1.b). The online platform Excalidraw, which allows multiple users drawing on the same whiteboard, was used to support this highly collaborative task.

#### 2.3 Results

For describing fellow students and creating personas, eleven variables were elicited: 'Personal Information' (including Study Location, Quality of equipment, Time availability, English skill, Learning foundation, Ability to use tools), 'Needs', 'Goals and motivation', 'Study strategy', 'Time management', 'Effort', 'Engagement', 'Concentration', 'Efficiency', 'Affective state', and 'Attitude'. Six significantly different student profiles emerged while instantiating these variables, which can be roughly classified as representing: (1) talented, (2) diligent, (3) social, (4) lacking commitment (see Figure 1.a), (5) lacking perseverance, and (6) disengaged students.

A functional LAD prototype was implemented (see Figure 1.c) and a preliminary evaluation questionnaire has been distributed to 31 students (11 identifying as male and 20 as female). The students highly evaluated most of the LAD's graphical components, with a preference for the radar charts and for the progress bars. 17 students (54.84%) indicated that

they felt the LAD could help them monitor their learning progress. 16 students (51.61%) indicated that they were satisfied with the usability and aesthetic of the LAD. Only 14 students (45.16%) indicated that they would be happy to use the LAD regularly. Clearly, more work is needed on the design of the LAD.

The respondents were also asked to identify with one of the six personas. Each persona was selected by at least one student (two of the personas were chosen by 11 and 10 students respectively), which goes some way to demonstrate the inclusivity of our approach.



Figure 1. a. Persona; b. Paper prototype; c. Functional prototype

# 3. Conclusion

There are numerous advantages to employing participatory design in educational technology. In addition to mitigating personal biases in instructional design, it allows instructors to gain a more profound understanding of learners' needs, preferences, and experiences. This results in more effective, inclusive and tailored solutions.

A significant insight from this project is how small, albeit not inclusive, student-staff partnerships can effectively assist instructional designers in better considering the diverse needs, priorities, experiences, behaviours, goals, and situations of learners. It is essential to make a distinction between inclusive partnerships and partnerships for inclusivity; the work outlined in this paper is an example of the latter. The work continues, building on our shared understanding of the students' diversity of needs to build more effective and trustworthy LADs.

- Bodily, R., & Verbert, K. (2017). Review of Research on Student-Facing Learning Analytics Dashboards and Educational Recommender Systems. In: *IEEE Transactions on Learning Technologies*, vol. 10, no. 4, pp. 405-418.
- Bourguet, M.L. (2022). Measuring students' self-regulated learning skills from their digital traces and learning pathways. *EC-TEL 2022 (17th European Conference on Technology Enhanced Learning)*, Toulouse, France.
- Molenaar, I., Horvers, A., Dijkstra, R., & Baker, R. (2019). Designing Dashboards to support learners' Self-Regulated Learning. In: *companion proceedings of the 9th International Conference on Learning Analytics and Knowledge*.
- Spinuzzi, C. (2005). The Methodology of Participatory Design. *Technical Communication* (Washington) 52(2):163-174.
- van Oorschot R., Snelders D., Kleinsmann M., & Buur J. (2022). Participation in design research. *Design Studies*, Volume 78.

# Estimation of Features and Skills of Drawing Experts by Tracing Figures

#### Yasuhisa TAMURA<sup>a</sup>\*, Kazunari KAIZU<sup>a</sup>, Akito HAMANO<sup>b</sup>

<sup>a</sup>Faculty of Science and Engineering, Sophia University, Japan <sup>b</sup>Graduate School of Science and Engineering, Sophia University, Japan \*ytamura@sophia.ac.jp

**Abstract:** This paper focuses on estimating expert-specific characteristics and skills of art drawing, with use of figure tracing log data of pen strokes. In order to verify hypothesis, the authors acquired stroke data of simple figure drawing with use of Tablet PCs and digital pens. Afterwards, we compared the log data of novice learners with one of experts. As a result, we could not find significance between novices and experts in 8 target parameters. However, we found some tendencies of experts. These tendencies imply some training methods in art education.

Keywords: Drawings, Pen stroke, Learning analytics, Multimodal data, Skill estimation

# 1. Introduction

Learning of arts, including painting, pottery, calligraphy, music, and sports is called "skill learning" in general. Unlike learning activity of mathematics and language, identification of the target skill and assessment is rather different in the field of art education and skill learning. Some research challenge in the skill learning field to use digital tools. Singh (2019) discusses the importance of various tools for evaluating different art forms, including drawing. Aboalgasm (2014) focuses on the use of digital art tools for drawing and their potential to enhance artistic ability and creativity among junior school students. The approach to utilize digital tools and acquire objective data is similar to Learning Analytics (LA) approach.

The authors focus on treating "pen stroke data" of figure tracing. In the field of LA research, multimodal LA research treat this level of data granularity. Around this field, some preceding papers present a variety of approaches to assessing learners' drawing skills. Soga (2007) discusses the development of a learning support system for sketching, highlighting the need for learning support systems for skill-based domains like art. Xu (2020) focuses on distinguishing hand drawing styles using a multilevel analytics framework, emphasizing the importance of automatic recommendation tasks for different brushwork skill expressions. Additionally, Yijing (2009) proposes the use of radar charts to assess students' drawing ability, providing a visual and quantitative approach to evaluating drawing skill. These papers highlight the potential of learning analytics approaches in evaluating and supporting art drawing skill development.

This paper aims at drawing skill estimation with use of tracing skill of simple figures like simple line, triangle, circle etc. The authors set the following two hypotheses.

- Experts have unique handwriting characteristics compared to novice?
- Is there a difference in handwriting skill between experts and novice?
- Is it possible to find expert-specific characteristics and skills by analyzing tracing result of simple figures?

To verify these hypotheses, the authors utilized tablet PCs (Apple iPad) and digital pens (Apple Pencil) to obtain pen stroke data. With use of the pen stroke data, it is possible to objectively capture and assess characteristics of figure tracing. In addition, by clarifying the relationship between drawing proficiency and pen stroke skill, it will be a basis for

reconsidering whether practice methods of pen strokes, such as simply continuing to draw straight lines and circles, will lead to improvement in drawing.

In the art education field, many instructors tends to access learners' drawings outcomes with use of their own subjective feelings. Additionally, the assessments are not enough verbalized nor informed. The authors agree the necessity of sensitivity in the art education. However, objective evidence and verbalization of instructors' assessments will bring better agreement and effort of learners.

#### 2. Method

The authors selected the following 13 types of simple shapes were selected: horizontal

straight line (short, medium, long), vertical straight line (short, long), triangle (small, large), circle (small, large), sine wave (small, large), and square wave (small, large).

To measure pen strokes in quantitative manner, the authors developed a JavaScript hook that acquires pen stroke data using iPad and Apple Pencil. iPadOS originally provides the API (application Program Interface) to acquire the data of Apple Pencil. The developed JavaScript hooks the data via the API and



Figure 1: Sample Screenshot of Tracing

transferred the data to a server with asynchronous http-request protocol. The acquired data includes event type (pen down, move, and up), timestamp, X-Y coordinate, pen pressure, and pen tilt.

Each of the 13 figures above is displayed on the iPad screen as shown in Figure 1 (triangle). The subject traces the referenced figure with Apple Pencil. The pen stroke data at this time is obtained as described above. In order to extract the feature of expert learners, the authors calculate 8 parameters below with use of the acquired pen stroke data.

- Average and standard deviation of pen pressure
- Average and standard deviation of pen angle
- Average writing speed
- Average overhang distance
- Number of strokes
- Average stroke length

#### 3. Experiment

The experiments were conducted from November to December 2022. The subjects were 30 in total, including students at Sophia University, Japan (Department of Science and Engineering / Economics / Law, novice of drawing, 15 students) and Tokyo University of the Arts (Department of Arts, expert, 15 students). The subjects were requested to trace 13 shapes described above.

#### 4. Results

The results of the t-test of the 8 parameters in Section 2 are shown below, to compare expert and novice subjects (total 28), excluding 2 subjects of data acquisition error. These result are of all 13 shape types. Unexpectedly, there were no significant difference of the feature values between expert and novice.

- Average pen pressure: t(28)=1.074, p=0.292
- Standard deviation of pen pressure: t(28)=0.785, p=0.439
- Average of pen angle: t(28)=0.887, p=0.383
- Angle standard deviation of pen angle: t(28)=1.229, p=0.229
- Average writing speed: t(28)=0.576, p=0.569
- Average overhang distance: t(28)=0.403, p=0.690
- Number of strokes: t(28)=0.855, p=0.400
- Average stroke length: t(28)=1.222, p=0.232

#### 5. Discussion

The result in Section 4 shows that the t-tests were unable to distinguish expert and novice learners of drawing, especially tracing skill of simple figures. Regardless of the unexpected results above, the authors found some tendencies that were specific in expert learners. First, the average pen pressure tends to be lower for expert learners. It is thought that they are trained the drawing technique of layering thin lines to express shades. The standard deviation of the writing pressure also tends to be lower (the variation in writing pressure is smaller). It is thought that the expert learners draw lines by leveling the pen pressure of the lines. Second, Number of strokes tends to be fewer for expert learners (average stroke length tends to be longer). It suggests that the expert learners are accustomed to draw long lines in stable manner.

On the other hand, surprisingly, we could not find significance of average overhang distance between expert and novice. It shows that the accurate tracing is not the key competency for drawing expert. It implies one of training and assessment policies in art education. On the other hand, in the engineering education field, accurate line drawing in mechanical blueprint and electric circuit is an essential skill. Instructors and learners should understand these field-specific training and assessment policies.

#### 6. Conclusion

The authors tried to extract expert-specific features in drawing activities. As a result, there were no significance between the expert and non-expert on focusing parameters. However, some tendencies were found to characterize the expert skills of drawing.

As a future task, it is necessary to reexamine the appropriateness of the types and sizes of figures used in the experiment. This is because the amount of drawing differs greatly depending on the figure, and it may not be suitable for analysis by integrating pen stroke data for all figures or analysis by comparison between figures.

- Aboalgasm, A., & Ward, R. (2014). Evaluating the Use of Digital Art Tools for Drawing to Enhance Artistic Ability and Improve Digital Skill among Junior School Students. *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, 8(10), 3323-3327.
- Singh, H. R. (2019). Understanding Skill as the tools for evaluation through arts. *International Journal* of English Literature and Social Sciences (IJELS), 4(4), 1173-1175.
- Soga, M., Matsuda, N., Takagi, S., Taki, H., & Yoshimoto, F. (2007). Sketch learning environment based on drawing skill analysis. *11th International Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES)*, 1073-1080.
- Xu, H. (2020). Distinguishing hand drawing style based on multilevel analytics framework. *Wireless Communications and Mobile Computing*, 2020, 1-10.
- Yijing, L., & Ming, L. (2009). Evaluation of drawing ability based on radar chart. *International Conference on Information Technology and Computer Science*, 1, 574-576.

# Logical Expression Tutoring System for Controlling Smart Devices in Multi-User Environments

Tomoya TAKAHASHI <sup>a\*</sup>, Yuko MURAKAMI<sup>ab</sup>, Hidenobu WATANABE<sup>ab</sup> & Kouji NISHIMURA<sup>ab</sup>

<sup>a</sup> Graduate School of Advanced Science and Engineering, Hiroshima University <sup>b</sup> Information Media Center, Hiroshima University

\* m223680@hiroshima-u.ac.jp

**Abstract:** Incorporating robust authentication and authorization mechanisms is important to reduce the risk of credential information leakage in an application such as Internet shopping when a smart device is shared by multiple people. We developed a system that allows smart device administrators to customize the application by registering access control conditions expressed as logical expressions into our system. However, the administrator need logical thinking abilities when creating logical expressions, which is not easy to describe the logical expressions to ensure they work as intended for the specific processing flow because they are not programmers. This paper proposes a tutoring system that helps to express logically using a chatbot, which contributes to enhancement and flexibility for smart device security.

Keywords: Smart Devices, Authentication, Tutoring System, Chatbot

#### 1. Introduction

Smart devices are used in private and public spaces by groups such as families and school classes. Smart speakers, one of the smart devices used in multi-person households, are the subject of privacy and consent studies (Nicole et al., 2021). In multi-user environments, there are challenges such as the risk of storing multiple users' biometric authentication information on the same device. Controlling the device based on its surrounding environment can be achieved by integrating factors such as the number of individuals nearby, the intentions they exhibit, and their interpersonal relationships as external information. We developed a system to control the application of smart devices collaborating with external information to adapt to the environment around the device to solve operational and security issues that arise when multiple people use a single device, as shown in Figure 1 (Takahashi et al., 2022). In this system, the Device Admin (administrator) describes combinations of external information as logical expressions to achieve more complex authentication and authorization than using voice-based verification codes built into smart speakers. Our paper shows a scene in which several students and a teacher with different privileges use a shopping app on a smart speaker in a laboratory. In order to allow students to freely use this app under their own authority, the teacher needs to set the logical expressions as a Device Admin. It is said that the degree of freedom in logical expression design varies greatly depending on the IT skills teachers possess, and the quality of educational content also changes (Karen E. I., 2006). Research is being conducted to assess the logical thinking abilities of Japanese high school students, and it has been found that some test takers do not fully understand the meanings of logical connectors such as "and" and "or" (Misono, et al., 2018). Therefore, a tutoring system is required to allow Device Admin with limited development skills to customize logical expression.

In this paper, we describe a system that performs control collaboration with external information based on a logical expression registered in advance by the Device Admin. Next,

we focus on the part where the Device Admin creates the logical expression and discusses a tutoring system to support it interactively using a chatbot. Even the Device Admin with limited logical thinking ability easily creates logical expressions by interacting with the chatbot.

# 2. Collaboration system with external information

We introduce the collaboration system with external information in Figure 1. This System consists of Process Flow Section, App/User Management Section, and App/User Management DB. App developers create apps and add API calls to the system at the point of customizing the flow. Device Admin registers a logical expression. A logical expression consists of a combination of logical operators and functions to retrieve external information, and the order of processing is determined by the state of the logical expression and external information. User Admin registers user information such as attributes and groups. When a user uses the app on a smart device, it calls the Process Logical Expressions section via API at the customization point. Next, it acquires the pre-registered logical expression from the App/User Management DB. The system processes the process according to the logical expression. If an instruction to the user is required when acquiring data, the contents of the instruction are returned to the application. App/User Management DB.

An example of control using this system implemented in a smart speaker is shown in Figure 2. We envisaged a situation where several students and a teacher use a shopping app on the smart speaker in the classroom. Existing shopping applications use voice-based confirmation codes for access control as shown in Figure 2(a). However, using voice-based confirmation codes is unsuitable in a space with many people, such as a classroom. By registering the logical expression shown in Figure 2(b) into the system, the teacher, as the Device Admin, can introduce processing in the application that utilizes external information such as authentication and relationship data. Even if people are around, secure authentication using ID and password can be performed. In addition, the purchasing authority can be changed according to the user's attributes, such as student or teacher.



Figure 1. The overall structure of the collaboration system with external information



Figure 2. Changes in the shopping app before and after the introduction of this system

#### 3. Proposal for an interactive tutoring system using a chatbot

In our system, the Device Admin is supposed to create a logical expression. Considering that devices are used in schools, the Device Admin would be a teacher, not a programmer. A more intuitive system is required for creating logical expressions.

Therefore, we propose a chatbot to assist in this task interactively, as shown in Figure 3. The chatbot assists in the creation of logical expressions through interactive dialogue, enabling Device Admin to intuitively configure logical conditions by asking questions and presenting choices. Additionally, by engaging in real-time feedback, the chatbot allows the Device Admin to make adjustments based on the situation, thereby preventing incorrect logical expressions. This chatbot helps with the following three points. First, express logical expressions in a natural language style to help Device Admin understand them. Second, by showing the priority of each function, it assists the understanding of the range of influence of "or" and "and". Finally, the logical expression eventually be correct even if the Device Admin adds the conditions in a different order. By supporting the above three points, even if the Device Admin does not understand the structure of logical expressions, the creation of logical expressions can be automated by taking information equivalent to logical expressions in a conversational format.



Figure 3. Overview of the tutoring system using a scenario-type chatbot

#### 4. Conclusion

We develop a system that performs control collaboration to external information based on a logical expression registered in advance by the Device Admin to solve operational and security issues that arise when multiple people use a single device. Chatbots are used to easily create logical expressions even without the administrator being a programmer. In the future, it is possible to introduce a large-scale language model such as ChatGPT to the Chatbot part. Since it is good at explaining logical formulas in natural language, it is possible to help the Device Admin's understanding at the stage of creating logical expressions.

- Nicole M, Dilara K, Kami V,(2021): Owning and Sharing: Privacy Perceptions of Smart Speaker Users. Proceedings of the ACM on Human-Computer, Interact. 5, CSCW1 (2021), Article 45, 1– 29.
- Takahashi, T., Watanabe, H., and Nishimura K. (2022): Development of an External Cooperation System for Enhancing the Authentication Functions of Voice Assistants [in Japanese]. IOT, 2022-IOT-58(6), 1-7.
- Karen E. I. (2006): The Impact of Educational Technology on Student Achievement: Assessment of and for Learning. Science Educator, 15(1), 13-20.
- Misono, T., Takezaki, S., Nishimura, T., Madama, Y.(2018): Development of Test Items for Evaluating Logical Thinking [in Japanese]. JSSE Research Report, 28(7), 49-52.

# Proposing a Processing Distribution System for Cross-Organizational Use of Educational Data

#### Takahiro MORITA<sup>a\*</sup>, Yuko MURAKAMI<sup>ab</sup>, Hidenobu WATANABE<sup>ab</sup> & Kouji NISHIMURA<sup>ab</sup>

<sup>a</sup>Graduate School of Advanced Science and Engineering, Hiroshima University, Japan <sup>b</sup>Information Media Center, Hiroshima University, Japan \*m221924@hiroshima-u.ac.jp

**Abstract:** Cross-organizational data utilization is crucial for fostering DX in education. In order to leverage data across multiple organizations, some research suggests platforms that enable data-using organizations to bring data and the programs deployed to a secure execution environment for processing. However, when it comes to using educational data, it is concerned about the risk of confidential information leakage by transferring the data outside of organization. In addition, to efficiently use data across organizations, each organization should perform common preprocessing on the data. In this paper, we propose a processing distribution system that can securely and efficiently collect educational data and utilize it across organizations.

**Keywords:** DX, Educational data, Cross-organizational use, Processing distribution, GakuNin

#### 1. Introduction

To promote DX (digital transformation) in the field of education, it is necessary not only to analyze educational data held by each organization independently but also to enable cross-organizational analysis. Cross-organizational data utilization can be an opportunity to create new services, in addition to discovering and improving problems in education. Some research suggests platforms that allow the processing of data collected from various organizations in a secure execution environment while ensuring data security and preventing information leakage (NTT, 2022; Sakamoto et al., 2021). These platforms provide an execution environment that keeps data and programs secret from others and enable program execution on that environment. After executing the process, the execution environment is deleted, and the data generated after execution is encrypted and saved in secure storage to prevent information leakage. Organizations that hope to use the data can receive processed results after moving the data and the programs to an environment outside of organization.

When focusing on the utilization of educational data, it is considered difficult from the viewpoint of ensuring privacy to collect the raw educational data held by each organization and analyze it across the organization. Hence, processing platform without transferring the data outside of organization is required. It is also important to be able to verify that educational data is processed appropriately, ensuring transparency in the process.

We propose a processing distribution system that facilitates the cross-organizational utilization of educational data. This system distributes a processing environment for handling data to each organization using container-based virtualization technology. Container packages an application and execution environment together virtually, enables independent processing to run in all computing environments. The processing content in the container is also shared and distributed after prior testing to ensure its validity. This approach allows the organization to access a trusted processing environment and analyze raw educational data without needing to transfer the data outside the organization. Additionally, it enables data standardization and anonymization for cross-organizational analysis.

This paper outlines the requirements for collecting and utilizing educational data, then, introduces the proposed system along with a specific example of its application. Finally, it concludes with a summary.

# 2. Requirements for collection and utilization of educational data

Figure 1 shows the data utilization flow required to realize educational DX. There are two ways to utilize educational data: intra-organizational use and cross-organizational use. Intra-organizational use necessitates a system that enables concise data processing within an organization, without the need to transfer the data externally, and the results are utilized within the same organization (Figure 1-(a)). For cross-organizational use, data is typically anonymized and aggregated, but this can lead to reduced analysis accuracy. Therefore, it is essential to establish a mechanism to anonymize data (Figure 1-(b)) after preprocessing such as data standardization (MEXT, 2021) and analysis within each organization (Figure 1-(a)), followed by aggregation of the results (Figure 1-(c)). This allows the analysis accuracy to be maintained at a high level after data is aggregated. It is also crucial to validate the program before sharing and distributing, ensuring that multiple organizations can agree on how the data will be processed. Additionally, a seamless mechanism is required to connect intra-organizational use to cross-organizational use. By implementing a system equipped with these mechanisms, data utilization can be improved, allowing for better data utilization by refining the program based on knowledge gained from cross-organizational use.



Figure 1. Data utilization flow required to realize educational DX.

# 3. Processing distribution system and a specific example of its utilization

To achieve the data collection and utilization described in Chapter 2, we propose a processing distribution system. Figure 2 shows an overview of the processing distribution system. Since containers are used in this system, the system can be used by installing Docker regardless of the server OS. The system operates on the premise that a group of servers executing the processing is connected through the Academic Access Management Federation in Japan (GakuNin, 2009), allowing container execution and other operations among organizations cooperating through GakuNin. External access from the container execution environment is restricted, and data export in a manner not desired by the data owner is prohibited.

The data processing workflow in this system commences with the container administrator creating a container image containing the application that will process the data. After testing verifies its validity, the container image is registered in the container registry.



Figure 2. Overall view of processing distribution system.

After authentication with GakuNin, data owners then select and download the desired container image (Figure 2-(a)(b)) and execute the processing within their environment. This represents the intra-organizational use. For cross-organizational use, a member of an organization who wants to use the data obtains authorization to run the container for the cooperating organization and gives processing instructions to the system of each organization (Figure 2-(c)). Each organization's system then downloads the required container images according to the instructions. (Figure 2-(b)). The data resulting from the processing performed on each server is subsequently aggregated on the organization's server for further analysis (Figure 2-(d)).

An example application of this system is the utilization of school survey data. For instance, an educational institution intends to conduct a standardized survey across multiple schools. A web application containing the survey content is containerized and distributed to each school. Since the data to be utilized are responses to a questionnaire and the data format is the same, the processing of the data can also be containerized and each school can use the container for intra-organizational use. Moreover, by analyzing the survey results within each school and then performing anonymization and other processing using containers, the survey analysis results can be safely shared. This enables the comprehension and comparison of overall trends.

# 4. Conclusion

We have proposed a processing distribution system for the utilization of educational data across organizations. With this system, essential data can be efficiently processed within the organization's environment, without needing to transfer the data outside of the organization. Depending on differences in the content and format of data held by organizations, it may be difficult to use data only with container processing. However, this system enables uniform data collection and analysis across organizations by defining anonymization policies and data format rules in advance.

# References

- NTT (2022). Next-generation Data Hub for Secure and Convenient Data Utilization across Organizational Boundaries. Retrieved August 17, 2023, from https://www.rd.ntt/e/research/JN202202 17186.html
- Sakamoto, H., Ishida, K., Kato, T., & Inagaki, Y. (2021). "Data value Co-creation Platform" which shares organization's data among other organizations, and enables multiple organizations to increase the value of data [in Japanese]. DC, 2021-DC-120(10), 1-8.
- MEXT (2021). Discussion Paper on the Utilization of Educational Data [in Japanese]. Retrieved August 17, 2023, from https://www.mext.go.jp/content/20210331-mxt\_syoto01-000013887\_1.pdf

GakuNin (2009). About GakuNin. Retrieved August 17, 2023, from https://www.gakunin.jp/en

# Catalyzing Python Learning: Assessing an LLM-based Conversational Agent

Daevesh SINGH<sup>a\*</sup>, Indrayani NISHANE<sup>b</sup>, & Ramkumar RAJENDRAN<sup>c</sup>

<sup>a,b,c</sup> *IDPET, Indian Institute of Technology Bombay, India* \*daeveshsingh@iitb.ac.in

**Abstract:** The rapid rise of digital learning platforms has ushered in an era of educational transformation. While these platforms offer the advantage of scalability, they often fall short in facilitating meaningful interaction, which is pivotal for effective learning. Addressing this concern, our study introduces PyGuru 2.0, an innovative online learning environment for Python programming that aligns with the ICAP framework with an advanced conversational agent. We further investigate the interactions between students and a chatbot, employing a qualitative approach to comprehensively explore the diverse ways in which students interact with the chatbot. The interaction categories encompass a wide spectrum, including code assistance, error resolution, and conceptual explanation. In future, we plan to further elaborate on this coding scheme and see its impact on students' learning outcomes.

Keywords: PyGuru, Chatbot, ICAP, ChatGPT, Python Programming.

#### 1. Introduction

In recent years, there has been a steady growth in the use of digital learning environments to supplement classroom teaching. However, the design of these learning environments challenges educators' capacity to facilitate personalised and meaningful interactions with learners (Song, Oh, and Rice, 2017). Such interactions align with established learning theories that underscore the significance of social engagement and meaningful discourse in yielding fruitful learning outcomes (Kim, 2001).

There is a quest to create technology that can mimic personal and meaningful interactions to improve learning in these digital learning environments. This gap of unavailability of meaningful and personalised learner-educator is now being overcome by Conversational Agents (CA). A CA is a computer system intended to converse with a human. In the domain of education, plenty of initiatives were undertaken to employ CA for various reasons, including the exploration of their potential to foster active participation, drive knowledge construction, and enhance learning outcomes (Song, Oh, and Rice, 2017; VanLehn et al., 2007; Grossman et al., 2019). However, notwithstanding their promising prospects, traditional CAs bear certain limitations. Their knowledge scope can be confined, resulting in limited responses and incomplete guidance. Moreover, their ability to grasp conversation context may be constrained, leading to misinterpretations of learners' queries.

To fill this lacuna, our research endeavours to address these limitations by introducing PyGuru 2.0: an extension of PyGuru that seamlessly integrates an advanced conversational agent powered by "ChatGPT 3.5 turbo". Further, we seek to examine how students interact with this chatbot and how it impacts their learning outcomes. This exploration is grounded in the principles of Chi and Wylie's ICAP framework, which was extended by Raković et al., 2020 by proposing 10 rhetorical moves characteristic of the interactive mode of Chi and Wylie's ICAP framework (2014). An elementary analysis has revealed that students use the chatbot for error resolution support, conceptual explanations etc. In future, we plan to further elaborate on this coding scheme and see its impact on students' learning outcomes.

# 2. PyGuru 2.0: Learning Environment

The learning environment PyGuru 2.0 is a simple extension of PyGuru, which is a computer-based learning environment for learning Python (Singh et al., 2022). PyGuru has a book reader, video player, code editor, and discussion forum. The book reader contains textual information which can be highlighted and annotated. The video player offers an advanced interactive feature of embedding questions into the video. It also has a discussion forum to allow learners to put forth their queries. Further, PyGuru offers Integrated Development Environment (IDE) to practise coding.



Figure 1. User interaction with the chatbot integrated with PyGuru 2.0.

In PyGuru 2.0, we have layered a chatbot on top of the existing learning environment PyGuru. This chatbot is powered by "ChatGPT 3.5 turbo". This chatbot provides students with real-time assistance, guidance, and feedback on their programming inquiries and challenges. PyGuru 2.0 aligns with the ICAP framework by promoting interactive conversations, where learners actively engage in discussions with the conversational agent.

# 3. Methods

# 3.1 ICAP Framework

ICAP is a hierarchical engagement framework classifying student behaviours into four modes: interactive, constructive, active, and passive (Chi & Wylie, 2014). Passive mode involves receiving information without active learning actions, leading to episodic storage (Chi & Wylie, 2014). Active engagement includes physical actions like note-taking and aiding learning with prior knowledge integration. Constructive engagement entails creating artefacts using prior knowledge and available information, fostering deeper understanding (Chi & Wylie, 2014). Interactive engagement happens during interactions between the partners (here chatbot and learner) and requires constructive engagement from both partners involved as well as sufficient turn-taking (Chi & Wylie, 2014). We intend to leverage the ICAP modes to analyse the learner interactions with the chatbot.

# 3.2 Study Design and Data Collection

This research aims to explore and analyse the interactions between students and a chatbot in the context of Python programming education. The study adopts a qualitative approach to investigate the diverse ways in which students engage with the chatbot during their learning process. The data for this study will be collected from PyGuru 2.0. The study involves undergraduate students enrolled in Python programming courses. Data will be collected for six sessions, each lasting 90 to 120 minutes, allowing for a comprehensive analysis of student-chatbot interactions.

# 3.3 Coding Interaction Categories

To analyse student-chatbot interactions, a set of interaction categories will be established based on the nature of communication observed during interactions. Some of the possible categories and their description is mentioned in Table 1.

Table 1. The table presents interaction categories for analysing student-chatbot interactions

Codes	Explanation
Error Resolution	Assistance in resolving programming errors.
Conceptual Explanation	Seeking conceptual explanations or theory.
Exploratory Discussion	Engaging in open-ended discussions to explore concepts

#### 3.4 Data Analysis

Student-chatbot interactions will be categorised by multiple coders to ensure the consistency and reliability of the interaction categories. Inter-rater reliability will be assessed using Cohen's Kappa. The coded interactions will be subjected to qualitative content analysis. The identified interaction categories will be examined in terms of frequency, patterns, and nuances, providing insights into how students engage with the chatbot and how this impacts their learning outcome.

#### 4. Conclusion

By embedding an advanced conversational agent within PyGuru's existing framework, we enhance the quality of interactions between learners and the platform, addressing the limitations of traditional digital learning methods. Our research is rooted in the ICAP framework's principles, advocating for interactive, constructive, and active cognitive engagement to optimise learning outcomes.

The integration of PyGuru 2.0 and its chatbot component aims to provide a comprehensive learning experience, promoting real-time assistance, guidance, and feedback. By employing a qualitative approach to analyse student-chatbot interactions, we seek to unravel the diverse engagement patterns occurring during programming education. This analysis can potentially offer valuable insights into how learners benefit from interactive discussions, error resolution support, and conceptual explanations facilitated by the chatbot.

- Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. Educational psychologist, 49(4), 219-243.
- Grossman, J., Lin, Z., Sheng, H., Wei, J. T. Z., Williams, J. J., & Goel, S. (2019). MathBot: Transforming online resources for learning math into conversational interactions. AAAI 2019 Story-Enabled Intelligence.
- Kim, B. (2001). Social constructivism. Emerging perspectives on learning, teaching, and technology, 1(1), 16.
- Raković, M., Marzouk, Z., Liaqat, A., Winne, P. H., & Nesbit, J. C. (2020). Fine grained analysis of students' online discussion posts. Computers & Education, 157, 103982.
- SINGH, D., SUBRAMANIAM, H., & RAJENDRAN, R. (2022) PyGuru: A Programming Environment to Facilitate Measurement of Cognitive Engagement. In 2022 International Conference on Computers in Education (ICCE). APCSE.
- Song, D., Oh, E. Y., & Rice, M. (2017, July). Interacting with a conversational agent system for educational purposes in online courses. In 2017 10th international conference on human system interactions (HSI) (pp. 78-82). IEEE.
- VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rosé, C. P. (2007). When are tutorial dialogues more effective than reading? Cognitive science, 31(1), 3-62

# How to "Unboxing" Gamer Competencies: via Systematic Literature Reviews to get Analytical Framework is workable?

Tieh-huai CHANG <sup>a\*</sup>, Mingfong JAN<sup>b</sup>

<sup>a</sup>Cross College Elite Tech Program, National Kaohsiung University of Science and Technology, Taiwan

<sup>b</sup> Graduate Institute of Learning and Instruction, National Central University, Taiwan \*ironchang@nkust.edu.tw

Abstract: The long-term lack of structural resources in game studies has been a constant obstacle preventing researchers from academicizing video game community knowledge and gamer competence (Lin, 2011). The difficulty lies in the fact that one cannot make bricks without straw. Each paper on game research is like a piece of brick, without a sufficient amount of bricks, it is naturally difficult to produce a relatively largescale structure. Such a predicament is interrelated and affect one another. Nowadays, a certain number of studies on gamer competence has accumulated in various fields such as neuroscience, communications, psychology, sociology, and pedagogy, however, a coordinate system for gamer competence in the academic circle of game studies has yet to be established. Based on this, this paper uses the document analysis method to propose a two-dimensional analytical framework for classifying current gamers' competencies. The results are presented according to a systematic review, which specifically points out the literature gaps and research method breakthroughs in each quadrant, providing a reference for relevant studies when choosing a research topic or focusing on a research path. To conclude, this paper uses a graphical framework to compile the current state of player competence at this stage, so that subsequent researchers can locate their own point of focus and relative position, making it easier for discussion and focus in their respective area of exploration.

Keywords: analytical framework; systematic review; gamer competence; video game

#### 1. Introduction

What exactly are the abilities that players derive from their gaming experiences? It is still a black box. This central concern of the latest learning sciences has yet to be explained. This dilemma has been taken up by heavyweight researchers in learning sciences, such as John Seely Brown and James Paul Gee, who have called for researchers to inquire into the unknown abilities of the gaming generation (Thomas & Brown, 2011). In recent years, there have been single-point breakthroughs on unknown player competencies, but standalone sporadic findings can rarely be used as structural resources to reverse the dilemmas encountered in the field (Ulicsak & Wright, 2010), and there are still very few systematic "unboxing" of gamer competence to improve the overall understanding.

Based on this, this study will follow up in Chapter 2 with a review of the relevant frameworks for gamer competence to date, seeking to improve the understanding of gamer competence with a systematic review, and subsequently propose an analytical framework and competency distribution in response to the urgency of deconstructing the aforementioned player knowledge and abilities.

#### 2. Literature Review

Analytical frameworks with a focus on community significance can provide key indexing functions. Analytical frameworks are mostly found in review article/ review paper of extensive past literature on a specific topic, summarizing the current research development of the

theme in a graphical framework. Since an analytical framework already contains a large amount of compilation and analysis of past literature, it is extremely helpful for areas where structural resources are scarce, or those who want to get a quick and concise summary, or for researchers who are trying to figure out the direction of their research.

# 2.1 Current Status of the Analytical Framework of Gamer Competence

Structural or review studies of gamer competence are very rare. In the past, many studies have individually explored players' abilities developed through gaming, such as reaction speed, hand-eye coordination, reasoning, spatial exploration, information ability, and social cognition (Crawford, Medendrop, & Marotta, 2004), which are all valuable results that researchers have advanced individually to counter the negative image of "gaming" in the past and their attempt to capture the unknown abilities of players.

The existing analytical framework for gamer competence is rather lacking. There is not yet a mature classification method that can incorporate these unknown abilities into the analysis. After a systematic review, only one framework article, "Key Dimensions of Contemporary Video Game Literacy: Towards A Normative Model of the Competent Digital Gamer" by Klimmt (2009) in the Journal for Computer Game Culture was found.

This three-dimensional analytical framework cannot easily be used to categorize gamers' competencies. Klimmt (2009) proposes three dimensions of player competence, namely (1) resilience against effects of game content on automatic cognition, (2) the ability to cope with social affordances of multiplayer games, and (3) the ability to manage inertia processes in playing motivation that result in a perceived risk of losing investments of time and effort when deciding against playing. These three types of competences are quite meaningful. However, and most critically, these three dimensions cannot be used to categorize existing findings of gamer competence. Therefore, even though this framework paper is the only literature available on player competence, there is still room for improvement if it does not achieve the indexing effect that an analytical framework should have. The research question is: what is the analytical framework of gamer competence obtained through systematic review?

# 3. Research Methodology

The review was conducted in an objective and systematic manner. Document analysis is an effective method to understand the current status of player competence. This chapter is divided into: (1) data collection, (2) data analysis, and (3) establishment the framework.

# 3.1 Data Collection

Four methods were used to carry out literature search: online database, ancestry approach, descendancy approach (Cooper, 1982), and invisible college.

As for Web of Science, the search was based on the Web of Science Core Collection's Social Sciences Citation Index (SSCI), and 445 entries were obtained from Web of Science(

Retrieved by 2021 summer). For the invisible college, four researchers with more than 10 years of experience in game research were consulted and provided supplemental information. A total of 14 additional articles were added by the ancestry approach, descendancy approach and invisible college. After 4 criteria used to filter the literature, a total of 107 articles met the criteria.

# 3.2 Data Analysis

The data analysis is based on the document analysis method. The author carefully read the recorded in detail in order to analyze the competencies of the players mentioned in various papers. Document analysis is a method of systematic research and analysis of document content to generate the context and meaning of the documents' content. The content is

explained and interpreted by extracting the theme, issue, or motif after analyzing the document (Denzin & Lincoln, 2005). For a more detailed categorization, this study adopted a phenomenologically sensitive expert coding for content assessment, called interpretative document analysis. In this paper, three checking methods used in qualitative research were applied to ensure the trustworthiness of the study. First, data triangulation: as mentioned above, three major databases and four data collection methods were used to collect data from multiple sources to help understand the overall phenomenon. This is a means to obtain evidence based on structural corroboration (Ary, Jacobs, & Sorensen, 2010). Second, descriptive validity: it refers to the factual accuracy of the researcher's interpretation in the text, without assembling or distorting what one has seen or heard (Maxwell, 2005). Third, code-recode/intra-rater agreement (Ary et al., 2010; Guba, 1981): Analysis in this paper was initially coded in January 2019 and recoded in January 2021, when comparing the analysis results of the two after the 2-year interval, the intra-rater agreement can be confirmed.

# 3.3 Selection and Establishment of the Analytical Framework

Establish the analytical framework dimensions that can be used to classify the present-day player competence. The researcher carefully reviewed 107 related studies in the Chinese and English literature on gamer competence and used the aforementioned notation method to delete the unused categories. Finally, two important attributes of player competence were identified, which can be effectively used to cover the current research development.

# 4. Research Results

Two important attributes of player competence were identified, which can be effectively used to cover the current research development. These are: the classic dimension of gamer culture "solo and team battle", and the latest achievement in new media literacies, "consuming and prosuming", and the most appropriate representation image was also chosen. The analysis results showed that the X-axis and Y-axis were plotted as continuous axes, and the relative positions of gamer competencies in video games discussed in the literature were placed respectively, and Figure 1 in the following was produced.



The main purpose for the development of an analytical framework is to "use the known to lead to the unknown, uncover the unknown, and gain the latest insights." For this paper, such an analytical framework is easier for readers to understand and use when presenting results, and it helps researchers to analyze and rationalize the data and evidence, which is an efficient way to communicate trends and insights.

# Acknowledgements

NSTC(Taiwan) funded this study under the grant numbered 112-2410-H-992 -001.

#### References

Ary, D., Jacobs, L. C., Sorensen, C., & Razavieh, A. (2010). *Introduction to research in education.* Wadsworth, OH: Cengage Learning.

Crawford, J. D., Medendorp, W. P., & Marotta, J. J. (2004). Spatial transformations for eye-hand coordination. *Journal of Neurophysiology*, *92*(1), 10-19.

Cooper, H. M. (1982). Scientific guidelines for conducting integrative research reviews. *Review of Educational Research*, 52(2), 291-302.

Denzin, N. K., & Lincoln, Y. S. (2011). *The SAGE handbook of qualitative research*. New York: Sage. Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Ectj*, *29*, 75-91.

Lin, H-L. (2011). Between Player and Researcher: Gaming Experience and the Study of Game Culture. *Mass Communication Research, 108*, 19-25.

Klimmt, C. (2009). Key dimensions of contemporary video game literacy: Towards a normative model of the competent digital gamer. Eludamos. *Journal for Computer Game Culture, 3*(1), 23-31.

Ulicsak, M. & Wright, M. (2010). Games in education: Serious Games. Bristol, English: Futurelab.

# A Novel Interpretation of Classical Readability Metrics: Revisiting the Language Model Underpinning the Flesch-Kincaid Index

#### Yo Ehara<sup>a\*</sup>

<sup>a</sup>Faculty of Education, Tokyo Gakugei University, Japan \*ehara@u-gakugei.ac.jp

**Abstract:** In the realm of natural language processing (NLP), the quantification of text readability remains crucial, with pivotal applications in education. While the Flesch-Kincaid GradLevel (FKGL) has been a foundational metric for English text readability, recent advancements, particularly with models like Bidirectional Encoder Representations from Transformers (BERT), have heralded a new age of language model-based assessments. Contrary to popular belief about the FKGL's legacy nature, our research elucidates that FKGL encapsulates language model complexities. We introduce a novel interpretation that views FKGL as a linear blend of perplexities from specific unigram models. Leveraging the OneStopEnglish dataset, we enhanced FKGL by incorporating perplexity values from state-of-the-art language models for sentence boundaries. Our results highlight that integrating BERT's capabilities significantly bolsters FKGL's performance. The implications are vast, suggesting potential expansion to multi-lingual FKGL applications and providing theoretical backing for FKGL-based research in languages like Japanese.

Keywords: Readability Assessment, Language Models, Perplexity

# 1. Introduction

The quantification of text readability stands as a fundamental task in natural language processing (NLP), having pronounced importance in various application domains, notably in education. The Flesch-Kincaid GradLevel (FKGL) [1] has historically been the benchmark metric, widely employed for the automatic assessment of English text readability. In contrast, recent advancements in deep learning, spearheaded by models like Bidirectional Encoder Representations from Transformers (BERT) [2], have ushered in an era where large-scale language models are increasingly anticipated to play significant roles in automated readability evaluation. Numerous studies, as delineated in [3], have put forth methodologies predicated on language models, reportedly outperforming FKGL in precision. A common conjecture attributed to this heightened accuracy is the inherent legacy nature of FKGL, given its non-reliance on any contemporary language models.

Challenging this narrative, our study underscores that FKGL, in fact, encapsulates complexities intrinsic to language models. More specifically, we posit that the FKGL can be aptly represented as a linear combination of the perplexity derived from a specific unigram model. This association between FKGL and perplexity is, to the best of our awareness, unprecedented, and its potential ramifications deem it worthy of comprehensive documentation.

Perplexity, as expounded in [4], acts as a gauge for the complexity of test data within the purview of language models. It is mathematically defined as the inverse of the probability of a word manifesting within the test data, computed on a per-word basis. Intuitively, perplexity provides insight into the average number of choices a language model, predicated on its training, can delineate from contextually when making predictions on the test data. For

illustrative purposes, if the occurrence probability of a word is set at 1/3, it implies the model's capability to narrow its choices down to one out of a potential three words.

# 2. Methods

The Flesch-Kincaid GradLevel (FKGL) is traditionally defined as:

FKGL=0.39×Average Word Count per Sentence+11.8×Average Syllables per Word-15.59 where the average word count per sentence is calculated as the ratio of total words to the number of sentences [1].

Consider a text where sentence boundaries are marked with the token "[SS]". If the text is partitioned such that the last "[SS]" demarcates the test data, with everything preceding it as training data, the output probability of "[SS]" for the unigram model trained on this partitioned data is given by the number of "[SS]" occurrences in the training data divided by the total word count, which equivalently is the ratio of the number of sentences to the total word count. The perplexity for the single-word test data "[SS]" on this unigram model (which is the inverse of its probability) is thus given by Word Count/Sentence Count, aligning with the average word count per sentence.

To elucidate, consider the text segment: "[SS] This is a pen. [SS] That is a cat. [SS]". Using the final "[SS]" as test data and the preceding as training, the unigram model yields p("[SS]") = 12/2. The perplexity of the test data ("[SS]") is 1/p("[SS]") = 12/2 = 6, matching the average word count.

Given that perplexity serves as a quantitative measure of test data complexity using a language model, the average word count in FKGL can be interpreted as representing text difficulty using a unigram model by quantifying the complexity of sentence boundaries. This understanding opens avenues for various applications.

We propose an enhanced FKGL by integrating the perplexity values derived from state-ofthe-art language models for sentence boundaries and examining the potential improvement in performance. For evaluation, we utilized the OneStopEnglish dataset comprising 567 texts, each manually rated on a three-tier readability scale tailored for English learners. (More details can be found at: https://paperswithcode.com/dataset/onestopenglish) Rather than using "[SS]", the BERT model was employed to mask the next word following each text's end, and the probability of the BERT "[SEP]" token appearing was determined. We engaged the pretrained model 'bert-large-cased-whole-word-masking' for this endeavor.

# 3. Results

To evaluate the efficacy of the proposed enhanced FKGL metric, we compared its performance against the traditional FKGL using the OneStopEnglish dataset. Specifically,

the Spearman rank correlation coefficient was calculated against the dataset's manual threepoint scale readability annotation labels: beginner, intermediate, advanced.

The traditional FKGL yielded a correlation coefficient of r=0.5776. In contrast, the enhanced FKGL, wherein the average word count component was supplemented with the BERT-derived sentence-boundary perplexity (scaled by a constant for value range alignment), demonstrated an improved correlation of r=0.5825. The scaling constant was computed as the reciprocal of the average perplexity across all texts.

A Wilcoxon signed-rank test was employed to test the statistical significance of the difference between the two metrics. The difference was found to be statistically significant, with a p-value less than 0.01.

# 4. Discussion

In this study, we demonstrate that the "Average Word Count" component of the Flesch-Kincaid Grade Level (FKGL) can be interpreted as the perplexity derived from a unigram language model. Similarly, the "Average Syllables per Word" can be understood in terms of the perplexity of word boundaries in a unigram language model designed for sequences of syllables. Consequently, the FKGL can be conceptualized as a linear combination of two distinct unigram language models: one considering words as units and the other considering syllables as units. Our findings can be immediately applied to other readability formulae that utilize average word or syllable counts. Notably, these include formulae such as the Flesch Reading Ease (https://en.wikipedia.org/wiki/Flesch%E2%80%93Kincaid\_readability\_tests) and the Automated Readability Index

(https://en.wikipedia.org/wiki/Automated\_readability\_index).

# 5. Conclusions

Our experimental results suggest a noteworthy observation: while the conventional unigram model fails to capture the intricate complexities of sentence boundary determinations, integrating BERT's prowess in detecting sentence boundaries significantly enhances the performance of the FKGL metric. This improvement was determined to be statistically significant.

While this study primarily focused on the average word count in a text, there's potential to further investigate the average syllable count from a similar perspective. Specifically, the intricacies of determining boundaries within sequences of syllables can be viewed analogously to perplexities in the context of a unigram model. Thus, the FKGL can be characterized as a linear combination of the perplexities derived from unigram models. The implications of this research are profound, particularly in lending theoretical validity to studies that employ FKGL for Japanese texts [5]. We anticipate that our findings will pave the way for more extensive research on the multi-lingual adaptation of FKGL and its tailored applications for English learners.

# Acknowledgements

This study was supported by This work was supported by JST ACT-X, Grant Number JPMJAX2006 and JSPS KAKENHI 22K12287. We would like to thank all the people who reviewed this paper.

- [1] Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L. and Chissom, B. S.: Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel, Technical report, Naval Technical Training Command Millington TNResearch Branch (1975).
- [2] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, Proc. of NAACL, pp. 4171–4186 (2019).
- [3] Martinc, M., Pollak, S. and Robnik-Sikonja, M.: Supervised and Unsupervised Neural Approaches to Text Readability, Computational Linguistics, Vol. 47, No. 1, pp. 141–179 (online), (2021).
- [4] Manning, Christopher, and Hinrich Schutze. Foundations of statistical natural language processing. MIT press, 1999.
- [5] Shinya Akagi, Kazuhiro Notomi. Development of Text Evaluation System based on Readability Index for Japanese: A Comparison of Indices and Students' Evaluation. N-007. FIT2016 (in Japanese)

# Experimental study for a computational model in ITS to predict the learners' state

Yoshimasa TAWATSUJI<sup>a\*</sup>, Keiichi MURAMATSU<sup>b</sup> & Tatsunori MATSUI<sup>c</sup>

<sup>a</sup>Graduate school of Engineering, The University of Tokyo, Japan
<sup>b</sup>Global Education Center, Waseda University, Japan
<sup>c</sup>Faculty of Human Sciences, Waseda University, Japan
\*y.tawatsuji@weblab.t.u-tokyo.ac.jp

**Abstract:** Cognitive architecture plays essential role in Intelligent Tutoring System (ITS). For effective learning support, dynamic model subserving the function of the prediction of learners' states should be incorporated. In this paper, we proposed a neural network-based architecture in progress that can incorporate the dynamic model to generate the prediction of leaners' perceptual states when categorical ambiguous stimulus and the instruction of category were given.

Keywords: cognitive architecture, intelligent tutoring system, neural network

# 1. Introduction

Cognitive architecture in the research field of artificial intelligence in education and learning have demonstrated its importance as a role of learner model in Intelligent Tutoring System (ITS) (Anderson et al, 09). As is well known that domain modelling, tutor modelling and learner modelling are essential role of ITS construction (Nkambou et al, 10), sophisticated learner model plays crucial role not only to estimate learner states but also to generate acute "prediction" for the learner states. Estimation of learner states has long been one of the most important issues in ITS. Bayesian approach has been adopted for knowledge-state estimation (e.g. Bayesian Knowledge Tracing), and affective tutors have been proposed that can estimate leaners' mental states (Ammar et al, 10). Estimation of mental states can be achieved with various physiological data (D'Mello et al, 2007), and intelligent mentoring system (IMS) incorporates mental states of learners. Machine learning approach including deep neural network came to realize the acute prediction the learner mental states from physiological data and the estimation results are used for teachers' pedagogical decision making (Matsui et al, 19). On the other hand, "how to predict the learners' states" can remain crucial issues to optimize the pedagogical decision making of ITS. The selection of teaching actions should be optimized with predictions of how the action will affect the state of the learner (e.g. to encourage the learner, a teacher considers how his/her action effects on the learner's mental states). To predict the learners' state and utilize it for the decision making, learner model should incorporate the dynamics of learning-related states including emotion. In this paper, our in-progress neural model incorporating the dynamics of leaner's states as the neural network model is proposed.

# 2. Proposed model for prediction of learner's perceptual states

In this article, we considered the simple situation where learner is provided with categorically ambiguous (including perceptually ambiguous like Rubin's vase) stimuli as educational material and with the perspective as the instruction such as "See this stimulus as the face." Fig.1 shows the protype of proposed neural model for prediction of learners' perceptual states. The model contains of three modules: (1) Gate module, (2) Attractor network module, and (3) Perspective controller module. Gate module is a module that pay attention to a part



*Figure 1.* The proposed neural network model. Blank circles and filled circles indicate the excitatory and inhibitory neurons, respectively.

of input. We construct the attractor network module based on the model of Inferior Temporal cortex proposed by [Matsumoto et al 03]. The output of this module indicates the prediction of learner's perceptual states. Perspective controller module is a key module to realize the perspective-based perception. This module consists of three elements: (1) bias vector and (2) gating modulator and (3) bias modulator. Bias vector represents a perspective that is provided as the pedagogical instruction. Gating modulator is feed-forward neural network that modulate gate pattern and its neural connectivity is updated based on the prediction error from bias modulator. Bias modulator computes a prediction error between the excitatory neurons of attractor network module and bias vector. The whole model is a part of model of learners to incorporate the dynamic states of the learner.

#### 3. Simulation and results

The purpose of the simulations was to test whether output pattern (i.e. prediction) can be generated in line with the bias vector (instruction) when an ambiguous stimulus (created by superposing two stimuli belonging to different categories) input. The input was given as an n-dimensional vector. The pattern  $\xi^{\mu,\nu}$  represents stimuli belonging to category  $\mu$ . According to [Matsumoto et al., 03], 15 stimuli (3 elements × 5 categories) were generated and the elements belonging to same category correlated with each other at 0.3. We adopted  $\xi^{1,3}$  and  $\xi^{4,2}$  as inputs and the connectivity of the gating module was trained before ambiguous stimuli was input. In the simulation, ambiguous stimuli  $H^{ext}$  is given as  $H^{ext} = \alpha \xi^{1,3} + (1 - \alpha)\xi^{4,2}$  ( $\alpha \sim 0.456$ ) and bias vector is given as  $\xi^{1,3}$  or  $\xi^{4,2}$ .

Figure 2 shows the transaction of overlap  $m^{\mu,\nu}(t)$  (the distance between  $\xi^{u,\nu}$  and neural activity of attractor network, ref. (Matsumoto, 05)) and prediction error. *Left-Up* shows the transaction when a bias vector was given as  $v_{bias} = \xi^{1,3}$ , and *Right-Up* shows the transaction when a bias vector was given as  $v_{bias} = \xi^{4,2}$ . The results indicated that in case of  $v_{bias} = \xi^{4,2}$ , the model predicted that the learner's perceptual state converged to  $\xi^{4,2}$  in accordance with the perspective. On the other hand, it predicted the perceptual state converged to failure of learning manner in gate module. In learning phase of the gating module, the stimuli were input from  $\xi^{1,3}$  to  $\xi^{4,2}$  and its connectivity can be adjusted to make it easier to perceive the input as  $\xi^{4,2}$ . This phenomenon should be related to catastrophic forgetting (French 1999). Since learners do not study all learning content at the same time, it is necessary to devise ways to describe the learner model as a neural network model.



*Figure 3.* Transaction of overlap  $m^{\mu,\nu}(t)$  and prediction error  $\varepsilon(t)$ : (Left)  $\xi^{1,3}$  was given as bias vector, (Right)  $\xi^{4,2}$  was given as bias vector.

# 4. Summary and limitation

In this study, we proposed the necessity of prediction module into the cognitive architecture of learner model in ITS. The neural network model that can predict the learner's states of perception was constructed. It also remains important issues other than the refinement of the model: (1) how to generate bias vectors, (2) validity evaluation of the model (e.g. alignment to subject's response in psychological experiments), and (3) the design of interaction between the proposed model and other component of ITS for effective learning support (e.g. adaptive feedback (Yamamoto 23)).

# Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP21K12094.

- 1. Ammar, M. B., et al. (2010). The affective tutoring system. Expert Systems with Applications, 37(4), 3013-3023.
- 2. Anderson, J.R., et al. (2009). Cognitive tutors: Lessons learned, Journal of the Learning Sciences, 4(2), 167-207.
- 3. D'Mello, S., et al. (2007). Toward an affect-sensitive AutoTutor. IEEE Intelligent Systems, 22(4), 53-61.
- Matsui, T., et al. (2019). Conceptualization of IMS that estimates learners' mental states from leaners' physiological information using deep neural network algorithm, In: Coy, A., Hayashi, Y., Chang, M. (eds) Intelligent tutoring system, ITS 2019. Lecture Notes in Computer Science, 11528, Springer:Cham.
- 5. Matsumoto, N. et al. (2003) Neuronal mechanisms encoding global-to-fine information in inferior-temporal cortex, *Journal of Computational Neuroscience*, 18, 85-103.
- 6. Nkambou, R., et al. (2010). What are intelligent tutoring systems, and why this book?, in Nkambou, R., Bourdeau, J. and Mizoguchi, R.(eds) Advances in Intelligent Tutoring Systems, Studies in Computational Intelligence, 308, Springer-Verlag Berlin, Heidelberg
- 7. French, R. M. (1999). Catastrophic forgetting in connectionist networks: Causes, consequences and solutions, Trends in Cognitive Sciences, 3(4), 128-135.
- 8. Yamamoto, S., et al. (2023). In-process feedback by detecting deadlock based on EEG data in exercise of learning by problem-posing and its evaluation, Research and Practice in Technology Enhanced Learning, 18(028).
# Gamified Learning Objects for Inclusive Programming and Science Education

#### Saumay GARG<sup>a</sup>, Seema MITTAL<sup>b</sup> & Mukta GOYAL<sup>a\*</sup>

<sup>a</sup>Department of Computer Science and Information Technology, Jaypee Institute of Information Technology, India <sup>b</sup>Department of Electronics and Communication Engineering, Galgotias University, India \*mukta.goyal@jiit.ac.in

**Abstract:** Children from low-income communities face additional barriers to science and computing education, including limited computer access, language barriers, and a lack of resources for experimentation, often while being out of school. Computer-based learning resources that address these challenges in their design have the potential to make STEM education accessible to a whole new section of society. This paper presents a gamified learning object and a supporting visual programming language (VPL) that allows users to program and execute science experiments in the form of a narrative-driven problem-solving exercise. This tool is primarily intended for children from economically disadvantaged backgrounds.

Keywords: STEM education, Introductory programming, Social-emotional learning

### 1. Introduction

Gamified computer applications have been extensively explored for enhancing the quality of instruction and engagement of students in introductory programming and science education. Unfortunately, children from underserved communities face unique barriers to accessing science, technology, engineering, and mathematics (STEM) education and are often excluded from such initiatives. Children from low-income families in developing countries may not attend school and get forced into physical labor, street vending, or begging, which leaves no time for learning. And even if they do attend school, they often lack proper equipment and competent faculty which prevents them from meaningfully engaging in STEM education interventions. In absence of government support, individual instruction via computing devices is an option, but is limited by access to network and devices. Furthermore, unstable economic conditions and social stigma may lead to persistent stress which hinders the development of a healthy social-emotional skill-set.

In light of the above limitations, we present a gamified inquiry-driven learning object to teach science concepts, basic coding, and social-emotional learning (SEL). Combining these goals into one app reduces intervention duration. Science inquiry games help lay the foundations for scientific thinking, and SEL is emphasized because it helps improve student performance and emotional well-being (McClelland et al, 2017). Also, block-based programming languages like Scratch have been shown to help young children develop problem-solving and computational thinking skills with minimal cognitive load (Resnick et al, 2009; Pelánek et al, 2022). We address the following research gaps in this paper: (1) Though a game unifying SEL and Physics education goals was proposed by Aleven et al (2013), tools that teach science, SEL, and programming skills through a singular interface don't exist; (2) Programming platforms like CODE.org (Kalelioğlu et al, 2015) focus on small puzzles that exist as one-off activities that are not adaptable to the local curriculum and lack the scalability to encompass much of the existing STEM curriculum; and (3) Educational interfaces for introductory programming rarely consider the unique challenges of children from low-income communities in their design.

# 2. Learning Object: Design and Implementation

The learning object includes a toolbar containing the activity blocks, a workspace and an SEL support toolbar; see Figure 1a for the outline. It can also be replicated offline with minimal resources and supports the following objectives:

**Support for scientific inquiry:** Developing scientific inquiry skills requires repetitive experimentation and real-time feedback. The activity's design is guided by the National Science Education Standards (Aleven et al, 2015) and encourages children to predict, test hypotheses, explain observations, and revise theories based on feedback from the interface.

**Support for block programming:** Blocks in the activity support the text-free design paradigm with supporting text in a local language (Hindi). The use of icons helps make the exercise more accessible to children who lack adequate literacy skills and extend its accessibility for younger students who might rely solely on visual cues. Some blocks, as shown in Table 1, use parameters to reduce ambiguity by limiting options to a reasonable set.

**Support for social-emotional learning:** The activity targets three primary SEL goals for elementary students: cooperation and empathy, self-awareness, and help-seeking behavior. It involves a human player and a non-player character (NPC); the player must respect the emotions of the NPC in order to successfully complete the game and earn maximum points. The NPC would express discomfort when under undesirable predicament and leave the game leading to a two-point deduction from the user score. This serves to reinforce the need for cooperation and showcasing empathy towards the NPC. Based on this understanding, the child would be prompted with a short quiz for labeling emotions of self and the NPC (Figure 1b). Additionally, help-seeking behavior is encouraged through tooltip (dialogue box appears on the hover event on each block), a help button (plays a short video explaining the activity setup), and trashcan feature that maintains a history of blocks used for reference during gameplay.



Figure 1. a) Layout of the proposed interface b) Self-reflection form for SEL support.

The game is set in a playground with a swing and an electromagnetic setup inspired by Gupta (2015). The NPC sits on an iron swing which moves due to electromagnetic forces according to the Fleming's left-hand rule. The player manipulates the swing by altering elements in the virtual setup (like polarity and orientation of magnetic field and electric current). One end of the wires remains connected to the swing while the other can be periodically attached to the battery, this induces a magnetic field around the swing. A magnet beneath it interacts with this field, rocking the swing so the NPC can collect apples from the trees. Being swung too high upsets the NPC which prompts them to leave the game. Therefore, the player would have to balance self-interest and regard for the NPC's emotions in order to gain maximum points. Players learn the Fleming's left-hand rule by continually developing and refining hypotheses about the virtual environment through experimentation as part of the scientific enquiry process. The blocks are designed to teach some basic concepts of programming like sequential execution, functions with parameters (*rotate, increase* and *connect* blocks), and loops. It also illustrates semantic errors and type compatibility, e.g., a battery can't substitute for a wire in the connect block. See Table 1 for a detailed description of blocks in the VPL.

We propose the deployment of this activity as a web application so it could be accessed on a wide range of devices and operating systems without the need to install any new software. Additionally, web applications can be conveniently packaged for offline usage. Our implementation used the React library while the graphical interface was made up of Scalable Vector Graphics (SVG) with Cascading Style Sheet (CSS) transitions for animations. The VPL and coding workspace were developed using the Blockly library.

Block Image	Block	Description	Block Image	Block	Description
	Name			Name	_
े वैटरी	Battery	Block representing the battery object.	घुमाएँ <b>0</b>	Rotate (battery)	Takes the battery object as argument and rotates it with the angle provided by the player. The values accepted are 90, 180 and 270 degrees.
जोहो 📑 M M	Connect	Used to connect wires to battery, the arguments for the connect block are type sensitive (wires can only occupy the extreme positions).		Increase (value)	Increases the current or magnetic field of the object attached. Only allows the battery or magnet blocks to be connected.
काला तार	Black Wire	Block representing the wire (black) object. Used to describe the direction of the current when connected to a battery object using the connect block.	लाल तार	Red Wire	Block representing the wire (red) object.
् चुंबक <mark>N</mark> S	Magnet	Block representing the magnet object. It can be passed to the 'rotate' or 'increase' block as an argument.	डिंस्कनेक्ट 🛞 📘	Disconnect	The current flow stops when the wires and battery are disconnected, returning the swing to its initial position.

#### Table 1. Custom blocks in the visual programming language

# 3. Conclusion and Future Work

This paper presented design considerations, implementation and deployment suggestions for a novel interdisciplinary educational activity supported by a small visual programming language. Though the presented approach could aid in making STEM education more accessible to children from low-income families, there are several questions that need further investigation. The proposed approach must be tested with participants in order to verify its efficacy; questions about cognitive load, perceptions formed, and complexity of code generated should be evaluated. Furthermore, it is essential to fully explore the implications of such intervention on the community's perception of formal schooling and ethics of child employment. Future research could also explore unsupervised joint media engagement (JME) and the potential of parental collaboration for temporarily replacing actual teachers during times of crisis when access to trained educators might be hampered.

# References

Aleven, V., Dow, S., Christel, M., Stevens, S., Rosé, C., Koedinger, K., & Zhang, X. (2013). Supporting social-emotional development in collaborative inquiry games for K-3 science learning. In Proceedings of the 9th Games+ Learning+ Society Conference-GLS (Vol. 9).

Gupta, A. (2015). Wonders from Waste. Scholastic.

- Kalelioğlu, F. (2015). A new way of teaching programming skills to K-12 students: Code. org. *Computers in Human Behavior*, 52, 200-210.
- McClelland, M. M., Tominey, S. L., Schmitt, S. A., & Duncan, R. (2017). SEL interventions in early childhood. *The Future of Children*, 33-47.
- Pelánek, R., & Effenberger, T. (2022). Design and analysis of microworlds and puzzles for block-based programming. *Computer Science Education*, 32(1), 66-104.
- Perez-Escolar, M., & Canet, F. (2023). Research on vulnerable people and digital inclusion: toward a consolidated taxonomical framework. *Universal Access in the Information Society*, 22(3), 1059-1072.
- Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., ... & Kafai, Y. (2009). Scratch: programming for all. *Communications of the ACM*, 52(11), 60-67.

# Exploring the Possibility of Harnessing Drones in Geography Education in High Schools

Morris Siu-Yung JONG<sup>a\*</sup>, Chin-Chung TSAI<sup>b</sup>, & De-Qi ZHANG<sup>a</sup> <sup>a</sup>Department of Curriculum and Instruction & Centre for Learning Sciences and Technologies, The Chinese University of Hong Kong, Hong Kong, China <sup>b</sup>Graduate Institute of Information and Computer Education, National Taiwan Normal University, Taiwan \*mjong@cuhk.edu.hk

**Abstract:** This working paper discusses our initiative to explore the possibility of harnessing drones as an inquiry-based learning tool in geography education for supporting outdoor fieldwork. In particular, we focus on delineating (i) the rationale behind this initiative, and (ii) the direction of how to adopt drones in geography learning and teaching based on the high-school context in Hong Kong.

**Keywords:** Drones, drone-assisted learning (DAL), inquiry-based learning, geography education, outdoor fieldwork, high schools

# 1. Drone-Assisted Learning (DAL)

Drone-assisted learning (DAL) has been regarded as one of the potential technologyenhanced learning approaches to be widely adopted in the field of education (Shadiev et al., 2022). Nonetheless, in K-12 contexts, most empirical studies and representative examples of DAL so far are mainly pertaining to STEM education or STEM-related subjects, in particular, with the purpose of supporting the learning and teaching activities of computational thinking, computer literacy and programming skills (e.g., Bhuyan et al., 2020; Breuch et al., 2020; Chou, 2018; Fokides et al., 2017). There has been a lack of DAL research on exploring the possibility of leveraging drones to support high-school students in conducting outdoor inquiry-based fieldwork in geography education.

# 2. High-School Geography Curriculum in Hong Kong

In alignment with the educational reform for promoting student-centredness in school education in Hong Kong, the revamped statutory high-school geography curriculum lays strong emphasis on theme-based and inquiry-based learning (Curriculum Development Council, 2022; Jong, 2023; Jong et al., 2019; 2020a; 2020b). In the curriculum, each geographical theme involves a number of curricular topics; the content of each topic pivots around a central inquiry question. Figure 1 illustrates an example of a curricular topic and inquiry question under the geographical theme of "Dynamic Earth."

# 3. Proposed Pedagogical Use of Drones in Outdoor Inquiry-Based Fieldwork

In Hong Kong, typically, a high-school geography teacher uses a 15-day teaching cycle to cover a curricular topic (see Figure 1), in which five lessons (70 minutes each) are evenly distributed in each cycle. Based on this learning and teaching context, we have proposed a pedagogy for facilitating students to harness drones to conduct outdoor inquiry-based

fieldwork. The design of this pedagogy is theoretically grounded on Pedaste et al.'s (2015) five-stage inquiry-based learning model, namely *Communication*, *Engagement*, *Exploration*, *Expression* and *Evaluation*. Figure 2 illustrates the implementation of this pedagogy within a 15-day teaching cycle.



Days **Description (T:Teacher; Ss: Students)** Day I (in-lesson communication): T connects Ss to the topic's background, introduce them to drones as an aerial observational data collection tool, and discuss with them the ethical/ safety issues and technical restrictions of drones with respect to the regulation in Hong Kong. Days 2-3 (off-lesson engagement): Ss (i) access short video materials posted by T on the Days I-3 LMS to gain a more comprehensive picture of the topic's background and deepen their understanding of ethical/ safety issues and technical restrictions of drones, and (ii) conduct beginning "virtual" practice on the operation of the drone to be used later by experimenting with the corresponding flying simulation application. Day 4 (in-lesson communication): T recaps the important points underpinned to the videos that Ss have watched in Days 2–3. After that, T demonstrates how to operate the drone and make use of its photo- and video-capturing functions to collect aerial observational data. Days 4-6 Then, Ss are divided into groups. Each group is given a drone to get real hands-on practice. Days 5-6 (off-lesson engagement): Ss continue to use the simulation application for conducting some advanced "virtual" practice on drone flying and aerial observational data collection. Day 7 (in-lesson communication): T discusses the topic's core inquiry question with Ss, equips them with the related key geological and geomorphological concepts, and guides them to break down the core question into a number of sub-questions in terms of the concepts. Days 7-9 Days 8-9 (off-lesson exploration): T leads Ss to the fieldwork site to conduct the aerial inquiry-based outdoor fieldwork. The sub-questions formulated on Day 7 will scaffold each group to probe into the phenomena by collecting useful aerial observational data through photo- and video-capturing. All groups' inquiry proceedings are recorded in their drones. Day 10 (in-lesson communication): Each group discusses the data collected in the fieldwork and co-constructs explanatory answers for responding to the sub-questions. During the discussion, they can access the inquiry proceedings from the provided tablets to support Days 10-12 and triangulate their answers. Days 11-12 (off-lesson expression): Each group prepares a presentation (with a visual aid) to elaborate and explain their findings (answers) to all the sub-questions with the arguments collectively developed on Day 10. Day 13 (in-lesson communication): Each group gives the presentation and receives . feedback from other groups. After that, T debriefs the class on (i) the strengths/ weaknesses of

Figure 1. A curricular topic and inquiry question under a geographical theme.

their findings, (ii) what knowledge they have learned and how the knowledge can be used when studying other topics, and (iii) the importance of observing the ethical and safety measures

Figure 2. Harnessing drones in outdoor inquiry-based fieldwork.

<sup>Days 13–15 when conducting drone-aimed outdoor fieldwork.
Days 14–15 (off-lesson evaluation): Each group reflects on their previous "answers" to the sub-questions based on the new insights obtained on Day 13, updates their presentation file with the refined answers and ethical/ safety reflection on using the drone as a learning tool, and finally uploads it to the LMS for other groups' reference.</sup> 

## 4. Conclusion and Future Research

Drones have been foreseen as a potential tool to be widely adopted for supporting K-16 learning and teaching (Shadiev et al., 2022). Inquiry-based learning has been regarded as an important instructional approach to facilitating constructivist, student-centred education and training (Huang et al., 2022; Pedaste et al., 2015). This working paper presents our initiative to develop a DAL pedagogy for supporting outdoor inquiry-based fieldwork in the context of Hong Kong high-school geography education. A critical piece of our coming work is to conduct empirical research to evaluate the effectiveness of this pedagogy. Moreover, it is important to collaborate with frontline geography teachers to enact design-based research (DBR) (Jong et al. 2023; Mckenney et al., 2019) for optimising the design and implementation of this pedagogy so that it can be pragmatically adopted in different school settings in the circumstance of formal education.

#### Acknowledgement

The work described in this paper was substantially supported by a grant from the Research Grants Council of Hong Kong Special Administration Region, China (Project No.: 14618322).

# References

- Bhuyan, J., Wu, F., Thomas, C., Koong, K., Hur, J. W., & Wang, C. H. (2020). Aerial drone: An effective tool to teach information technology and cybersecurity through project based learning to minority high school students in the U.S. *TechTrends*, *64*, 899–910.
- Breuch, B., & Fislake, M. (2020). First steps in teaching robotics with drones. In M. Merdan, W. Lepuschitz, G. Koppensteiner, R. Balogh & D. Obdržálek (Eds.), *Robotics in education* (pp. 138–144). Springer.
- Chou, P. N. (2018). Smart technology for sustainable curriculum: Using drones to support young students' learning. *Sustainability, 10*, Article 3819.
- Curriculum Development Council. (2022). *Geography: Curriculum and assessment guide (Secondary* 4–6). HKSAR Education Bureau.
- Fokides, E., Papadakis, D., & Kourtis-Kazoullis, V. (2017). To drone or not to drone? Results of a pilot study in primary school settings. *Journal of Computers in Education*, *4*(3), 339–353.
- Jong, M. S. Y. (2023). Pedagogical adoption of SVVR in formal education: Design-based research on the development of teacher-facilitated tactics for supporting immersive and interactive virtual inquiry fieldwork-based learning. *Computers & Education*, 207, 104921.
- Jong, M. S. Y., Tsai, C. C., Xie, H., Wong, F., Tam, V., Zhou, X. (2019). Exploring the possibility of leveraging spherical video-based immersive virtual reality in secondary geography education. *Proceedings of the 27<sup>th</sup> International Conference on Computers in Education (ICCE 2019)* (Vol. II, pp. 709–711). Kenting, Taiwan.
- Jong, M. S. Y., Tsai, C. C., Xie, H., & Wong, F. K. K. (2020a). Integrating interactive learner-immersed video-based virtual reality into learning and teaching of physical geography. *British Journal of Educational Technology*, *51*(6), 2063–2078.
- Jong, M. S. Y., Tsai, C. C., Xie, H., Wong, F. K. K., Tam, V., Zhou, X., & Chen, M. (2020b). Problems in educational use of spherical video-based immersive virtual reality in practice: The LIVIE experience. *Proceedings of the 28<sup>th</sup> International Conference on Computers in Education (ICCE 2020)* (Vol. II, pp. 724–727). Darwin, Australia.
- Huang, H., Hwang, G. J., & Jong, M. S. Y. (2022). Technological solutions for promoting employees' knowledge levels and practical skills: An SVVR-based blended learning approach for professional training. *Computers & Education*, 189, 104593.

Mckenney, S., & Reeves, T. C. (2019). Conducting educational design research (2nd ed.). Routledge.

- Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, 14, 47–56.
- Shadiev, R., & Yi, S. (2022). A systematic review of UAV applications to education. *Interactive Learning Environments*. <u>https://doi.org/10.1080/10494820.2022.2028858</u>

# Impact of Augmented Reality App on EFL Young Learners' Vocabulary Learning Engagement in a Seamless Learning Environment

Yanjie SONG<sup>a\*</sup>, Jianfeng ZHOU<sup>b</sup>, Yin YANG<sup>c</sup>

<sup>ab</sup>Department of Mathematics and Information Technology, The Education University of Hong Kong, China <sup>c</sup>Artificial Intelligence and Digital Competency Education Centre, The Education University of Hong Kong, China <u>\*ysong@eduhk.hk</u>

Abstract: Despite that Augmented Reality (AR) has been integrated into English as a Foreign Language (EFL) instruction, few studies have been conducted, focusing on AR -supported vocabulary learning both inside and outside the classroom at a primary level. Against this background, this guasi-experimental study aimed to examine the impact of an AR app - VocabGo on primary students' vocabulary learning engagement under the 'pedagogical framework of AR-supported vocabulary acquisition in a seamless learning environment' in the context of a private school in Shenzhen, China. Mixed research methods were adopted. The study lasted 22 weeks. Seventy-two Grade 4 students from a private school in Shenzhen were randomly divided into three groups (N = 24 per group). Group 1 used VocabGo both in-class and out-of-class, Group 2 used VocabGo in-class only, and Group 3 used VocabGo out-of-class only. Data collection involved student pre- and post-engagement guestionnaires, student focus group interviews, and log data. The research findings show that students in Group 1 outperformed significantly those in Group 2 and Group 3 in learning engagement in cognitive, behavioural, emotional and agentic dimensions. This indicates that AR-supported vocabulary acquisition in a seamless learning environment is conducive to students' vocabulary learning engagement in all dimensions. This study contributes to the literature in substantiating the effectiveness of AR-supported vocabulary learning in a seamless learning environment for young learners in an EFL context, particularly in Mainland China.

**Keywords:** Augmented reality, English as a Foreign Language (EFL), vocabulary learning, engagement

# 1. Introduction

Vocabulary knowledge plays a crucial role in successful language comprehension and use (Nation & Webb, 2011). English as a Foreign Language (EFL) vocabulary learning across different contexts in a seamless learning environment has gained substantial recognition. A 'seamless learning' refers to a learning process that happens across various contexts, including formal and informal settings, virtual and physical spaces and individual and social spaces (Wong & Looi, 2021). Recent studies on augmented reality (AR)-supported English as a second language (ESL) vocabulary learning in a seamless learning environment have shown much promise in enhancing learners' motivation, cognitive engagement and outcomes (Wen, 2021; Wu et al., 2021). An investigation conducted by Wu et al. (2021) found that AR-assisted vocabulary learning activities were perceived as more enjoyable and stimulating by Chinese EFL learners, leading to higher engagement and intrinsic motivation levels. Wen (2021) reported that students in the experimental group outperformed those in the control group in

cognitive engagement due to longer time spent on getting used to AR-supported ESL language learning.

Huang et al. (2021) in their review study on AR and VR applications in language education found that the studies have focused more on learning motivation that have led to increased learning outcomes and less on engagement and satisfaction, which is echoed by the most recent review on AR-supported language learning by Shadiev and Liang (2023). In addition, other challenges and concerns about AR-supported vocabulary learning remain such as the majority of studies having been conducted in classrooms or laboratories, the learning activities having been prescribed by the researchers and teachers where learners have had limited opportunities to consolidate and use the newly learned words outside the classroom leveraged by AR technologies, existing studies having been generally conducted over a short term, rarer studies having been conducted in EFL contexts, like China.

In view of these issues, this study aimed to investigate the impact of an AR app - VocabGo AR on EFL pupils' vocabulary learning engagement both inside and outside the classroom in a primary school in Mainland China. Next section presents the research design and the design of the VocabGo app.

# 2. This study

#### 2.1 Design of the VocabGo app

The AR app named VocabGo was newly developed by the first author and her research team to enhance ESL young learners' vocabulary learning across different settings in a seamless learning environment. Premised on Mayer (1997)'s generative theory of vocabulary learning with technology, dual-coding theory (Paivio, 2014), second language acquisition (Nation, 2006) and the concept of seamless learning (Wong & Looi, 2011), the design of the AR app focuses on enhancing learners' vocabulary learning using AR identified real-world objects/picture with triggered vocabulary in authentic learning environments across different settings.

VocabGo app runs on mobile devices with iOS and Android systems. It consists of four modes: (1) 'Find' mode; (2) 'Go' mode; (3) 'Explore' mode; and (4) 'Challenge' mode (see Song et al., 2023 for detailed explanations). The teacher can define the words within the curriculum unit and input them into VocabGo before students' learning.

Considering the integration of AR into EFL young learners' vocabulary learning in China is still in its infancy, in this study, students only allowed to use 'Find' mode and 'My Collection' both inside and outside the classroom. In 'Find' mode, students can scan the real objects with identified newly learned English words included in the current curriculum unit; and in 'My Collection', the scanned objects using 'Find' mode can be automatically saved in the category of pictures with newly learned words (picture-word cards) in the current learning unit.

#### 2.2 Pedagogical framework

The pedagogical design aims to examine learners' engagement in vocabulary learning both inside and outside the classroom supported by the AR app – VocabGo in primary schools in Hong Kong (refer to Figure 1). The design consists of activities where the VocabGo app was adopted to support both in and out-of-class activities. To be specific, in-class activities aim to engage learners in the first three stages of vocabulary acquisition (a) encounter a new word; (b) get its form; and (c) understand its meaning with the form supported by 'Find' mode on VocabGo. Out-of-class activities in real-life settings aim to involve learners in consolidating the word, and using the word using 'Find' mode, and 'My Collection' on the VocabGo app.

#### 2.3 Quasi-experimental design

This quasi-experimental study adopted mixed research methods were adopted, lasting 22 weeks. Seventy-two Grade 4 students from a primary school in Mainland China were randomly

divided into three groups (N = 24 per group). Group 1 used VocabGo both in-class and outof-class, Group 2 used VocabGo in-class only, and Group 3 used VocabGo out-of-class only. All groups of students use the same study time in and out of class no matter whether they used the VocabGo or not. Teachers and parents monitored their study time. Data collection involved student pre- and post-engagement questionnaires on behavioural, cognitive, emotional and agentic dimensions with 18 items on a 5-point Likert scale (Zainuddin et al., 2020), student focus group interviews, and log data (number of picture-word cards). The research procedure is presented in Figure 2.



Figure 1. Pedagogical framework.

Figure 2. Research procedure.

Both qualitative and quantitative data analysis methods were adopted.

# 3. Results

The post-questionnaire results showed that there were significant differences in engagement levels across all dimensions between Group 1 and Groups 2 and 3 with Group 1 having the highest engagement level. The results also showed that there was a significant difference between Group 2 and Group 3 with Group 2 having a higher engagement level.

The focus group interview data analysis results indicate that students in Group 1 reported their most enjoyable and engaging experiences than those in the other two groups.

In terms of the number of picture-word cards collected in 20 weeks, Group 1 collected 6392, Group 2 collected 5172 and Group 3 collected 4861 picture-word cards.

# 4. Discussion

The findings of this study indicate that AR-supported vocabulary learning in a seamless learning environment is conducive to young learners' learning engagement in a EFL context such as Mainland China, which contributes to the literature. Future research will focus on the impact of the AR app – VocabGo on students' vocabulary learning outcomes, and the relationship between learning engagement and outcomes to inform curriculum design.

#### **Reference** (more references will be included in the presentation due to space limitation)

Zainuddin, Z., Shujahat, M., Haruna, H., & Chu, S. K. W. (2020). The role of gamified e-quizzes on student learning and engagement: An interactive gamification solution for a formative assessment system. *Computers & Education*, *145*, 103729.

# Optimization of Personalized Content Providing Function for Adult Learners with Diverse Backgrounds

Chikako NAGAOKA<sup>a\*</sup>, Masako FURUKAWA<sup>a\*</sup>, Yuan SUN<sup>a\*</sup>, & Kazutsuna YAMAJI<sup>a\*</sup> <sup>a</sup>National Institute of Informatics (NII), Japan \*nagaoka@nii.ac.jp

**Abstract:** The skills required in relation to the same topic can differ depending on learner's attributes such as job and preference. For example, when training on research data management for staffs in universities and research institutes, the skills needed differ depending on the learner's jobs, such as researchers, research administrators, and IT technologists. In the past research, we developed a Moodle plug-in that provides learning contents and quizzes optimized for each learner's job and preference dynamically. In this research, we further enhance this plug-in to register skills retrieved from the skill repository and support learners selecting learning contents based on their preference.

**Keywords:** Skill Development, Skill Repository, Moodle Plug-in, LMS, Personalized learning, Research Data Management

#### 1. Introduction

In recent years, there have been an increasing number of opportunities for working adults to relearn new skills through reskilling and recurrent education programs (MEXT, 2022) in Japan. According to the theory of andragogy by Knowles (2015), adult learners generally become ready to learn when their life situation creates a need to know (readiness to learn). Therefore, the skills required for adult learners may vary significantly because different learners have different life situations such as jobs in the workplace even if they need to learn the same topic. However, it is difficult to prepare all the courses according to learners' jobs. Furthermore, even if courses on all jobs are developed, if a single learner studies courses on multiple jobs at the same time, time will be wasted due to duplication of learn "Research Data Management (RDM)" for faculty staffs in universities and research institutes in Japan, but RDM learners have a variety of occupations and required skills, and learning content needs to be provided in a flexible manner.

To avoid the creation of various learning courses and duplication of learning contents, it is desirable for learning content to be dynamically organized and provided to learners according to the skills required for the learner's jobs and preference. With Moodle, the learning management system, it is possible to display learning activities only to specific learners by using a feature called restrict access (Moodle, n.d.). However, the lack of the concept of skills reduces flexibility and teachers need to connect learners and learning activities manually.

# 2. Basic Function of Personalized Content Providing Function Developed in The Past Study

In the past study (Nagaoka et al., 2023), we developed a function that links skills, quizzes to check skill acquisition, and learning content that enables users to learn the knowledge necessary to pass the quizzes. Furthermore, we then developed a function that

dynamically generates personalized page which shows quizzes and learning contents according to the learner's job associated with the learner (Figure 1 and 2).

The personalized content providing function developed in the past research (Nagaoka et al., 2023) consists of three main functions: (1) skill registration function that links all possible skills that may be presented to learners with learning content and quizzes for learning those skills, (2) user attribute registration function that links learners with the job, and (3) user content providing function that displays only the learning content and quizzes related to the skills needed by the learner in a tabular format.







*Figure 2.* Automatically generated page by personalized content providing function (Nagaoka et al., 2023).

# 3. Additional Functions of Personalized Content Providing Function Developed in The Present Study to Use in RDM Context.

Our past research (Nagaoka et al., 2023) has established the basic function for providing learners with optimal learning content. However, there were two features that were missing to use this system in RDM: (1) the function for learners to select skills by themselves, and (2) the function to register skill data retrieved from skill repositories.

#### 3.1 The function for learners to select skills by themselves.

On the system which we developed in the past study, only Moodle administrators (LMS administrators) can link learner's attributes and skills. However, when using this system for RDM learning, there was a possibility that learners need to learn skills other than those

assigned to their own jobs depending on their own interests and circumstances such as department movement. Especially in smaller organizations, the range of skills that a single staff should learn considered to be wide and flexibility is required to choose skills. Therefore, we developed a function that allows learners to select the skills which they want to learn.

#### 3.2 The function to register skill data retrieved from skill repositories.

In the past study, Moodle administrators register skills one by one, manually. However, in recent years, a system for sharing skills has been developed. For example, 1EdTech (n.d.a), which provides technical standards for education-related systems, offers the Competencies and Academic Standards Exchange (CASE) as a standard for skill sharing systems. Based on this standard, the IMS Case Provider (1EdTech, n.d.b), a repository for sharing skills, has been published. In this research, we developed a function that allows Moodle administrators to register skills to Moodle with JSON data format retrieved from the WEKO3 repository system (Research Center for Open Science and Data Platform [RCOS], n.d.), which used as the skill repository to share and update skills in this research. By using this function, we can register and update skills on the repository and register latest skill data to Moodle.

#### 4. Future Works

This paper detailed the characteristics and additional functions of the content providing system. One future work of this study is the consideration of how to integrate this system with the skills repository. Since skills on the skill repository can be constantly updated, it is necessary to consider how to reflect the updated skills in the Moodle side. For example, we can use or develop an API to obtain information, so that the skill repository side and the Moodle side always synchronize skills. In addition, it is necessary to consider whether we should customize the skill repository system to adjust standards such as 1EdTech's CASE.

# Acknowledgements

This research was supported by MEXT as "Developing a Research Data Ecosystem for the Promotion of Data-Driven Science".

# References

Knowles, M. S., Holton, E. F., & Swanson. R. A. (2015) The adult learner: The definitive classic in adult education and human resource development (8th ed.). Oxon: Routlege.

MEXT (2022). Ministry of Education, Culture, Sports, Science and Technology's initiatives to promote recurrent education. Retrieved from https://www.mhlw.go.jp/content/11801000/000983561.pdf [in Japanese]

Moodle (n.d.a). Restrict Access Settings. Retrieved September 14, 2023, from https://docs.moodle.org/402/en/Restrict access settings

- Nagaoka, C., Furukawa, M., Sun, Y., & Yamaji, K. (2023). Development of content presentation functionality for learning essential skills based on the learner's position. Information Processing Society of Japan (IPSJ): Collaboration and Learning Environment, 39(9), 1-4.
- Research Center for Open Science and Data Platform (n.d.). WEKO3 (Publishing Platform). Retrieved August 18, 2023, from https://rcos.nii.ac.jp/en/service/weko3/

1EdTech (n.d.a). Competencies and Academic Standards Exchange (CASE). Retrieved August 18, 2023, from https://www.imsglobal.org/activity/case/

1EdTeCH (n.d.b). CASE Network Frameworks. Retrieved August 18, 2023, from https://casenetwork.imsglobal.org/vsalt/index.html#/

# SATELUC

# Promotion of ERD Design Comprehension Using Recomposing Method

#### Putra Prima ARHANDI <sup>a\*</sup>, Banni Satria ANDOKO <sup>b</sup>, Tsukasa Hirashima <sup>c</sup>

<sup>a</sup> Information Technology Department, Politeknik Negeri Malang, Indonesia
 <sup>b</sup> Information Technology Department, Politeknik Negeri Malang, Indonesia
 <sup>c</sup> Informatics and Data Science Program, Hiroshima University, Japan
 \*putraprima@polinema.ac.id

**Abstract:** Conceptual modelling is a critical ability for a computer science student, enabling them to create a variety of conceptual models such as Entity Relationship Diagrams (ERDs). However, learning to create an ERD is a challenging task especially for novices because it involves various cognitive processes and actions. This research proposes a recomposing approach using kit-build exercise for ERD design learning. In a kit-build exercise, a student is provided a set of components of a structure and asked to recompose the structure using these components. The kit-build exercise has advantages as (1) reducing student's cognitive load by providing available components and (2) realization of automatic assessment by comparing recomposed structure with the original structure. This exercise has been already implemented for concept maps and the learning effect has also been confirmed both experimentally and practically. Considering the structural similarities between ERDs and concept maps, we have applied the kit-build exercise approach to ERDs. In this paper, we describe the kit-build exercise for ERD with automatic assessment and feedback. **Keywords:** ERD, Kit Build, Concept Map.

#### 1. Introduction

In software engineering, conceptual modelling is an important skill. It involves using a modelling language to reconstruct statements about certain domains, such as data or process modelling. This activity is required in computer science, software engineering, and related majors, as it is mandated by curricula standards (ACM and AIS).

Creating a conceptual model, such as an entity-relationship diagram, involves various cognitive processes and actions such as goal setting, abstracting, conceptualizing, contextualizing, interpreting, evaluating, anticipating, visualizing, communicating, discussing, and agreeing. Therefore, creating a conceptual model is considered a complex task that requires understanding of theoretical foundations, modelling languages, methods, and tools, as well as critical thinking and reflection on the application domain. Codified and tacit knowledge are also required to complete this task successfully (Rosenthal et al., 2022). Several study reports that novice student having difficulties and create errors when designing an ERD. Fessakis reported that students' face difficulties when deciding entities and attributes in an ERD (Fessakis et al., 2005). Rashkovits and Lavy mentioned that identifying relationships between entity is a serious challenge for novice designer (Rashkovits & Lavy, 2020). The difficulties when designing relationships between entities in ERD also explored by Batra and Antony their study revealed that novice students' common errors were incorrect relations between entities (Batra & Antony, 1994). Batra and Davis observed that ERD created by novice students' were different with the ERD created by experts in terms of ternary relationships, unary relationships and categories (Batra, n.d.). Furthermore, Batra mentioned that designing ERD is an activity with high complexity for novice students since the activity involve extensive use of memory, no mental aids and unconstrained choices (Batra, 2007).

Kit Build Concept Map (KB Map) is a framework for concept map where a concept map created by students can be diagnosed programmatically, KB Map uses the recomposing approach(Hirashima, 2019). In this research proposal we propose kit-build exercise for ERD. We chose Kit Build Concept Map (KB Map) because KB Map have advantages over normal concept map. KB Map able to reduce complexity by providing available components for student to recompose. KB Map also provide ability to automatically assess the students' answers with teachers' solution by comparing the students' map with goal map.

In KB Map a domain expert or lecturer creates a goal map, then this goal map is decomposed to create a kit (parts). These parts are provided to learners and the learners are required to recompose their own map (recomposing) based on provided kits. This map is called learners' map. Kit build utilizes the goal map and learners map to detect defects in the learners' map. Several learners' maps can be overlayed to create a group map. This group map can be compared with a goal map to re-examine the goal map. Since the ERD have similarities and difference with KB map in terms of concept and link between concept, the extending process will involve several adjustments to fit an ERD to KB map.

# 2. Extending Kit Build for ERD Learning

#### 2.1 Subject for Experiment

The subject for this research is sophomore student in Politeknik Negeri Malang.

#### 2.2 Steps to Extend Kit Build for ERD Learning

To extend kit build for ERD learning there are several differences between ERD and Concept Map, the first difference is in concept map there are two items concept and link meanwhile in ERD the concept can be an entity or attribute. To fit entity and attribute to kit build we will convert entity and attribute to a concept in kit build. The second difference is link between concept where the connection between entity in ERD can be divided into three types such as "One To One Relationship", "One To Many Relationships", and "Many To Many Relationships". To fit the relation to kit build we will convert the relation into link in kit build this relation include link between entity and attribute.



Figure 1 Change from ERD to Concept Map

To demonstrate the process the following example questions for ERD learning / problems is used: "Create a ERD for a simple blogging system, in this system a post consists of a title, post body, and category. To create a post a user must assign the post category, one category can have multiple blog post, each category consists of category name."



Figure 2 Example of ERD answer and ERD converted to goal map

A lecturer will create the goal map by converting ERD (Figure 2 A) answer into a goal map. The process of converting ERD to goal map following the rules in Figure 1, the result of this process can be seen in Figure 2B. Then for the learning purpose the goal map in Figure 2C is decomposed into parts of goal map. From this goal map students will be given the questions text and must create their own version of the concept map. This map called student map.



Figure 3 Variation of a Student Map

In the learning process the students will get a questions text and a decomposed goal map Figure 2C and they will recreate their own map (Student Map) example of a possible student map shown in Figure 3. There are several variations of a student map can be recomposed the first is a student map with "leaving link" Figure 3 A, this happen when student leaving a link on their map. The second variation is student map with "lacking link" this link happens when student connecting a concept to another concept but put wrong link on one of the concepts. The third variation is when student link two concepts with completely wrong link.



For analysis purpose in kit build student maps can be overlayed on top of goal map to find difference between student answers and teachers answer. Several student map also can be overlayed to get the insight of what students think about the answers in kit build this map called a group map Figure 4A. In the group map the similar answers between student represented by a bold link between concepts, meanwhile the lacking link (Figure 4B) and excessive link (Figure 4C) can be obtained by comparing group map with goal map. Both lacking and excessive link are the links that many learners failed to connect correctly.

#### References

- Batra, D. (n.d.). Conceptual data modetling in database design: Simlaritii and differences between expert and novice designers.
- Batra, D. (2007). Cognitive complexity in data modeling: Causes and recommendations. Requirements Engineering, 12(4), 231–244. https://doi.org/10.1007/s00766-006-0040-y
- Batra, D., & Antony, S. R. (1994). Novice errors in conceptual database design. European Journal of Information Systems, 3(1), 57–69. https://doi.org/10.1057/ejis.1994.7
- Fessakis, G., Dimitracopoulou, A., & Komis, V. (2005). Improving database design teaching in secondary education: Action research implementation for documentation of didactic requirements and strategies. Computers in Human Behavior, 21(2), 159–194. https://doi.org/10.1016/j.chb.2004.06.006
- Hirashima, T. (2019). Reconstructional concept map: Automatic Assessment and reciprocal reconstruction. International Journal of Innovation, 5(5).
- Rashkovits, R., & Lavy, I. (2020). Students' Difficulties in Identifying the Use of Ternary Relationships in Data Modeling: International Journal of Information and Communication Technology Education, 16(2), 47–58. https://doi.org/10.4018/IJICTE.2020040104
- Rosenthal, K., Strecker, S., & Snoeck, M. (2022). Modeling difficulties in creating conceptual data models: Multimodal studies on individual modeling processes. Software and Systems Modeling. https://doi.org/10.1007/s10270-022-01051-8

# Designing Educational Personas using Generative Al

#### Ivan TERZIC<sup>a</sup>, Antun DROBNJAK<sup>a</sup> & Ivica BOTICKI<sup>a\*</sup>

<sup>a</sup>Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia \*ivica.boticki@fer.hr

**Abstract:** In this paper, the use of generative AI for creating educational personas is presented. Prior to the appearance of generative AI, conversational AI was used to enhance educational environments with the use of technologies such as chatbots. Generative AI brings a new dimension as it allows for the creation and use of personas mimicking real people, historical figures, and even fictional characters. Such designs potentially bring a new immersive and interactive dimension to learning, but further research is required to understand the full potential of these technologies. In this paper, the experiences and challenges in appropriating language learning models to act as potential educational personas are described.

Keywords: Generative AI, learning, education, personas, conversational AI

### 1. Introduction

The emergence of generative AI has highlighted the potential of AI in education, leading to the efforts to its widespread adoption worldwide (Baidoo-Anu & Owusu Ansah, 2023). Generative AI has played a pivotal role in driving the creation of interactive tools and systems that are enabling learners to access personalized and dynamic learning experiences. Ultimately, these interactive systems possess the capability to revolutionize the way learners engage with educational content and access learning resources. However, akin to other novel tools, there are persisting challenges concerning their use in achieving enhanced learning outcomes and improved educational process. Although generative AI holds the potential to contribute to the improved interaction in the educational processes that could potentially lead to educational benefits (Xu et al., 2022), it is important to acknowledge that pedagogical approaches to achieve so are still work in progress.

# 2. From Conversational AI to Personas in Education

Terzidou et al. (2016) found that the implementation of effective pedagogical approaches can yield positive impact on students' learning outcomes. Xu et al. (2022) suggest that facilitating contingent interaction between child viewers and media characters has the potential to bring about supplementary educational advantages that are not readily available through conventional video programming. Pataranutaporn et al. (2021), Fu et al. (2022), Almahri et al. (2021), Park et al. (2022) and Chien & Yao (2020) found that personalization can improve students' motivation, engagement, and outcomes. Haller & Rebedea (2013) presented a method for building a conversational agent that has personality and knowledge about historical figures and can be used in educational contexts. Likewise, Fu et al. (2022) highlighted that agents' personal memory in conversation can significantly improve agents' responses. Xu et al. (2022) proposed an enhanced framework for designing pedagogical agents, which can pose questions, provide feedback, are based on scaffolding, and dynamically adjust to students. Implementing such an innovative framework could significantly enhance the efficacy of pedagogical agents in supporting student learning.

# 3. Creating Personas and Conversing with Personas

To explore the potential for the use personas in generative AI, Open AI and GPT-3 model were leveraged to create a prototype where personas could be defined and used. The prototype had the advantage of own database for storing user information and logs. Figure 1 shows a feature where predefined personas can by modified by specifying name, desired voice, avatar, and the initial prompt thereby determining persona characteristics. Such a design allows for adjusting the persona vocabulary level towards the educational level of students taking part in the learning activity.

Persona picker:	
Albert Einstein 🗸 🗸	
Name:	
Albert Einstein	
Voice:	
Hans, German, Sample <b>q</b>	
Image:	
71 ~	
Initial prompt:	
Provide scientifically accurate answers to questions from the perspective of Albert Einstein. Use scientific language and expressions. The questions will be	
provided by the user in the following messages. You must answer exclusively from Einstein's imaginary	•
Prompt length: 326, max length is 1000	a
Submit Reset	

# Persona variation creation

Figure 1. Modifying a predefined persona to match the desired educational level

Conversations with the persona are conducted via the interactive feature where a specific persona can be chosen for textual and voice communication (speech-to-text and text-to-speech) when interacting with the persona. The avatar of the persona is customized and animated to reflect real time talk (Figure 2).

Select a chat option:	
Text Interactive Conversation	
Select a persona: Albert Einstein Submit Reset conversation	
Ask Albert Einstein something	Send message
Or	
Listening	1

Figure 2. Interactive mode for conversing with the persona via text or voice

#### 4. Conclusions

To achieve sensible and smooth conversation with the persona, prompts, or queries to OpenAI are used to set up the persona. Such a query needs to include the detailed specification of the expected persona characteristics. Due to the design of Open AI, the following needs to be eliminated from the conversation: the answers where the model itself generates questions for the person it is imitating, answers where the model mentions that it is not a persona but that it is a language model, and responses where the entire conversation is generated as a conversation between two people. The randomization of answers and the reduction of the occurrence of repeating expressions had to be performed as well. In terms of the overall interaction, the roles of the participants in the conversation sent to OpenAI needed to be precisely defined. To simulate the synchronous nature of the conversation, speech and animation needed to be synchronized, pronunciation of numbers had to be defined, and the flow of conversation had to be slowed down for easier readability.

The development of generative AI has open the path for significant advancements in pedagogy, interaction, and personalization in education. By incorporating pedagogical principles into AI systems, educators can leverage generative AI as effective teaching tools. Interaction with AI systems, such as chatbots, humanoid robots, and LLMs, should enable engaging and immersive learning experiences. The adaptability of generative AI models supports personalized learning by potentially tailoring instruction to individual student needs.

### Acknowledgements

We would like to thank Dr. Peter Seow and Dr. Ken Kahn who prepared and revised previous versions of this document.

#### References

- Almahri, F. A. A. J., Bell, D., & Arzoky, M. (2021). Applications of machine learning in education: Personas design for chatbots. In *Machine Learning Approaches for Improvising Modern Learning Systems*. https://doi.org/10.4018/978-1-7998-5009-0.ch004
- Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4337484
- Chien, Y. H., & Yao, C. K. (2020). Development of an ai userbot for engineering design education using an intent and flow combined framework. *Applied Sciences (Switzerland)*, 10(22). <u>https://doi.org/10.3390/app10227970</u>
- Fu, T., Zhao, X., Tao, C., Wen, J. R., & Yan, R. (2022). There Are a Thousand Hamlets in a Thousand People's Eyes: Enhancing Knowledge-grounded Dialogue with Personal Memory. *Proceedings of* the Annual Meeting of the Association for Computational Linguistics, 1. https://doi.org/10.18653/v1/2022.acl-long.270
- Haller, E., & Rebedea, T. (2013). Designing a chat-bot that simulates an historical figure. *Proceedings* - 19th International Conference on Control Systems and Computer Science, CSCS 2013. https://doi.org/10.1109/CSCS.2013.85
- Park, C., Jang, Y., Lee, S., Park, S., & Lim, H. (2022). FREETALKY: Don't Be Afraid! Conversations Made Easier by a Humanoid Robot using Persona-based Dialogue. 2022 Language Resources and Evaluation Conference, LREC 2022.
- Pataranutaporn, P., Danry, V., Leong, J., Punpongsanon, P., Novy, D., Maes, P., & Sra, M. (2021). Algenerated characters for supporting personalized learning and well-being. In *Nature Machine Intelligence* (Vol. 3, Issue 12). https://doi.org/10.1038/s42256-021-00417-9
- Terzidou, T., Tsiatsos, T., Miliou, C., & Sourvinou, A. (2016). Agent Supported Serious Game Environment. *IEEE Transactions on Learning Technologies*, 9(3). https://doi.org/10.1109/TLT.2016.2521649
- Xu, Y., Vigil, V., Bustamante, A. S., & Warschauer, M. (2022). "Elinor's Talking to Me!":Integrating Conversational AI into Children's Narrative Science Programming. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3491102.3502050

# Timorese University Students' Perception on E-learning: A Case Study

Agostinho Dos Santos GONÇALVES<sup>a\*</sup>, Sebastião PEREIRA<sup>b</sup>, Saida ULFA<sup>c</sup>

<sup>a,b</sup>Faculty of Education, Instituto Superior Cristal, Timor Leste <sup>c</sup>Faculty of Education, Universitas Negeri Malang, Indonesia \*dossantosgoncalvesagostinho@gmail.com

**Abstract:** For the sake of maintaining academic continuity during the Covid19 pandemic that struck the world, particularly in East Timor, all students were required to adopt the distance learning format. Online learning, which allows flexibility in learning, is one of these formats. This study aims to describe how Timorese University students perceive online learning. 190 college students from Instituto Superior Cristal (ISC), East Timor, participated in this study. According to the study's findings, students generally have a favorable perception of online learning.

**Keywords:** behavioral intention to use e-learning, face-to-face learning, online learning, student's perception

### 1. Introduction

Setting up distance learning to serve all children, especially the hardest-to-reach children in remote or rural was not an easy undertaking for a tiny country like East Timor with a population of 1.3 million, and distance learning had never been undertaken at scale [UNICEF, 2020]. Not everyone has access to technology but suprisingly In East Timor, 109.7% of the population had a mobile connection in 2021, indicating that many individuals had several mobile phones, but the country's internet penetration rate was just 45.1% [Kemp, 2021]. The statistics does not yet demonstrate a fair allocation of community access to technology. As a result, UNICEF assisted the government in continuing to educate all children during the COVID-19 epidemic by offering learning opportunities through a variety of media, including TV, radio, the internet, mobile phones, and printed materials. Although practically everyone owns a cell phone, not everyone has access to technology [UNICEF, 2020].

Instituto Superior Cristal (ISC), one of East Timor's institutions, hosted a teaching and learning process utilizing a variety of online platforms for the first time during the Covid-19 pandemic, including zoom meetings, Google class rooms, Google Meet, social media YouTube, and WhatsApp. However, due to issues with the erratic internet network, exorbitant prices, and a lack of supporting infrastructure, the educational process was not carried out successfully and efficiently.

After Covid-19, East Timor's educational system now relies more on face-to-face instruction. However some professors still choose to employ online learning methods. The learning experience of using online platforms left a good effect on students and instructors despite the inadequate current infrastructure. Therefore, the goal of this study was to ascertain how East Timor students felt about the usage of online learning at an institute of higher education.

# 2. E-Learning in East Timor

Although East Timor's ICT infrastructure has advanced significantly since gaining independence in 2002, there are still many obstacles to overcome, including the fact that less than 1% of the country's population has access to fixed-broadband services and that both mobile and fixed-broadband services are regarded as being out of reach for most people [ESCAP, 2019]. When the Covid-19 pandemic hit the world, East Timor was one of the countries that were struggling due to the state of its ICT infrastructure. As a result, the East Timor government made the difficult decision to close schools in an effort to stop the virus'

spread, which resulted in the disruption of education for 400,000 students [Ramos, 2020].

#### 3. Research Methodology

ISC, a private school of higher learning in East Timor served as the site for this study. In this study, there were190 student participants—57.9% female and 42.21% male that came from the faculties of education (41.6%) and health (58.4%).

The questionnaire in Table 1 requested replies on a scale of 1 to 5 from each participant. The survey adopted by Abbad (2021) research's study questions. Responses were provided using a Google form. The questions asked of the study's participants are included in Table 1.

Table 1. Questionnaire on the Use of E-Learning

Items	Questions
1	: What device was used to access e-learning?
2	: What was the most preferred learning model?
3	: Are you skilled in using the e-learning platform?
4	: Are you interested in using the e-learning platform in the future?

#### 4. Findings and Discussion

The "new normal" age, in which new behavior developed following the Covid-19 epidemic, has arrived in the modern world, including East Timor. This is particularly true of the use of information and communication technology (ICT) in education. Particularly in East Timor, Covid-19 compelled the shift from non-digital to digital learning. According to research conducted at ISC, cellphones (53.44%), laptops (43.39%), and personal computers (3.17%) were the most common types of devices used to access E-Learning.



Figure 1 Type of Technology Ownership

Figure 2 Learning Mode Preference

Only a third of the population in East Timor has access to internet infrastructure as a result of the country's poor internet penetration [UNDP, 2021]. Consequently, not all students can effectively access online learning. However, employing e-learning gives students a unique experience. According to a poll of students' preferred learning modes (See Figure 2), about 55.26% of students select face-to-face instruction while 44.74% select online instruction. This can be because students feel comfortable utilizing e-learning platforms because it is not a tough process to utilize it (see Figure 3). strongly agree (23.68%), agree (65.79%), moderate (8.42%), disagree (2.11%), and strongly disagree (0%) are all depicted in Figure 3. Figure 3 shows strongly agree (23.68%), agree (65.79%), moderate (8.42%), disagree (0%). In line with the results of this study, it shows that in general there are no obstacles in terms of mastering the use of the e-learning platform. This ease of use is what allows it to have an impact on student learning preferences and behavioral intentions to use e-learning in the future.







Figure 4 Behavioral intention to use e-learning

Plans to utilize the e-Learning platform in the future (item 4) may be affected by the study findings collected from the survey (item 3). The behavioral intention to utilize e-learning is shown in Figure 4 as strongly agreeing (35.79%), agreeing (59.47%), moderately agreeing (3.68%), disagreeing (0%), and strongly disagreeing (1.05%). Since East Timor has been promoting the use of distant learning modalities for enhancing the quality of teaching and learning in numerous policy papers, opportunities for the use of e-learning in the future in East Timor are becoming more and more abundant [UNESCO, 2023].

#### 5. Conclusion

As one of the countries where the number of mobile phones exceeds the total population, it is not surprising that East Timor students prefer smartphones as a device to access e-learning. With the experience of using e-Learning during the Covid19 pandemic, student learning preferences have also changed, quite a few students have chosen online learning as their learning preference mode, even though face to face learning is still the dominant choice. This ease of use is what allows it to have an impact on student learning preferences and behavioral intentions to use e-learning in the future.

#### Acknowledgements

We would like to thank the students from Faculty of Education and the Faculty of Nursing that participated in this research.

#### References

Abbad, M. (2021). Using the UTAUT model to understand students' usage of e-learning systems in developing countries. *Education and Information Technologies*, 7205–7224.https://doi.org/10.1007/s10639-021-10573-5.

- ESCAP. (2019). *Regulatory Policies and ICT Trends:*. https://www.unescap.org/resources/regulatory-policies-and-ict-trends-insights-timor-leste#.
- Kemp, S. (2021). *DIGITAL 2021: TIMOR-LESTE.* https://datareportal.com/reports/digital-2021-timor-leste.
- Kemp, S. (2023, February 14). *Data Reportal*. Retrieved from DIGITAL 2023: TIMOR-LESTE: https://datareportal.com/reports/digital-2023-timor-leste#:~:text=A%20total%20of%20 1.44%20million,percent%20of%20the%20total%20population.
- Ramos, L. (2020, June 22). *n Timor-Leste, "school goes home" during coronavirus*. Retrieved from Global Partnership: https://www.globalpartnership.org/blog/timor-leste-school-goes-home-during-coronavi rus
- UNDP. (2021). UNDP. Retrieved from https://www.undp.org/timor-leste/blog/starting-career-lockdown-impact-covid-19-yout h-education-and-employment-timor-leste

UNESCO. (2023). UNECO. Retrieved from https://education-profiles.org/eastern-and-south-eastern-asia/timor-leste/~technology UNICEF. (2020). UNICEF Education COVID-19 Case Study :Timor Leste – Learning at home and opening up better. https://aa9276f9-f487-45a2-a3e7-8f4a61a0745d.usrfiles.com/ugd/aa9276\_d3294a85 093045cfa7950423c200acb7.pdf.

# Technology-Enhanced Environmental Learning: Co-design of Educational Mobile Application Case

# Ewa DUDA<sup>a\*</sup>, Helena ANACKA<sup>b</sup>, Jolanta KOWAL<sup>c</sup>, Iwona NOWAKOWSKA<sup>a</sup> & Hanna OBRACHT-PRONDZYŃSKA<sup>d</sup>

<sup>a</sup>Maria Grzegorzewska University, Poland <sup>b</sup>Gdańsk University of Technology, Poland <sup>c</sup>University of Wrocław, Poland <sup>d</sup>University of Gdańsk, Poland \* eduda@aps.edu.pl

**Abstract:** The process of co-creating an educational mobile application to support environmentally friendly behavior is presented in this paper. The research material consisted of quantitative data collected on the application during the first testing phase by early adopters. The results suggest that the most frequently used features of the app were related to transport and educational activities. While women tended to split their time between transport and learning, men focused mainly on transport. This study contributes to the growing field of the use of educational mobile applications for the promotion of environmentally friendly behavior. The study results can support the design and development of future applications that encourage sustainable practices, benefit society, and contribute to a more sustainable future.

**Keywords:** mobile application, pro-environmental behavior, technology-enhanced learning, urban education

#### 1. Introduction

Mobile applications have played a significant role in our everyday routines for many years. It appears reasonable to utilize them to tackle contemporary problems, such as those associated with mitigating climate challenges (Takaoka et al., 2023). One example of a working application in Europe is PULA, which motivates pro-environmental behaviors in city dwellers. It aims to combine the benefits of a community currency and digital means of a mobile application to tackle a European city's environmental and urban challenges. This study presents the research outcomes of the PULA project implementation. Prior systematic literature review, case analysis, and in-depth interviews led to the identification of two research questions, namely: (RQ1) Whether PULA mobile application can increase ecological awareness and enhance environmentally-friendly activities among urban dwellers?; (RQ2) Which gamification actions of PULA application are the most popular among its users?

# 2. Related Works

Raising awareness and educating city dwellers about environmental issues is an area where applications can help and encourage pro-environmental behavior. Interactive mobile applications can educate users about how their daily behavior affects the environment and provide straightforward, practical suggestions for minimizing this impact. By disseminating reliable scientific information, such applications help to strengthen pro-environmental behavior (Wilson et al., 2019; Ahn, 2022). By breaking down educational material into short, engaging segments, it is possible to take advantage of the user's fragmented time and meet the demands of a fast-paced life (Chung et al., 2020).

Applications can be utilized to educate people on sustainability, reduce energy consumption, promote green transportation, encourage behaviors enhancing air quality,

manage waste, and conserve water (Douglas & Brauer, 2021). By adopting these applications, urban residents can scrutinize their consumption patterns and make more green choices (D'Arco & Marino, 2022). Such applications can stimulate sustainable actions through rewards and challenges (Cudok et al., 2022). Social engagement and inclusion influence individuals' behavior patterns and bring a sense of efficiency, effectiveness, and affiliation to the performed activities and community goals. Mobile applications proved to be egalitarian and inclusive tools that strengthen communities' cooperation, especially through gamification related to pro-environmental and local actions (Douglas & Brauer, 2021).

# 3. Methodology

To respond to the research aim and questions, following the method of Diffusion of Innovation Theory (Rogers, 2003), we analyzed data provided by the early adopters of the PULA application. According to the method, we fulfilled all its five phases: knowledge, persuasion, decision regarding the adoption or rejection of the innovation, implementation, and confirmation. By analyzing the behaviors of PULA early adopters, we evaluated the quality of the persuasion stage, observed the tendencies on the decision stage, aiming to plan the final implementation and prove stages of diffusion. The process involved 29 users (20 women and 9 men), while 44% represented the 18-25 age range and active employees or students already engaged in environmental activities, thus open to testing the PULA application. The testing phase took nine weeks. The paper presents the evaluation of data from the application.

# 4. Results and Discussion

The patterns of PULA application usage suggest that it has the potential to raise proenvironmental awareness in city dwellers. None of the testers gave up using the application (RQ1). The most utilized functionalities by PULA users were transportation activities and quizzes (RQ2). These activities were the most engaging, indicating participants' strong interest in environmentally friendly transportation options and knowledge acquisition.

# 4.1 Transport-oriented Functions

Overall, 70% of all transportation activities are associated with urban public transport, followed by cycling and walking. Gained results confirm that application users willingly opt for public transportation, thus reducing energy consumption and carbon footprint. Furthermore, the testing phase results suggest that an individual uses one preferred transportation means regularly, and its use notably prevails over other transportation activities of the same user. Nevertheless, it is difficult to formulate general conclusions based on a limited testing group and obtained data sample.

# 4.2 Knowledge Gaining Functions

The learning environment, which has been enhanced by the technology represented by the mobile app, is an effective tool for supporting environmental education. Learning activities, including quizzes and news stories, constituted a substantial part of the participants' tasks. The inclusion of news materials was notably prevalent. These activities were regularly sustained over time, especially among women.

Our findings provide evidence that mobile application-enhanced learning can support the development of environmentally friendly practices, including the use of public transport, beyond the scope of previous research (Ahn, 2022). Among the various quizzes tested on the PULA app, the briefest ones were the most favored. Previous research states that shorter learning materials are especially captivating on mobile applications (Chung et al., 2020) as they aid learning during non-urgent instances. In our study, these included breaks during lunchtime or, while driving to work/ university or on the journey back.

# 4.3 Gender Observations

Quantitative research results suggest that females are engaged in various types of activities in the application, paying particular attention to knowledge-gaining and educational activities, e.g., quizzes. On the other hand, males at large strongly emphasize transport functions, which are rewarded with the highest amount of points in the application. Therefore, the conclusion is that female and male users apply different strategies and are guided by different considerations while using the application. Gender differences were also observed in studies conducted under the framework of diffusion innovation theory, suggesting a moderating effect of gender between the perceived strengths of innovation and use intentions (Loarne-Lemaire et al., 2021). This trend is particularly apparent regarding technological advancements to tackle climate change and raise ongoing awareness of it.

# 5. Conclusion

Research results address two research questions: PULA mobile application can increase ecological awareness and enhance environmentally-friendly activities in urban dwellers (RQ1); urban public transport and quizzes are among the most frequently used app functionalities with the highest gained total points, with challenges being less frequent, yet characterized by the highest users' turnout (RQ2). The research findings, outlined in a full-length paper (Duda et al., 2023), could benefit ecological tech startups dealing with mobile tools, local authorities, educators, researchers, and sociologists unraveling effective technological solutions to boost pro-environmental activities. A limitation of the study presented is its pilot nature. Early adopters formed a small research sample, and their conduct is prone to bias. Due to the users' awareness of their participation in the pilot, they may have intentionally modified their behavior. Therefore, a more thorough investigation will take place soon, involving a larger study group and app users rather than testers.

# Acknowledgments

The research is supported by funding from Iceland, Liechtenstein, and Norway under the EEA Funds and the state budget of Poland, grant no. NOR/IdeaLab/GC/0003/2020.

# References

- Ahn, J. (2022). Impact of cognitive aspects of food mobile application on customers' behavior. *Current Issues in Tourism, 25*(4), 516-523.
- Chung, E., Noor, N. M., & Mathew, V. N. (2020). Are you ready? An assessment of online learning readiness among university students. *International Journal of Academic Research in Progressive Education and Development*, *9*(1), 301-317.
- Cudok, A., Lawrenz, S., Rausch, A., & Vietor, T. (2022). Circular Economy Driven Communities– Sustainable Behavior Driven by Mobile Applications. *Procedia CIRP*, *105*, 362-367.
- D'Arco, M., & Marino, V. (2022). Environmental citizenship behavior and sustainability apps: An empirical investigation. Transforming Government: People, *Process and Policy*, *16*(2), 185-202.
- Douglas, B. D., & Brauer, M. (2021). Gamification to prevent climate change: a review of games and apps for sustainability. *Current Opinion in Psychology*, *42*, 89–94.
- Duda, E., Anacka, H., Kowal, J., Nowakowska, I., ...& Zawieska, J. (2023). Encouraging proenvironmental behaviour through an educational mobile application: preliminary insights from early adopters. *International Journal of Pedagogy, Innovation and New Technologies, 10*(1).
- Loarne-Lemaire, S. L., Bertrand, G., Razgallah, M., Maalaoui, A., & Kallmuenzer, A. (2021). Women in innovation processes as a solution to climate change: A systematic literature review and an agenda for future research. *Technological Forecasting and Social Change, 164*, 120440.

Rogers, E.M. (2003). Diffusion of innovations (5th ed.). New York: Free Press.

- Takaoka, A. J., Ahlers, D., ... & Jaccheri, L. (2023). Towards understanding digital support contributing to climate neutral, inclusive, and beautiful cities: A systematic literature review. *GREENS*, 30-37.
- Wilson, C., Tabor, A., & Bateman, S. (2019). Diversifying pro-environmental behaviors: redesigning applications to incorporate environmental "spill-over". *Proceedings of the 2019 ACM International Symposium on Wearable Computers*, 1097-1098.

# Urban Living Lab Enhanced by a Mobile Application as a New Way to Educate Towards Green and Inclusive Cities

### Ewa DUDA<sup>a\*</sup>, Łukasz GONTAR<sup>b</sup>, Maksymilian KOCHAŃSKI<sup>b</sup> & Mari Hanssen KORSBREKKE<sup>c</sup>

<sup>a</sup> Institute of Education, Maria Grzegorzewska University, Poland
 <sup>b</sup> Research and Innovation Centre Pro-Akademia, Poland
 <sup>C</sup> Western Norway Research Institute, Norway
 \*eduda@aps.edu.pl

Abstract: Amidst rapid technological advancements and escalating environmental crises, there is a pressing need for innovative educational approaches addressing these challenges. This study introduces the Urban Living Lab (ULL) concept as an educational environment to engage households in sustainable food production enhanced by a mobile application. Leveraging a learning-by-doing approach and new technologies, the study focuses on the "SmartFood" project, which aims to reduce the environmental footprint by promoting self-produced food and fostering proenvironmental behaviors in urban communities. The research employs co-design workshops with future ULL participants to understand their needs and expectations. Utilizing thematic analysis, the study identifies both enablers and barriers to the project's implementation. The SmartFood mobile app emerged as a pivotal tool, providing users with valuable information on plant growth conditions and facilitating communication among participants. The study reveals that ULL enhanced by mobile application can serve as effective platform for educating households on sustainability. It also highlights the need for more frequent and detailed educational sessions to maximize the sustainability impacts. Overall, the study contributes to the limited body of research on ULLs as educational environments beyond higher education and offers practical insights for the co-design of sustainable urban food systems.

**Keywords:** Environmental Education, Urban Living Lab, Sustainable Food Production, Co-Design Workshops, Mobile Application, Pro-Environmental Behaviors

# 1. Introduction

With the rapid advancement of technology, as a society we are faced with many new challenges. Firstly, the negative impact of human activity on the environment is giving rise to ever new crises, for which we must begin to prepare effectively on the one hand, and which we must begin to counteract on the other. Secondly, the development of new technologies means that we should look for new tools to educate and engage the public, better adapted to their changing needs. The study presented here is an attempt to introduce the increasingly popular educational environment of the Urban Living Lab concept, supported by new technologies, which is based on a learning-by-doing approach.

# 2. Urban Living Lab Concept

# 2.1 Related Works

Urban Living Labs (ULL) are often described as "urban innovation ecosystems" that serve as frameworks for co-creation of innovation in cities (Yilmaz & Ertekin, 2023). As platforms for iterative learning, they are characterized by experimentation, geographical embeddedness, participation and user involvement, leadership and ownership, and evaluation and refinement

(Beecroft, 2018). However, limited research has been completed so far on Urban Living Labs as educational environments going beyond higher education contexts (Martinez-Acosta et al., 2023; Bogedain and Hamm, 2020). Some of the prior studies have focused on living labs as spaces for testing a mobile app-based health program (Kim et al., 2022), indirect education on energy consumption habits (Brohmer et al., 2023), and increasing awareness of food waste reduction targets (Black et al., 2023). Still, no prior research has used the concept of Urban Living Lab to engage households in the co-designing of a mobile-app-assisted system for the self-production of food in cities.

# 2.2 SmartFood Project Case

To address the identified challenges, the "SmartFood: Engaging citizens in food diversity in cities" project (SmartFood) aims to contribute to the reduction of the environmental footprint by presenting a model for self-produced food, thereby reducing food transportation. The project is aimed at shaping a vision of cities of the future in which sustainable patterns of food consumption and production, based on healthy eating habits, controlled dependence on food retail, reduced food waste and social networks in urban buildings play a dominant role. The ambition of the SmartFood project team is to implement a controlled experiment involving a local neighborhood community, its building administration and local stakeholders in the process of co-creating an educational model for shaping pro-environmental behaviors that counteract climate change. SmartFood serves as a starting point for the co-design of sustainable hydroponic installation in the common corridors of urban residential blocks of flats and the preparation of a social experiment within SmartFood Urban Living Lab in Poland.

# 3. Methodology

The aim of the research progression is to diagnose problems of food production and waste as well as to understand its essence. We invited people to share their experiences, needs, understandings of the problem and expectations. The research material presented here consists of qualitative data collected during two sessions of co-design workshops (Örnekoğlu-Selçuk et al., 2023) with future Urban Living Lab participants, conducted online. Data analysis was carried out by applying the thematic analysis methodology proposed by Braun and Clarke (2019). Due to the high cost of the hydroponic installation envisaged by the SmartFood urban experiment, the project was designed for 20 residents of a selected residential block who participated in the workshops. The research presented here therefore has the status of a case study (Sjoberg et al., 2020). It aims to answer the following research question: How a mobile application could increase the engagement of households in sustainable food production?

# 4. Results and Discussion

Participants of the workshops found the mobile application to be a useful tool to support the learning process, both in terms of self-growing food in hydroponic cabinets and as a valuable source of knowledge. In particular, the SmartFood App is regarded as a promising contact channel between the ULL's participants and researchers. Its most important feature is the information, it provides the user related to cultivation conditions, such as light, humidity, substrate parameters, and send notices of emerging irregularities.

Participants of the SmartFood co-design workshops identified more enablers than barriers to the project implementation. Supporting factors identified included economic factors, the ability to monitor changes in one's diet and grocery purchases, the associated opportunity for self-reflection, the strengthening of environmental awareness, the opportunity to develop interests by growing one's own fruit and vegetables, and the desire to reduce one's carbon footprint. Predominant among the barriers were people's inherent prejudices, low public commitment to any kind of pro-environmental innovation, fear of insects, reluctance to change eating habits, lack of time and technical problems in entering data into the app.

A more effective implementation of the SmartFood vision could be enhanced by patient education of future users, dissemination activities, encouraging the consumption of edible insects rather than farming them, creating a space in the app for advice on how to use the products from the cabins (vegetables, worms). Workshop participants expressed approval of this form of involving potential users in the co-design of SmartFood solutions. They saw their advantages in terms of being able to actively participate, to speak freely, and to be able to obtain a lot of information that was important and interesting to them. In the future, however, it would be useful to hold more frequent but shorter workshops, and to provide more detailed and specific educational information.

#### 5. Conclusion

Upon reflection on the findings from the co-design part of the ULL conducted at the beginning of the project period we found that it taught us valuable lessons in how to implement important elements of the project when moving forward with participants. In introducing them to elements of the educational format and co-producing aspects of the Lab we strengthened the ULL design to be more inclusive and productive. With high levels of participation throughout the project progression we found that discussing the barriers and drivers of use of the hydroponic solutions, the mobile app and surrounding technologies was necessary for increasing the legitimacy of the experiment and generating more sustainable use and engagement. Despite high levels of participation being somewhat resource draining, we concluded that more frequent and detailed educational sessions are necessary to ensure the process fulfills its educational potential and to maximize the sustainability impacts and engagement in pro-environmental behaviors. The findings also offer practical implications for the co-design of sustainable urban food systems, highlighting the key role of mobile applications and the need to address both enablers and barriers to implementation.

### Acknowledgements

The research is supported by funding from the Norway Grants 2014-2021 and the state budget of Poland, grant no. NOR/IdeaLab/SmartFood/0005/2020.

# References

- Beecroft, R. (2018). Embedding Higher Education into a Real-World Lab: A Process-Oriented Analysis of Six Transdisciplinary Project Courses. *Sustainability*, *10*(10), 3798.
- Black, D., Wei, T., Eaton, E., Hunt, A., Carey, J., Schmutz, U. He, B., Roderick, I. (2023). Testing Food Waste Reduction Targets: Integrating Transition Scenarios with Macro-Valuation in an Urban Living Lab. *Sustainability*, *15*, 6004.
- Bogedain, A., Hamm, R. (2020). Strengthening local economy an example of higher education institutions' engagement in "co-creation for sustainability". *REGION*, 7(2), 9-27.
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*(4), 589-597.
- Brohmer, H., Munz, K., Röderer, K. et al. (2023). How attractive is the participation in a Living Lab study? Experimental evidence and recommendations. *Discover Sustainability, 4*, 23.
- Kim, Y., Lee, H. & Chung, M.L. (2022). Living labs for a mobile app-based health program: effectiveness of a 24-week walking intervention for cardiovascular disease risk reduction among female Korean-Chinese migrant workers: a randomized controlled trial. *Archives of Public Health, 80*, 181.
- Martínez-Acosta, M., Vázquez-Villegas, P., Mejía-Manzano, L. A., Soto-Inzunza, G. V., Ruiz-Aguilar, K. M., Kuhn Cuellar, L., Caratozzolo, P. and Membrillo-Hernández, J. (2023). The implementation of SDG12 in and from higher education institutions: universities as laboratories for generating sustainable cities. *Frontiers in Sustainable Cities*, *5*,1158464.
- Örnekoğlu-Selçuk, M., Emmanouil, M., Hasirci, D., Grizioti, M., & Van Langenhove, L. (2023). A systematic literature review on co-design education and preparing future designers for their role in co-design. *CoDesign*, 1-16.
- Sjoberg, G., Orum, A. M., & Feagin, J. R. (2020). *A case for the case study*. The University of North Carolina Press.
- Yilmaz, O.C., Ertekin, O., Urban Living Labs as a Tool to Achieve Sustainable Development Goal 16: A Case Study of Istanbul, Turkiye (2023). *Theoretical and Empirical Researches in Urban Management*, 18(3), 88 – 118.

# Supporting Engineering Degree Student Wellbeing with Compulsory Lessons on Stress Management

## Marc BEARDSLEY

<sup>a</sup>Department of Information and Communication Technologies, Universitat Pompeu Fabra, Spain

**Abstract:** In this paper, we present findings from the evaluation of compulsory lessons on stress management at an engineering school in Spain. Both a face-to-face and an online lesson were integrated into a credit-bearing course for first-year university students. A pre-post survey study was conducted (n = 54). Results indicate that the lessons fostered more positive stress mindsets among students. Findings also suggest that students increased their proficiency in contemplating the topic of academic stress and in understanding the active role they can play in managing stress.

Keywords: Stress Mindsets, Stress Management, Student Wellbeing

# 1. Introduction

Teaching university students about stress and coping strategies can benefit their learning (Vogel & Schwabe, 2016). Moreover, learning to manage academic stress can carryover outside of academics and improve overall student wellbeing (Ebrahimi *et al.*, 2019). This paper presents an evaluation of a stress management intervention centered on self-efficacy theory (Bandura, 1977). The intervention was integrated into a compulsory course for university students and aims to demonstrate an approach through which universities can address the need to support healthy coping in university (Böke *et al.*, 2019).

# 2. Methodology

A mixed methods design within a design-based research methodology was used as the educational intervention has been iterated upon across 4 academic years.

# 2.1 Participants

54 first-year students from a university in Spain responded to pre and post surveys (female = 17, male = 37,  $M_{-age}$  = 18.17, SD = 1.71) in the 2022-23 academic year. All were enrolled in a compulsory general studies course that is part of an engineering degree programme.

# 2.2 Data Collection and Analysis

Data was collected in pre- and post-surveys using the Stress Mindset Measure (SMM), a closed-ended question ('Did you know that UPF offers a psychological support service?') and an optional open-ended question ('In what ways can the university better support students in coping with academic stress?'). UPF is an abbreviation of the university in which the study took place. Additional questions appeared in the surveys but are not reported in this paper. SMM is an 8-item measure (see Table 1) that assesses the extent to which an individual believes that the effects of stress are debilitating or enhancing. Items are rated on a 5-point scale (0 = strongly disagree to 4 = strongly agree). Scores are determined by reverse scoring

four negative items (indicated by an \*) and then taking the mean of all items. Scores (above 2) represent a stress-is-enhancing mindset whereas lower scores represent a stress-is-debilitating mindset (Crum *et al.*, 2017). All surveys were Google Forms either in English (n = 20) or Spanish (n = 34). The survey language was determined by the line of studies respondents were enrolled in as the university offers distinct English and Spanish/Catalan lines of studies. Responses were excluded from students that did not complete both pre and post surveys or had completed the surveys incorrectly. Quantitative data was analyzed using Excel. Thematic analysis was used for qualitative data.

# 2.3 Procedure

In week 1 of a 10-week general studies course, students completed the pre-survey at the onset of a 2-hour face-to-face class on stress management. The post-survey was assigned as an online task posted to the course learning management system (Moodle) at the end of the trimester, 8 weeks after the pre-survey. The compulsory lessons included the week 1 face-to-face class and an online lesson. The face-to-face class introduced students to concepts related to the science of stress (human stress response, stress mindsets) and evidence-based self-regulation techniques such as cognitive reappraisal and breathing techniques. The online lesson took place 5 weeks later. It required students to complete two units of an online course on the science of stress (https://tidex.upf.edu/courses/course-v1:Spotlighters+SP1+2021\_1/). The first unit was on how lifestyle choices affect one's stress response. For the second unit, students were able to select from 5 available units. Upon completing a unit, students responded to reflective questions ('How has your knowledge about stress changed? How can you apply what you learned? What did you agree or disagree with?', 'How useful do you think it is for students to learn the content of this unit?'). Open educational resources (OER) created in European projects (Spotlighters, BEAM) were used to produce the lessons.

# 3. Results and Conclusions

# 3.1 Student Beliefs about Stress

The average SMM score from the pre-survey was 1.47 (SD = .63) out of 4 indicating that overall respondents held stress-is-debilitating mindsets. The average SMM average score on the post-survey increased to 2.06 (SD = .65). SMM scores had increased significantly from pre to post; t(54) = 30.45, p < .0001. Table 1 presents the number of students agreeing/strongly agreeing with each SMM statement to highlight the effects of the intervention – and to help identify areas to focus on in improving the intervention in the future.

Statement	n <sub>-pre</sub>	n <sub>-post</sub>	pre-post (%)
*The effects of stress are negative and should be avoided.	25	12	-13 (-52)
*Experiencing stress depletes my health and vitality.	40	28	-12 (-30)
*Experiencing stress inhibits my learning and growth.	20	15	-5 (-25)
*Experiencing stress debilitates my performance and productivity.	22	19	-3 (-13.6)
Experiencing stress facilitates my learning and growth.	12	27	15 (125)
Experiencing stress enhances my performance and productivity.	16	30	14 (87.5)
Experiencing stress improves my health and vitality.	0	3	3 (NA)
The effects of stress are positive and should be utilized.	2	24	22 (1,100)

Table 1. Count of Students Agreeing or Strongly Agreeing with Belief Statements about Stress

These findings are in line with previous studies that have found that exposing students to different beliefs about stress (i.e., the enhancing properties of stress) can lead to a shift in their stress mindsets and better help them benefit from challenging situations (Crum *et al.*, 2017; Beardsley *et al.*, 2021). As seen in Table 1, the biggest changes are found in items related to perceiving stress in a positive manner as many students had a limited understanding of the human stress response and the role it plays in preparing oneself to face challenges.

## 3.2 University Support to Help Students Cope with Academic Stress

From pre- to post-survey, students marked being more aware that the university offered a psychological support service (pre = 51.2%, post = 75.9%). In relation to the question about how the university can better support students in coping with academic stress, more students responded in the post survey ( $n_{-pre} = 16$ ,  $n_{-post} = 34$ ) and responses were more elaborated with a higher average word count:  $M_{-pre} = 11.87$  words (SD = 9.74),  $M_{-post} = 21.97$  words (SD = 22.03). A thematic analysis identified two primary themes: (1) accommodate us – which reflects students suggesting adjustments be made for them (e.g., Being flexible with classwork delivery dates), (2) help us learn – which reflects students suggesting being taught how to improve (e.g., With classes focused on stress and anxiety). Identified subthemes are shown in Table 2. In total, the pre-survey generated 16 categorisations and the post-survey 40. The coding of responses suggests that students became more aware of the help available to them outside the classroom (e.g., psychological support services: Pre = 0, Post = 17.5\%) and better understood the potential role they could play in managing their academic stress (e.g., help us learn: Pre = 12.5%, Post = 35%; stress management: Pre = 6.25%, Post = 27.5%).

Theme	Subtheme	Pre-survey (%)	Post-survey (%)
Accommodate us		13 (81.25)	25 (62.5)
	Teacher support	9 (56.25)	7 (17.5)
	Lesson design/content	4 (25)	11 (27.5)
	Psychological support services	0	7 (17.5)
Help us learn		2 (12.5)	14 (35)
	Self-management	1 (6.25)	3 (7.5)
	Stress management	1 (6.25)	11 (27.5)
Other	(i.e., Not sure)	1 (6.25)	1 (2.5)

Table 2. Theme Ca	tegorizations of Responses	related to University	Support for Students
-------------------	----------------------------	-----------------------	----------------------

Results suggest that students may have acquired a greater proficiency in contemplating the topic of academic stress and understanding the active role they can play in managing stress. However, this study has limitations due to the nature of the sample (opportunity sample) and small sample size. All in all, these findings demonstrate a reusable approach, based on OER, through which universities can address the need to support healthy coping in university.

#### Acknowledgements

This work has been partially funded by Pompeu Fabra University's PlaCLIK (E2022014338), Erasmus+ programme (2021-1-ES01-KA220-SCH-000032801, RemixED), and Spanish Ministry of Science and Innovation (PID2020- 112584RB-C33 granted by MICIN/AEI/10.13039/501100011033).

# References

- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191.
- Beardsley, M., Vujovic, M., & Portero-Tresserra, M. (2021). Evaluating an intervention to improve university student stress mindsets and self-management skills. *Revista del Congrés Internacional de Docència Universitària i Innovació (CIDUI)*, (5).
- Böke, B. N., Mills, D. J., Mettler, J., & Heath, N. L. (2019). Stress and coping patterns of university students. *Journal of College Student Development*, 60(1), 85-103.
- Ebrahimi, O. V., Pallesen, S., Kenter, R. M., & Nordgreen, T. (2019). Psychological interventions for the fear of public speaking: A meta-analysis. *Frontiers in Psychology*, *10*, 488.
- Crum, A. J., Akinola, M., Martin, A., & Fath, S. (2017). The role of stress mindset in shaping cognitive, emotional, and physiological responses to challenging and threatening stress. *Anxiety, Stress, & Coping*, *30*(4), 379-395

Vogel, S., & Schwabe, L. (2016). Learning and memory under stress: implications for the classroom. *NPJ Science of Learning*, *1*(1), 1-10.

# DSC

# Exploring the Potential of Mobile Assisted Language Learning in Rural Area of Indonesian

**Fais Nurul HADI** <sup>a\*</sup>, <sup>a</sup>Charles Darwin University, Australia \*faisnurul.hadi@cdu.edu.au

**Abstract:** This study aims to explore the potential of Mobile Assisted Language Learning (MALL) in Indonesian rural areas. Rural communities often face challenges to access a good education. Therefore, an innovative approach is needed that can overcome these obstacles. MALL is emerging as a potential solution that harnesses the power of mobile technology to enhance the language learning experience, providing flexibility and learning opportunities according to the needs of learners. This research uses a systematic approach in the literature review to explore the benefits of MALL in improving language learning outcomes in Indonesian rural communities. By exploring the potential of MALL in language learning in rural Indonesian communities, this research is expected to contribute to improving the practice of language education in rural areas, as well as providing equal educational opportunities for students throughout Indonesia.

Keywords: Mobile-Assisted Language Learning, Mobile Learning, Language Learning

# 1. Introduction

Research on Mobile Assisted Language Learning (MALL) has gained considerable concerns in language learning developments (Hou & Aryadoust, 2021; Kamasak et al., 2020; Shadiev et al., 2020). Previous research on MALL has provided many valuable findings about its potentials such as improving language learning experience, more personalized approaches, flexibility, and collaborative opportunities (Kukulska-Hulme & Shield, 2008). Other MALL research directions also tend to analyze the extent to which the effectiveness of the use of mobile devices in improving the quality of language learning with various research contexts, strategic approaches, and various targeted language aspects (Hwang & fu, 2018; Sung et al., 2015).

Research on MALL in the Indonesian context provides a different perspective in terms of the potential for developing language learning and the barriers to implementation (ibrahim, 2021). The complexity of Indonesian local cultures and language diversity has become such a unique challenge in the implementation of MALL. Moreover, it also concerns about the language need and cultural context for Indonesian students to improve their language learning experiences (Estuarso, 2017). In this case, the use of mobile devices can be the main alternative in providing solutions to issues of inequality in education because mobile devices can be optimized for language learning anytime and anywhere while also minimizing barriers from geographical factors (Fotouhi-Ghazvini et al., 2009). In addition, the very rapid increase in the penetration rate of mobile technology in Indonesia is also an important reason for the need to develop the potential of MALLs in improving the quality of education in Indonesia (Pratama, 2018).

The effectiveness of MALL can also be promoted to improve the quality of education in rural areas in Indonesia. Most research on MALL has focused on regional contexts with advanced technological developments or urban areas. Other researchers had developed MALL research to a more general context regardless of the conditions of the area that is the research target. Therefore, research on MALL in rural areas needed to be highlighted, especially to improve the quality of education in rural areas (Fotouhi-Ghazvini et al., 2009).

This paper is expected to provide an understanding of the urgency of implementing MALL in rural areas of Indonesia and to what extent MALLs can be developed further. Moreover, this discussion will also identify how MALL implementation practices can be carried out more effectively for recommendations for education policy makers, stakeholders and teachers to develop MALL in rural Indonesian contexts.

#### 2. Method

This study uses a literature-based approach to explore teaching and learning in rural areas of Indonesia. The research design involves a systematic review and analysis of relevant literature from various sources, including academic journals, books, reports, and other scholarly publications. Several leading academic databases, such as ERIC, JSTOR, and Google Scholar, were selected to obtain relevant literature. The identified search terms are used to perform a systematic search in selected databases and sources. The screening process involves reviewing the titles and abstracts of the sources found to assess their relevance to the research topic. Retained sources were carefully read and analyzed to extract key information and findings. Thematic analysis was used to identify recurring themes, patterns, and challenges related to teaching and learning in rural contexts of Indonesia.

#### 3. Illustration the Educational Condition in Rural Area of Indonesia

As the biggest archipelagic country consisting of 17,499 islands, Indonesia has continuously experienced problems in the equitable distribution of resources, especially in education, due to various geographical problems and social-cultural differences in society (Luschei & Zubaidah, 2012). This problem becomes very complicated when it occurs in the context of education in rural areas. Rural areas in Indonesia refer to areas that are far from urban centers and are often difficult to reach due to difficult geographical conditions, such as mountains, wilderness, or small islands. Rural areas usually have lower populations and lack basic infrastructure such as roads, electricity, clean water, and adequate educational facilities (Madhakomala et al., 2022).

The geographical differences between urban, rural and rural areas have significant implications for educational conditions. Urban areas generally have better access to educational facilities, such as schools with complete facilities, stable internet access, and good transportation accessibility. On the other hand, rural and rural areas often face challenges in terms of school accessibility, limited facilities, and the availability of educational resources. This inequity is also closely related to the funds allocations from local government (Sukasni & Efendy, 2017). Long distances and limited means of transportation can be obstacles for students to reach school safely and on time. In addition, supporting infrastructure such as limited or non-existent electricity, networks, availability of clean water, and poor sanitation can also affect learning conditions in rural areas (Nuraeni et al., 2020). Geographical factors also have an impact on the availability of educational resources. Rural areas tend to have limitations in terms of textbooks, laboratory equipment, computers and other technological devices. These limitations can affect student learning experiences and the teacher's ability to present material well.

Other common issues faced by rural education are the availability of the skilled teachers, the low motivation of students to learn, and the low support from the communities for education (Fadilah et al., 2023). This issue is further worsen with lack learning materials or resources, teaching aids or media, teaching methods, slow internet connectivity, learners' motivation, evaluation technique and parental support (Mudra, 2018). Moreover, Indonesian education management can also be concern particularly in rural areas for the ineffective teaching methods, insufficient facilities, and infrastructure (Solihin & Sara, 2021).

Some efforts to improve the education quality in rural area of Indonesia have not shown the good progress. Even after having a mass recruitment of teachers from the Indonesian government, the inadequacy and the low welfare of teachers remains the issue for rural education (Febriana et al., 2018). Teachers often have to teach multiple subjects which they might not be qualified for (Anwar et al., 2020). These complex issues of inequality in access to and quality of education in Indonesia can also be considered as the effect of various factors at individual, household, school, community, and government levels (Muttaqin, 2018). Therefore, it is necessary to focus on development of teachers' professionalism in remote areas, including facilitating further studies, sharing abilities among teachers, and improving school facilities (Kurniati et al., 2020). Moreover, the professional teacher training should be considered based on geographical location, culture, and the limitations of human resources and infrastructure (Fauziyah & Uchtiawati, 2017).

### 4. The Significance of MALL for Rural Area of Indonesia

The urgency of MALL in rural areas is very important and has a significant impact in improving the quality of education. MALLs provide greater accessibility and affordability to educational resources. In rural areas that often have limited access to educational facilities and learning resources, MALLs allow students and teachers to flexibly access learning materials, applications and online learning resources through mobile devices that are more affordable and easy to use (Rajendran & Yunus, 2021).

The other significance is that MALL allows students to learn languages anytime and anywhere. This is especially important for rural areas where schools may be far away, or travel times are long. By using the MALL application, students can learn languages at home or in their surroundings, according to their needs and time availability (Kukulska-Hulme & Shield, 2008). Some studies even found out that the students' acceptance and perception toward the use of mobile assisted learning were high (Lizamuddin et al., 2019; Yudhiantara & Nasir, 2017)

MALL can also improve the quality of learning in rural areas through the various features and applications available (Rajendran & Yunus, 2021). For example, interactive and multimedia language learning applications can help increase student motivation and understanding. In addition, MALLs can also provide access to a more varied learning material, such as videos, audio, and e-books, which help enrich students' learning experiences. The use of MALL also allows students in rural areas to develop digital skills which are essential in today's information technology era. Using mobile devices, students can learn about using technology, navigating the internet, and using relevant applications. This helps prepare them for the needs of the digital world and enhances their opportunities in the future.

Related to the learning content and student's needs, MALL development can produce language learning content that is tailored to the needs of students in rural areas. Such content may include learning modules, exercises and interactive games specifically designed to reinforce language skills. These contents can be adapted to students' language level, relevant topics, and local cultural context (Estuarso, 2017).

MALLs can also connect students in rural areas with the wider language community via social media or sharing platforms. Students can interact with native speakers of the target language or with fellow students from other areas who are learning the same language. This can help increase student motivation and engagement in language learning (Kukulska-Hulme & Shield, 2008). In this case, parental and community involvement in education in rural areas can also be included. With the MALL platform that allows communication between teachers, students, and parents, better interactions can be established in monitoring student learning progress. Parents can also be actively involved in supporting their child's learning process through monitoring, guidance, and collaboration with teachers

Rural areas often face limited educational resources such as textbooks, laboratory equipment, and limited learning materials (Febriana et al., 2018). MALLs can help overcome this limitation by providing access to a more diverse and up-to-date digital content and online learning resources. Thus, students in rural areas can access more complete and up-to-date learning materials even though local resources are limited.

MALL applications can provide monitoring and feedback features that can help students and teachers monitor language learning progress (Ishaq et al., 2021). Students can track their progress, identify areas for improvement, and receive immediate feedback via the app. This can help students in rural areas manage their learning independently and increase learning effectiveness.

### 5. Identifying the Challenges of MALL Implementation in Rural Area

Implementation of Mobile-Assisted Language Learning (MALL) in rural areas may face several obstacles. One of the main barriers is the availability of adequate technological infrastructure in rural areas. Unstable or even widely available internet connectivity can hinder access and use of the MALL. Limited networks and lack of infrastructure such as signal towers and limited access to electricity can be serious obstacles.

Other obstacles are about the accessibility and the availability of mobile devices. Although mobile technology is becoming more common, there are still limitations in mobile device accessibility in rural areas. Many students and families in the area may not have access to the necessary mobile devices, such as smartphones or tablets, to access and use the MALL. In addition, the cost of the device and the need for internet access can also be obstacles (Nuraeni et al., 2020).

The Major concern of MALL implementation in Rural area is also about the limited technological knowledge and skills both from teacher and students. In rural areas, teachers and education staff may have limited knowledge and skills in using mobile technologies for education. Adequate training and coaching are needed so that they can integrate MALL in learning effectively. Lack of knowledge about proper application and learning methods can be a barrier in tapping into MALL's full potential (Tømte, 2015). Moreover, not all learning content available in digital formats or MALL applications is relevant to the context and needs of students in rural areas. Content related to the culture, environment, and everyday life of students may be less available. The lack of content adapted to local contexts can reduce the effectiveness of MALLs in enhancing student learning in rural areas.

Other issues of MALL application is related to sociocultural issues and the readiness to the use of mobile devices in learning (Eltayeb et al., 2014). The use of MALL requires changes in learning behavior and culture among students, teachers, and people in rural areas. Not all students or teachers may be accustomed or ready to adopt mobile technology in the learning process. Awareness, understanding and support from all stakeholders is important to overcome resistance to change and promote acceptance of the use of MALL.

#### 6. Implication and Recommendation for Better MALL Application

There are several recommendations that can be applied in the context of using Mobile Assisted Language Learning (MALL) in Indonesian rural areas. In the provision of access to technology, the government and relevant stakeholders need to work together to ensure the availability of adequate infrastructure and access to technology in rural communities (Fauziyah & Uchtiawati, 2017). This includes providing a reliable internet network, affordable mobile devices, and access to relevant MALL applications or platforms. In this case, recommendations can be in the form of government programs that encourage access to technology in rural areas and collaboration with institutions or organizations to provide technological assistance to rural communities.

Another important implication is the need to increase the Teacher Training (Kurniati et al., 2020). Teachers in rural areas need to be provided with adequate training regarding the use of MALL in the language learning process. This training may include knowledge of relevant MALL applications or platforms, teaching strategies for effective use of technology, and an understanding of how to overcome constraints that may arise in the use of MALL in rural settings.

More efforts are also needed to develop educational content that is appropriate to the rural context and takes advantage of the potential of mobile technology. The current interest on developing educational contents was focusing on learners and their involvement in knowledge sharing in a contextualized mode (Foomani & Hedayati, 2016). These contents can include interactive learning modules, supporting materials in the form of video or audio, and digital resources that are relevant to the life and culture of rural communities. Recommendations can be in the form of collaboration between content developers, educational institutions, and rural communities in designing and developing appropriate educational content.

Other important matter is about collaboration with local communities: It is important to involve local communities in the development and implementation of MALLs in rural areas.
Through collaboration with the community, such as community leaders, parents, and village leaders, a better understanding of local needs and context can be found. This collaboration can also strengthen social support and acceptance of the use of MALL in language learning. Teacher can also be a bridge for this collaboration for students to integrate their language learning with their own local culture (Estuarso, 2017).

Monitoring and evaluation can be a crucial process in MALL implementation. A continuous monitoring and evaluation system are needed to measure the effectiveness of using MALL in rural communities. This evaluation should cover aspects such as improving language learning outcomes, student motivation, active participation, and user satisfaction. Evaluation results can be used to adjust and improvements needed in the implementation of MALL. Further research on technological innovation will always be needed in order to develop better education quality including MALL. It is very necessary to continue research on the latest technological innovations that can support MALLs in rural areas.

#### 7. Conclusion

Rural areas in Indonesia face challenges in improving the quality of education, including limited access to educational resources, limited infrastructure, and a lack of qualified teachers. In this regard, MALL has great potential in improving education in rural areas by providing better accessibility, quality, and engagement in learning. MALL can also overcome the limited accessibility of learning resources, while also offer more engagements, freedom, and flexibility in language learning for rural areas. By optimizing various supports and collaborative efforts, it is hoped that MALL can become an effective tool in improving the quality of education in rural areas in Indonesia and promoting equity of educational opportunities throughout the country.

#### Acknowledgements

I would like to thank my supervisor, Dr Jon Mason, for his guidance and advices, all reviewers from ICCE 2023 conference, and all the people who prepared and revised previous versions of this document.

#### References

- Anwar, K., Ubaidillah, M. F., & Sulistiyo, U. (2020). Exploring EFL Teachers' Classroom Management: The Case of Indonesian Remote Secondary Schools. *Journal of Language and Education*, 6(3), 22–35. http://dx.doi.org/10.2139/ssrn.3862391
- Eltayeb, M., Osman, M., Hegazi, A., & Hegazi, M. O. (2014). Mobile Learning Aspects and Readiness. International Journal of Computer Applications, 103(11), 22–28. https://doi.org/10.5120/18118-9401
- Estuarso, D. (2017). When Local Goes Global: English Teachers As Cultivators Of Local Culture In A Globalized Region. *KnE Social Sciences*, *1*(3), 60–69. https://doi.org/10.18502/kss.v1i3.725
- Fadilah, R., Ayudhia, H. Y., Pratama, V. M., & Harmawan, V. (2023). Teachers' Teaching Reflection: English Teachers' Challenges in Rural Areas of Indonesia. *LET: Linguistics, Literature and English Teaching Journal*, 13(1), 85–104. https://doi.org/10.18592/let.v13i1.8923
- Fauziyah, N., & Uchtiawati, S. (2017). Developing a Model of Educators' Professional Training Special for Remote Areas through the Implementation of Lesson Study. *International Education Studies*, 10(8), 108–115.
- Febriana, M., Nurkamto, J., Rochsantiningsih, D., & Muhtia, A. (2018). Teaching in Rural Indonesian Schools: Teachers' Challenges. International Journal of Multicultural and Multireligious Understanding, 5(5), 11. https://doi.org/10.18415/ijmmu.v5i5.305
- Foomani, E. M., & Hedayati, M. (2016). A Seamless Learning Design for Mobile Assisted Language Learning: An Iranian Context. *English Language Teaching*, 9(5), 206–213.
- Fotouhi-Ghazvini, F., Earnshaw, R. A., & Haji-Esmaeili, L. (2009). Mobile Assisted Language Learning in a Developing Country Context. 2009 International Conference on CyberWorlds, 391–397. https://doi.org/10.1109/CW.2009.28
- Hou, Z., & Aryadoust, V. (2021). A review of the methodological quality of quantitative mobile-assisted language learning research. *System*, *100*, 1–38. https://doi.org/10.1016/j.system.2021.102568

- Hwang, G.-J., & fu, Q. (2018). Trends in the research design and application of mobile language learning: A review of 2007–2016 publications in selected SSCI journals. *Interactive Learning Environments*, 27, 567–581. https://doi.org/10.1080/10494820.2018.1486861
- ibrahim, ibrahim. (2021). Mobile-Assisted Language Learning (MALL) in Indonesia: Case on English Language Learning for Higher Students during COVID 19. *Journal of English Education*, *6*(1), 56–62.
- Ishaq, K., Mat Zin, N. A., Rosdi, F., Jehanghir, M., Ishaq, S., & Abid, A. (2021). Mobile-assisted and gamification-based language learning: A systematic literature review. *PeerJ Computer Science*, 7, e496. https://doi.org/10.7717/peerj-cs.496
- Kamasak, R., Ozbilgin, M., Atay, D., & Kar, A. (2020). *The Effectiveness of Mobile-Assisted Language Learning (MALL): A Review of the Extant Literature* (pp. 194–212). https://doi.org/10.4018/978-1-7998-4769-4.ch008
- Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, 20(3), 271–289. https://doi.org/10.1017/S0958344008000335
- Kurniati, M., Arafat, Y., & Mulyadi, M. (2020). Developing Teacher's Professionalism To Improve The Quality Of Education In Remote Areas. *International Journal of Educational Review*, *2*(2), 192–202.
- Lizamuddin, A., Asib, A., & Ngadiso, N. (2019). Indonesian English Learners' Perception of The Implementation of Mobile Assisted Language Learning in English Class. *Metathesis: Journal of English Language, Literature, and Teaching, 3*(1), 70–77.
- Luschei, T. F., & Zubaidah, I. (2012). Teacher training and transitions in rural Indonesian Schools: A case study of Bogor, West Java. *Asia Pacific Journal of Education, 32*(3), 333–350.
- Madhakomala, R., Hakim, M. A., & Syifauzzuhrah, N. (2022). Problem of Education In Indonesia and Alternative Solutions. *International Journal of Business, Law, and Education.*, *3*(2), 135–144.
- Mudra, H. (2018). Pre-Service EFL Teachers' Experiences in Teaching Practicum in Rural Schools in Indonesia. *The Qualitative Report*, 23, 319–344.
- Muttaqin, T. (2018). Determinants of Unequal Access to and Quality of Education in Indonesia. *The Indonesian Journal of Development Planning*, 2(1), 1–23. https://doi.org/10.36574/jpp.v2i1.27
- Nuraeni, C., Carolina, I., Supriyatna, A., Widiati, W., & Bahri, S. (2020). Mobile-Assisted Language Learning (MALL): Students' Perception and Problems towards Mobile Learning in English Language. *Journal of Physics: Conference Series*, 1641(1). https://doi.org/10.1088/1742-6596/1641/1/012027
- Pratama, A. R. (2018). Investigating Daily Mobile Device Use Among University Students in Indonesia. *IOP Conference Series: Materials Science and Engineering*, 325(1). https://doi.org/10.1088/1757-899X/325/1/012004
- Rajendran, T., & Yunus, M. M. (2021). A Systematic Literature Review on the use of Mobile-assisted Language Learning (MALL) for Enhancing Speaking Skills among ESL and EFL Learners. *International Journal of Academic Research in Progressive Education and Development*, *10*(1), Pages 586-609. https://doi.org/10.6007/IJARPED/v10-i1/8939
- Shadiev, R., Taoying, L., & Hwang, W.-Y. (2020). Review of research on mobile-assisted language learning in familiar, authentic environments. *British Journal of Educational Technology*, 51(3), 709–720. https://doi.org/10.1111/bjet.12839
- Solihin, S., & Sara, R. (2021). Gaps in the Education Management in Indonesia that are not Following Existing Laws in Indonesia. *Proceedings of the 1st International Conference on Law, Social Science, Economics, and Education, ICLSSEE 2021, March 6th 2021, Jakarta, Indonesia*. https://api.semanticscholar.org/CorpusID:236719039
- Sukasni, A., & Efendy, H. (2017). The Problematic of Education System in Indonesia and Reform Agenda. International Journal of Education, 9(3), 183–199. https://doi.org/10.5296/ije.v9i3.11705
- Sung, Y.-T., Chang, K., & Yang, J.-M. (2015). How Effective are Mobile Devices for Language Learning? A Meta-Analysis. *Educational Research Review*, *16*(6), 1–53. https://doi.org/10.1016/j.edurev.2015.09.001
- Tømte, C. (2015). Educating Teachers for the New Millennium? Teacher training, ICT and digital competence. Nordic Journal of Digital Literacy, 4(10), 138–154. https://doi.org/10.18261/ISSN1891-943X-2015-Jubileumsnummer-10
- Yudhiantara, R., & Nasir, I. A. (2017). Toward Mobile-Assisted Language Learning (MALL): Reaping Mobile Phone Benefits in Classroom Activities. *Register Journal*, *10*(1), 12. https://doi.org/10.18326/rgt.v10i1.12-28

## Learning Effectiveness of Integrating Peer Assessment and Board Games in a Computational Thinking and Artificial Intelligence Unit: Taking Speech Recognition as Examples

#### Mu-Sheng CHEN<sup>a\*</sup>, Ting-Chia HSU<sup>b</sup>

<sup>a, b</sup>Technology Application and Human Resource Development, National Taiwan Normal University, Taiwan \*mushengchen946@gmail.com

**Abstract:** This study aims to investigate the impact of peer assessment and educational board games on computational thinking (CT) and artificial intelligence (AI) courses using a two-factor experimental design. Combining educational technology, we designed an AI speech recognition course suitable for first-year students with no prior knowledge of CT or AI. The goal is to develop their foundational understanding and application of CT and AI. By employing educational board games, the study seeks to enhance learning enjoyment, increase students' self-efficacy in CT, and reduce their learning anxiety. Additionally, through the strategy of peer assessment, the study aims to provide benefits to students in both evaluating and being evaluated, thereby improving their learning outcomes and engagement.

**Keywords:** Computational thinking, artificial intelligence, educational board game, peer assessment, speech recognition

#### 1. Introduction

The rapid development of technology in today's world has significantly transformed our lives. The continuous advancements in technology have led to the widespread influence of AI which is impacting numerous domains (Schmidhuber, 2015). AI has profound implications for various professional fields, such as medical diagnosis, AI-powered job interviews, and big data analysis. The progression of these technologies has greatly altered our societal perceptions and perspectives. Consequently, it has become crucial to cultivate people's abilities to operate and comprehend these technologies, with CT being the fundamental and vital concept. In contemporary society, CT has become an essential skill for everyone to possess (Yadav et al., 2014).

Research suggests that game-based learning can promote and sustain students' motivation and engagement (Clark et al., 2016). Well-designed educational games can encourage deeper thinking among students. In recent years, many researchers have incorporated the concept of board games into education. Gee (2005) showed that by establishing clear learning goals, board games can be effective tools for promoting self-directed learning, problem solving, and deep learning. Due to the collaborative nature of board games, students can experience educational board games through social interaction, peer-to-peer engagement, and mutual learning (Wu et al., 2014). What's more, research on Al education indicates that visualizing and gamifying Al course content can effectively enhance students' interest in learning (DeNero & Klein, 2010).

The primary value of peer assessment lies in the close interaction among peers. In typical classroom settings, the number of students far exceeds the number of teachers, resulting in limited feedback received by each student. Peer assessment can address this issue by allowing students to provide feedback to one another. Research has found that peer feedback can be more direct and personalized compared to feedback from teachers (Topping, 2009). From another perspective, teacher feedback may be perceived by students as authoritative and non-negotiable, while peer feedback tends to be more diverse, fostering interactive discussions among peers (Cole, 1991).

This study aims to improve the situation of students with low self-efficacy in CT when they first start learning, using educational board games and peer assessment. Typically, low self-efficacy in CT is associated with increased learning anxiety and decreased engagement among students (Bandura, 1977; Pellas, 2014). Therefore, this study employs educational board games to enhance learning enjoyment, boost students' self-efficacy in CT, and reduce their learning anxiety. Additionally, the use of peer assessment strategies is expected to provide benefits to students in both evaluating and being evaluated, thus enhancing their learning outcomes and engagement. This study will address the following research questions:

- Are there significant differences in achievement in CT and AI learning among university students who take the four aforementioned courses using AI speech recognition?
- Are there significant differences in self-efficacy in CT among university students who take the four aforementioned courses using AI speech recognition?
- Are there significant differences in learning anxiety and engagement among university students who take the four aforementioned courses using AI speech recognition?

#### 2. Literature review

#### 2.1 The application of board games in education

According to the claims of Situated Learning Theory, "knowledge is a product of the interaction between learners and the context, deeply influenced by activities, social contexts, and culture." As stated by Lave and Wenger (1991), through authentic contexts that guide students' explorations, the most significant learning occurs when students engage in the process of exploration. Learning that arises from such experiences is profound and effective. Educational board games simulate contextual scenarios that enable students to explore and learn within those contexts. In the past few years, the use of board games has made it possible to address different types of learning, especially in the field of CT (Menon et al., 2019).

Educational board games can help students develop critical thinking, problem-solving skills, analytical abilities, reasoning, organization, planning, and communication skills (Hinebaugh, 2009). Compared to traditional learning, integrating educational board games into the curriculum can further enhance students' logical thinking, critical thinking, and deductive reasoning skills (Hinebaugh, 2009). In the past, many subjects such as mathematics, language, history, and science were taught using traditional methods. To enhance students' learning outcomes and spark their motivation and interest, educators have incorporated board game elements into their lessons, allowing students to engage in communication with others and improve their interpersonal skills. Additionally, Cheng (2018) used educational board games for language learning, and the results showed that the games effectively improved students' learning outcomes while reducing their learning anxiety.

#### 2.2 The application of peer assessment in education

Peer assessment is a mode of evaluation where learners assess each other, particularly those who have similar levels of learning. This approach has been successfully applied in elementary, middle, and high schools, including very young students, as well as those with special educational needs or learning disabilities (Ehly & Topping, 1998). Numerous studies have demonstrated that peer assessment can enhance the effectiveness and outcomes of learning (Cartney, 2010; Cartney & Rouse, 2006; Cassidy, 2006; Panadero et al., 2016; Topping et al., 2000). Importantly, both the assessors and the assessed benefit from peer assessment. Assessors can identify their own shortcomings by evaluating their peers' work, while also identifying areas for improvement in others (Yorke, 2003).

The primary value of peer assessment lies in the close interaction among peers. In typical classroom settings, the number of students far exceeds the number of teachers, resulting in limited feedback received by each student. Peer assessment significantly addresses this issue, and research has found that peer feedback can be more direct and personalized compared to feedback from teachers (Topping, 2009). Teacher feedback may be perceived by students as authoritative and non-negotiable, while peer feedback tends to be more diverse. Moreover, peer assessment promotes communication and discussion among students (Cole, 1991). Willey and Freeman (2006) proposed that peer assessment can effectively enhance learning outcomes and engagement, particularly in STEM-related courses.

#### 3. Research Methodology

To investigate the learning outcomes of college students using peer assessment in an Al speech recognition board game course, this study focuses on first-year university students. The instructional experiment is conducted in an Introduction to Computer Science course, with students aged between 18 and 19 years. The students' semester of enrollment is used as the basis for grouping, resulting in four groups. The experimental design follows a two-factor design. To minimize potential experiment interference from other factors, the course is taught by the same instructor, ensuring consistent teaching content, duration, and progression.

We designed an AI speech recognition course which includes topics such as an introduction to CT and AI concepts, explanation and application of speech recognition, personalization of recognition models, and subsequent practical applications, as shown in Fig. 1. The course incorporates CT, speech recognition, machine learning, and peer assessment through the BookRoll learning management content platform. Operational tools include the use of the MIT App Inventor platform and the Personal Audio Classifier (PAC) platform. The board game utilized is AI 2 Robot City. Each class session lasts for 3 hours, and instruction is conducted over a period of 2 weeks, resulting in a total of 6 hours of instruction.



Figure 1. Al speech recognition course.

Peer assessment will be conducted on the BookRoll platform, using an anonymous and randomly assigned approach. Only students who submit their assignments to the platform will be eligible to participate in the peer assessment, ensuring a complete matching of submissions. During the peer assessment process, students will receive an anonymous peer's assignment along with a peer assessment form. Assessing students will use the criteria specified in the peer assessment form to assign scores, and they can also provide qualitative feedback. To ensure the quality of peer feedback during peer assessment, teachers will explain the meaning of the scoring items and advise students to provide effective scoring explanations.

Al 2 Robot City combines three main concepts: CT, AI, and IoT. By utilizing the MIT App Inventor platform and PAC platform, players can train their own individualized speech recognition models. These models enable players to control the smart cars within the board game using their trained models. The process of training the models provides insights into the speech recognition process. Once the speech recognition models are trained, players can engage in the board game where they can cultivate the concepts of CT.

This study will utilize the CT Self-Efficacy Scale developed by Yağcı (2019), which consists of four dimensions: problem-solving, collaborative learning and critical thinking, creative thinking, and algorithmic thinking. Additionally, the study will employ the Learning Anxiety and Engagement Scale developed by Chang (2014), which includes three dimensions: anxiety, playfulness, and enjoyment. Both scales will be assessed using a 5-point Likert scale. Pretests and posttests will be conducted for achievement measures as well as the two scales. Subsequently, a two-factor analysis of covariance (ANCOVA) will be employed for data analysis.

#### References

- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review,* 84(2), 191-215. https://doi.org/10.1037/0033-295X.84.2.191
- Cartney, P. (2010). Exploring the use of peer assessment as a vehicle for closing the gap between feedback given and feedback used. *Assessment & Evaluation in Higher Education, 35*(5), 551-564. https://doi.org/10.1080/02602931003632381
- Cartney, P., & Rouse, A. (2006). The emotional impact of learning in small groups: highlighting the impact on student progression and retention. *Teaching in Higher education, 11*(1), 79-91. https://doi.org/10.1080/13562510500400180
- Cassidy, S. (2006). Developing employability skills: peer assessment in higher education. *Education* + *Training, 48*(7), 508-517. https://doi.org/10.1108/00400910610705890
- Cheng, Y.-C. (2018). The Effect Of Using Board Games In Reducing Language Anxiety And Improving Oral Performance. *Electronic Theses and Dissertations*, 899. https://egrove.olemiss.edu/etd/899
- Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of educational research*, *86*(1), 79-122. https://doi.org/10.3102/0034654315582065
- Cole, D. A. (1991). Change in self-perceived competence as a function of peer and teacher evaluation. *Developmental Psychology*, 27(4), 682–688. https://doi.org/10.1037/0012-1649.27.4.682
- DeNero, J., & Klein, D. (2010, July). Teaching introductory artificial intelligence with pac-man. In *First* AAAI Symposium on Educational Advances in Artificial Intelligence.
- Ehly, S. W., & Topping, K. J. (1998). *Peer-assisted learning.* L. Erlbaum Associates.
- Gee, J. P. (2005). Learning by design: Good video games as learning machines. *E-learning and Digital Media, 2*(1), 5-16. https://doi.org/10.2304/elea.2005.2.1.5
- Hinebaugh, J. P. (2009). A board game education. R&L Education.
- Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge university press.
- Menon, D., Romero, M., & Viéville, T. (2019). Computational thinking development and assessment through tabletop escape games. *International Journal of Serious Games, 6*(4), 3-18. https://doi.org/10.17083/ijsg.v6i4.319
- Panadero, E., Jonsson, A., Strijbos, JW. (2016). Scaffolding Self-Regulated Learning Through Self-Assessment and Peer Assessment: Guidelines for Classroom Implementation. In: Laveault, D., Allal, L. (eds) Assessment for Learning: Meeting the Challenge of Implementation. The Enabling Power of Assessment, vol 4. Springer, Cham. https://doi.org/10.1007/978-3-319-39211-0\_18
- Pellas, N. (2014). The influence of computer self-efficacy, metacognitive self-regulation and self-esteem on student engagement in online learning programs: Evidence from the virtual world of Second Life. *Computers in Human Behavior, 35*, 157-170. https://doi.org/10.1016/j.chb.2014.02.048
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks, 61*, 85-117. https://doi.org/10.1016/j.neunet.2014.09.003
- Topping, K. J. (2009). Peer assessment. *Theory into practice, 48*(1), 20-27. https://doi.org/10.1080/00405840802577569
- Topping, K. J., Smith, E. F., Swanson, I., & Elliot, A. (2000). Formative peer assessment of academic

writing between postgraduate students. *Assessment & Evaluation in Higher Education, 25*(2), 149-169. https://doi.org/10.1080/713611428

- Willey, K., & Freeman, M. (2006-02). Improving teamwork and engagement: The case for self and peer assessment. *Australasian Journal of Engineering Education*, 1-19. https://search.informit.org/doi/10.3316/aeipt.157674
- Wu, C.-J., Chen, G.-D., & Huang, C.-W. (2014). Using digital board games for genuine communication in EFL classrooms. *Educational Technology Research and Development*, 62(2), 209-226. https://doi.org/10.1007/s11423-013-9329-y
- Yadav, A., Mayfield, C., Zhou, N., Hambrusch, S., & Korb, J. T. (2014). Computational thinking in elementary and secondary teacher education. *ACM Transactions on Computing Education (TOCE)*, 14(1), 1-16. https://doi.org/10.1145/2576872
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher education, 45*(4), 477-501. https://doi.org/10.1023/A:1023967026413

# Exploring a Supportive Ecosystem for Creative Teaching in Higher Education

#### Weijing Han and Nur Aira Abdrahim Universiti Putra Malaysia, Malaysia 303679843@qq.com, nuraira@upm.edu.my

**Abstract:** This study aims to explore the construction of a supportive ecosystem in the development of teacher creativity within the education domain. The research primarily focuses on the emergence and development of teachers' individual creativity within the organizational context and their interaction with the organizational environment, in support of the teacher's professional development. As a result, the mechanisms underlying the occurrence of individual creativity within the organizational context are summarized, and supportive policies to foster teacher creativity development are proposed. This research has significant implications for the development of individual teacher creativity, professional development, and organizational creativity management.

#### Keywords: Ecosystem, creative teaching, higher education

#### 1. Introduction

#### 1.1 Background

In the era of rapid change and globalization, the field of education is facing unprecedented transformations and challenges. Particularly, since the outbreak of the COVID-19 pandemic, traditional teaching methods have been reshaped by modern technology and digital education. In this context, the ability of educators to adapt and embrace change, as well as creatively address challenges in educational practices, becomes crucial (Bruggeman et al., 2021).

In their professional careers, teachers encounter continuous learning and growth, and those possessing creativity are better equipped to adapt to the diverse demands of different subjects and teaching domains. They expand their teaching skills and educational perspectives through creative teaching practices and actively exploring new educational methods. This proactive approach helps teachers gain more opportunities and achievements in their professional development. Creativity is a skill that can be acquired through learning and experience (Kaufman & Sternberg, 2006).In order to foster teacher creativity, requisite and purposeful professional development is essential (Cropley & Patston, 2019).

Over the past few decades, the academic community has maintained a sustained enthusiasm for researching creative teaching. Scholars have primarily approached this issue from two perspectives: One aspect of research focuses on the implementation of creative teaching practices within the classroom. For example, researchers have explored the use of critical thinking cards (Holland & Ulrich, 2016), the integration of new teaching technologies (Madumarova et al., 2020) and the application of game-based approaches (McEnroe-Petitte & Farris, 2020) to foster creative teaching. On the other hand, there is an analysis of the influencing factors of creative teaching. These factors encompass personal aspects (Pashazadeh & Alavinia, 2019), leadership factors (Chen & Yuan, 2021), and environmental factors (Lapeniene & Dumciene, 2014). Scholars assert that a teacher's creative teaching is a part of the education system, which includes other subsystems such as school culture, leadership style, student characteristics, infrastructure, and even some broader collectives, like employment, admissions, social needs, and education policies. In this context, the creative behavior of teachers is not merely an individual act but requires maintaining a dynamic balance in the social environment (Zamana, 2022).

However, due to the context-specific nature of creativity (Cropley & Patston, 2019), the formation process and underlying mechanisms of creativity in the field of education still require further examination (Qian et al., 2019). Moreover, in practice, teachers continue to face formidable challenges in implementing creative teaching. Many educators lack a systematic understanding of creative teaching, and some even believe they do not possess creative

capabilities (Katzbuonincontro et al., 2020). Furthermore, the education sector still lacks a dedicated system for cultivating teacher creativity, and there is a notable absence of robust organizational support for teachers' creative teaching behavior (Borodina et al., 2019). These issues collectively impact the realization of creativity within educational organizations and also influence teachers' professional development prospects.

Therefore, this study aims to address the research gaps by adopting a managerial perspective, and investigating how creativity among teachers can be managed and fostered from an organizational standpoint. By doing so, it seeks to unveil the mechanisms through which teacher creativity is formed within the organizational context, promote the sustainable development of teacher creativity, and ultimately enhance teachers' professional growth as well as the continuous development of organizational creativity.

#### 1.2 Research Objectives

This study aims to explore the creativity process in higher educational institutions through investigating the teachers' everyday creative teaching experience, reveal teachers' creativity mechanism in organizations and finally establish an ecosystem for teachers' creative teaching. In particular, this research lend support to the notion of teacher's professional development by exploring on the following research questions: (1) How teachers establish their creative teaching behaviors in the educational context? (2) How teachers' creative teaching behaviors interplay with the organizational context? (3) How to foster teachers' creativity in the organizational context?

#### 1.3 Significance

This study carries both theoretical and practical significance within the body of knowledge. Theoretically, by exploring the mechanisms behind daily creative teaching, the study contributes to the broader understanding of creativity within educational settings. Practically, for individual teachers, it enables them to comprehend the process of creative teaching and provides timely assistance to enhance the efficiency of their creative teaching practices as part of their professional development. Second, for educational managers and administrators, the findings of this study provide valuable insights for developing supportive policies and creating an environment conducive to fostering creative teaching. This, in turn, helps cultivate a culture of sustained organizational creativity, leading to positive transformations and improvements within educational institutions.

#### 2. Literature Review

#### 2.1 The Concept of Creative Teaching

Previous scholars have mainly defined it in two ways. One view is that creative teaching refers to the creation and application of novel, new forms, innovative or creative teaching strategies to stimulate students' interest in learning (Gibson, 2010; Lee, 2011). Another way of thinking about creative teaching can also be defined as those behaviors that promote students' creativity (Liu et al., 2020). In this study, the creative teaching refers to the former.

#### 2.2 The Dynamic Componential Model of Creativity (DCMC)

The inception of the DCMC model originated from understanding the daily experiences of employees engaged in innovation projects within organizations (Amabile & Pratt, 2016). Similarly, teachers, being frontline practitioners in education, also integrate creative teaching as part of their daily work.

The model describes the process of creativity in organizations in terms of both organizations and individuals. The organizational innovation process consists of five stages:

Stage 1: Agenda setting. Stage 2: Stage setting. Stage 3: Producing ideas. Stage 4: Testing and implementing the ideas. Stage 5: Outcome assessment. The individual creative process also consists of five stages: Stage 1: Task presentation. Stage 2: Preparation. Stage 3: Idea generation. Stage 4: Idea validation. Stage 5: Outcome assessment.

#### 2.3 Theoretical Framework

This study attempts to employ the DCMC as the theoretical framework. Firstly, it intends to incorporate the five stages of both the organizational and individual levels from the DCMC model. Secondly, it analyzed the specific elements involved in creative teaching activities within the educational context. Finally, this study aims to analyze the mechanism of the teachers' creativity in the organizational context and propose the supportive ecosystem.



#### 3. Method

#### 3.1 Research Design

This study plans to employ a qualitative research approach and select a private university in China as the research case. In this research, creative teaching is considered a normalized teacher development activity, with a specific focus on cultivating creative teachers.

Specifically, this research intends to implement five stages of intervention measures based on the DCMC framework at the organizational level. These measures are aimed at fostering and supporting creative teaching among teachers within the selected private university. Additionally, at the individual level, five stages of intervention measures will be provided to support and enhance the creative teaching capabilities of individual teachers.

#### 3.2 Sampling

The participants will be provided by the faculty, which will offer a list of candidates participating in the creative teaching competition. Formal participants will be chosen after obtaining the candidates' consent.

#### 3.3 Data Collection

This study plans to employ three methods for data collection: interviews, observation memos, and reflective diaries. Interviews constitute the primary data collection method in this research and will involve both individual interviews and focus group interviews. Observation memos refer to the researcher's records during the research process. Reflective diaries will be maintained by the teachers, wherein they record their reflections and experiences throughout this process. By utilizing these three methods of data collection and employing triangulation, this research aims to enhance the trustworthiness and validity of the study.

#### 4. Conclusion and Future Plan

This paper discusses the following conclusions based on the aforementioned discussion above. Creative teaching is a daily manifestation of creativity within the organizational context and constitutes an integral component of teachers' professional development. Organizations can foster the management of creative endeavors by constructing a supportive ecosystem, thereby enabling sustainable professional development for their employees.

This is an undergoing Ph.D study for the first author. Currently, the data collection for this study has been completed. With enthusiasm and dedication, the researcher is eager to immerse in the data analysis phase and explore the possibilities it holds.

#### References

- Anderson, R. E. (1992). Social impacts of computing: Codes of professional ethics. *Social Science Computing Review*, *10*(2), 453-469.
- Borodina, T., Sibgatullina, A., & Gizatullina, A. (2019). Developing creative thinking in future teachers as a topical issue of higher education. *Journal of Social Studies Education Research*, *10*(4), 226-245.
- Bruggeman, B., Tondeur, J., Struyven, K., Pynoo, B., Garone, A., & Vanslambrouck, S. (2021). Experts speaking: Crucial teacher attributes for implementing blended learning in higher education. *The Internet and Higher Education*, 48.
- Chen, H.-H., & Yuan, Y.-H. (2021). The study of the relationships of teacher's creative teaching, imagination, and principal's visionary leadership. *SAGE Open, 11*(3).
- Cropley, D. H., & Patston, T. J. (2019). Supporting creative teaching and learning in the classroom: Myths, models, and measures. In C. A. Mullen (Ed.), *Creativity Under Duress in Education?* (pp. 267-288). Springer.
- Gibson, R. (2010). The 'art' of creative teaching: Implications for higher education. *Teaching in Higher Education*, *15*(5), 607-613.
- Holland, C., & Ulrich, D. (2016). Critical thinking cards: An innovative teaching strategy to bridge classroom knowledge with clinical decision making. *Teaching and Learning in Nursing*, *11*(3), 108-112.
- Katzbuonincontro, J., Perignat, E., & Hass, R. W. (2020). Conflicted epistemic beliefs about teaching for creativity. *Thinking Skills and Creativity*(36), 100651.
- Kaufman, J. C., & Sternberg, R. J. (2006). *The international handbook of creativity*. Cambridge University Press.
- Lapeniene, D., & Dumciene, A. (2014). Teachers' creativity: Different approaches and similar results. *Procedia - Social and Behavioral Sciences, 116*, 279-284.
- Lee, Y.-J. (2011). A study on the effect of teaching innovation on learning effectiveness with learning satisfaction as a mediator. *World Transactions on Engineering and Technology Education*, 9(2), 92-101.
- Liu, H. Y., Tsai, H. M., Wang, I. T., & Chen, N. H. (2020, May Jun). Predictors of self-perceived levels of creative teaching behaviors among nursing school faculty in Taiwan: A preliminary study. *J Prof Nurs*, 36(3), 171-176.
- Madumarova, M. K., Yusupova, A. M., Tilekova, N. D., Ismailova, B. T., Batyrkulova, A. B., & Batyrkulova, A. B. (2020). Modern ways of application of innovative teaching methods for the development of creative activity in the teaching process. In *Growth Poles of the Global Economy: Emergence, Changes and Future Perspectives* (pp. 37-47).
- McEnroe-Petitte, D., & Farris, C. (2020). Using gaming as an active teaching strategy in nursing education. *Teaching and Learning in Nursing*, *15*(1), 61-65.
- Pashazadeh, F., & Alavinia, P. (2019). Teacher creativity in light of autonomy and emotional intelligence. *Teaching English Language*, *13*(1), 177-203.
- Qian, M., Plucker, J. A., & Yang, X. (2019). Is creativity domain specific or domain general? Evidence from multilevel explanatory item response theory models. *Thinking Skills and Creativity*, 33, 100571.
- Zamana, F. (2022, Nov 10). The Future of Education as a Creative Ecosystem: A Sociocultural Framework for the Development of Creativity. *J Intell, 10*(4).

## Gamification to Train Soft Skills of Software Professionals Through Active Video Watching

#### Pasan PEIRIS\*

*University of Canterbury, Christchurch, New Zealand* \*pitigalagepasan.peiris@pg.canterbury.ac.nz

**Abstract:** Soft skills play a major role in the success of software development projects. However, learning soft skills is difficult. Video-based learning (VBL) provides self-paced and flexible learning but is often a passive method. Therefore, Active Video Watching (AVW) has been proposed to increase engagement and learning. We investigate gamification as an intervention to enhance the soft skill training of software professionals using AVW. The proposed research method includes both quantitative and qualitative approaches.

Keywords: Active video watching, gamification, soft skills, software professionals

#### 1. Introduction

Software development does not only require technical skills but also involves different processes and interactions with multiple stakeholders, such as clients, project managers, and development team members (Matturro et al., 2019). The skills required to fulfil non-technical activities are called soft skills. Soft skills can be defined as "intra- and inter-personal (socioemotional) skills, essential for personal development, social participation, and workplace success," which includes communication and teamwork (Kechagias, 2011). Furthermore, according to Gallagher et al. (2010), soft skills leverage technical skills. Since the production of software involves a human factor to some extent and soft skills have a major impact on a person's behaviour with others in a work environment, soft skills play a significant role in software projects to be successful (Ahmed et al., 2013; Mtsweni et al., 2016). However, learning soft skills is difficult and time-consuming (lorio et al., 2022).

Video-based learning (VBL) has attracted attention as a popular method of learning in recent years due to its self-paced, flexible way of learning (Giannakos et al., 2016; Pal et al., 2019). However, VBL can still be a passive learning activity (Seo et al., 2020). Recent research has found that the more the learners engage in various activities, the more learning they gather (Koedinger et al., 2015). Therefore, there have been significant attempts to increase engagement in video-based learning (Mitrovic et al., 2019).

Gamification is the use of game design elements in non-game contexts (Deterding et al., 2011). Since gamification can increase user engagement through intrinsic motivation (Hong & Masood, 2014), the education sector has been among the top domains where gamification has been researched to increase engagement and learning (Hamari et al., 2014). In this project, we investigate the use of gamification as an intervention in AVW to increase the learning of soft skills of software professionals. We will address the following research questions (RQ):

- 1. What are existing gamification approaches that have been proposed for learning soft skills?
- 2. What are the effective gamification approaches for software professionals to undertake active video-based training?

- 3. To what extent will gamification increase engagement and learning in active videobased training of software professionals?
- 4. What are the perceptions of software professionals when using gamified active video watching for development of soft skills?

Answering RQ1 and RQ2 will result in an empirically evaluated set of gamification elements that can be incorporated into training platforms for soft skills for software professionals. The magnitude of the effect of gamification on the software professionals' soft skills training will be evaluated through RQ3. Furthermore, from RQ4, the project will validate its findings through an industrial scenario by evaluating the perceptions of IT professionals. This will be beneficial in assessing how the gamified AVW fits within the organizational context setting and professional development plans of the professionals.

#### 2. Literature Review

#### 2.1 Gamification

The objective of incorporating gamification into a system is to increase the users' intrinsic and extrinsic motivation to get them involved in the tasks by integrating playful activities (Buckley & Doyle, 2014). Klock et al. (2020) state following game elements as the most used in tailored gamification: Customization (the ability to customize virtual space or character in a gamified system), Badges (a visual representation of an accomplishment), Challenges (a task that requires effort to complete), Levels (an element that supports user to track progression on system's purpose over time), Competition (an element that allows social comparison between users), Leader boards (an element that orders users according to a criterion).

There are few studies that look at gamification for professionals. There have been positive perceptions from professionals in using gamification in the software and healthcare industry (Lobin & Kevin Sungkur, 2023; Lemermeyer & Sadesky, 2016). For example, Stanculescu et al. (2016) present an experimental study where authors developed a web-based online application utilizing Work & Play gamification framework. They presented the gamified platform to IBM company employees for two months to identify how gamification can enable them to foster employee engagement with their day-to-day tasks. The platform was a learning and social interaction web application utilizing leader boards, badges, and scores as gaming elements. The authors conducted a between-subject, post-test with 206 employees by using application logs and a questionnaire to measure perceived engagement and perceived learning. The authors state that the level of engagement achieved via game elements was high, and the web-based application was successful in promoting knowledge acquisition. Leader boards and badges have positively affected user engagement, and the combination of the above game elements has been found to be more effective in enhancing the average session length a user spends on the platform (Stanculescu et al. 2016).

#### 2.2 Active Video Watching

AVW utilizes reflective activities such as writing comments, and rating comments in contrast to a traditional video-based learning platform (Lau et al., 2016). By integrating such activities, learners improve their engagement and learning while leveraging the experience from video-sharing social media websites such as YouTube (Dimitrova et al., 2021). One example implementation of AVW is AVW-Space, an online platform that facilitates interactive notetaking by enabling users to write comments about the content of the videos they watch and rate comments written by other learners on the same videos (Lau et al., 2016; Mitrovic et al., 2016).

There are few studies that explore AVW and soft skills: Mitrovic et al. (2017) present results from a study conducted with university students on presentation skills (a soft skill) in

AVW-Space. It found that learners who undertook commenting on videos and rating their peers' comments in AVW-Space have accumulated increased conceptual knowledge of the soft skill. Following that, more interventions have been developed to increase learning in AVW-Space, e.g., reminder nudges to motivate learners to write more comments and quality nudges to encourage higher-quality comments (Mitrovic et al., 2019; Mohammadhassan et al., 2022).

#### 3. Proposed Work and Research Method

This mixed-method study includes both qualitative and quantitative elements. We will use surveys, focus groups, control-experimental studies, and case study research methodologies. The proposed work will consist of six major research studies.

The first study was conducted as a preliminary survey with students, to collect insights on motivation, and perception on integrating gamification into active video watching. 103 respondents from New Zealand and the Philippines participated in this survey. Results show a positive perception towards integrating gamification into AVW with combinations of game elements instead of a specific element.

The second study is a systematic mapping study (SMS) to explore RQ1 and inform RQ2 as well. Based on the findings from the SMS, we will design the third study as a controlled experiment with software engineering students as proxies for entry-level software professionals. Selected gamification elements will be implemented in AVW-Space (an example AVW platform), and students will be given a fixed period to undertake soft skills training through the gamified platform. The control group will not interact with the gamification interventions. At the end of the study, data on the number of videos watched, comments written, comments rated on videos, and respective comment quality will measure learner engagement. A conceptual knowledge score obtained via questionnaires in pre-and post-tests will allow us to identify changes in learning. This will facilitate testing the developed interventions with students prior to the real study with software professionals.

The purpose of the fourth study is to validate the practical usability and improve the gamified intervention for software professionals. To do that, this study will be conducted as a focus group discussion with software professionals in the industry to get their feedback. The pre-tested gamified instance of the AVW-Space will be presented to these professionals, and the implementation will be critically analysed against different learning activities. This study will help answer RQ2.

The fifth study will be a similar experimental study to study three which will answer RQ3. Participants will be software professionals from the industry, recruited from crowdsourcing platforms (e.g., Amazon Mechanical Turk) to increase the number of participants. The purpose of this study is to analyse the engagement, and learning between control and experimental groups and identify whether gamification has a significant influence on the above factors. Based on the availability of professionals, we may have a more iterative single group approach to look at learning behaviours in AVW-Space as an alternative, since professionals might have multiple ways of achieving the learning key performance indicators.

The sixth and final study will be a case study conducted in a New Zealand software company with software professionals. Participants will undertake a soft skill training course in AVW-Space in their organizational environment. The purpose of this study is to collect professionals' perceptions on the usefulness of a gamified AVW platform for their professional development in a corporate organizational context. This study will help answer RQ4.

#### 4. Conclusions and Expected Outcomes

This research aims at achieving two major outcomes. At the end of the study, the authors will provide empirically grounded foundations for using gamification approaches to improve the training of soft skills of software professionals. This will be followed by the development of an

actual learning platform with gamification techniques, tested and validated for the use in organizational contexts.

#### References

- Ahmed, F., Capretz, L. F., Bouktif, S., & Campbell, P. (2013). Soft Skills and Software Development: A Reflection from Software Industry. International Journal of Information Processing and Management, 4(3), 171–191.
- Buckley, P., & Doyle, E. (2014). Gamification and student motivation. Interactive Learning Environments, 24(6), 1162–1175.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining Gamification. Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, 9-15. 10.1145/2181037.2181040
- Dimitrova, V., & Mitrovic, A.(2022). Choice Architecture for Nudges to Support Constructive Learning in Active Video Watching.International Journal of Artificial Intelligence in Education, 32(4), 892-930
- Gallagher, K. P., Kaiser, K. M., Simon, J. C., Beath, C. M., & Goles, T. (2010). The requisite variety of skills for IT professionals. Communications of the ACM, 53(6), 144–148.
- Giannakos, M. N., Sampson, D. G., & Kidziński, Ł. (2016). Introduction to smart learning analytics: Foundations and developments in video-based learning. Smart Learning Environments, 3(1), 12.
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work?A Literature Review of Empirical Studies on Gamification. 2014 47th Hawaii International Conference on System Sciences.
- Hong, G. E., & Masood, M. (2014). Effects of Gamification on Lower Secondary School Students' Motivation and Engagement. World Academy of Science, Engineering and Technology, International Journal of Educational and Pedagogical Sciences. 8(12), 3765–3772.
- Iorio, S., Cilione, M., Martini, M., Tofani, M., & Gazzaniga, V. (2022). Soft Skills Are Hard Skills-A Historical Perspective. Medicina, 58(8), 1044.
- K. Kechagias. 2011. Teaching and Assessing Soft Skills. MASS Project European Union's Lifelong Learning Programme.
- Koedinger, K., McLaughlin, E., Kim, J., Jia, J., & Bier, N. (2015). Learning is Not a Spectator Sport: Doing is Better than Watching for Learning from a MOOC. 111–120.
- Lau, L., Mitrovic, A., Weerasinghe, A., & Dimitrova, V. (2016). Usability of an active video watching system for soft skills training. In Proc. 1st Int. Workshop on Intelligent Mentoring Systems, ITS.
- Lemermeyer, G., & Sadesky, G. (2016). The Gamification of Jurisprudence: Innovation in Registered Nurse Regulation. Journal of Nursing Regulation, 7(3), 4–10.
- Lobin, S., & Kevin Sungkur, R. (2023). Gamification of Knowledge Management System: A Framework to Incentivize Knowledge Sharing for the Software Industry. In G. Ranganathan, X. Fernando, & S. Piramuthu (Eds.), Soft Computing for Security Applications (pp. 41–56). Springer Nature.
- Matturro, G., Raschetti, F., & Fontán, C. (2019). A Systematic Mapping Study on Soft Skills in Software Engineering. Journal of Universal Computer Science, 25(1), 16–41.
- Mitrovic, A., Dimitrova, V., Lau, L., Weerasinghe, A., & Mathews, M. (2017). Supporting Constructive Video-Based Learning: Requirements Elicitation from Exploratory Studies. Lecture Notes in Computer Science, 224–237.
- Mitrovic, A., Dimitrova, V., Weerasinghe, A., & Lau, L. (2016). Reflective Experiential Learning: Using Active Video Watching for Soft Skills Training. Proceedings of the 24th International Conference on Computers in Education, 192–201.
- Mitrovic, A., Gordon, M., Piotrkowicz, A., & Dimitrova, V. (2019). Investigating the Effect of Adding Nudges to Increase Engagement in Active Video Watching. In S. Isotani, E. Millán, A. Ogan, P. Hastings, B. McLaren, & R. Luckin (Eds.), Artificial Intelligence in Education (pp. 320–332). Springer International Publishing.
- Mohammadhassan, N., Mitrovic, A., & Neshatian, K. (2022a). Investigating the effect of nudges for improving comment quality in active video watching. Computers & Education, 176, 104340
- Mtsweni, E. S., Hörne, T., & Van Der Poll, J. A. (2016). Soft Skills for Software Project Team Members. International Journal of Computer Theory and Engineering, 8(2), 150–155.
- Pal, S., Pramanik, P. K. D., Majumdar, T., & Choudhury, P. (2019). A semi-automatic metadata extraction model and method for video-based e-learning contents. Education and Information Technologies, 24(6), 3243–3268.
- Seo, K., Fels, S., Yoon, D., Roll, I., Dodson, S., & Fong, M. (2020). Artificial Intelligence for Videobased Learning at Scale. Proceedings of the Seventh ACM Conference on Learning @ Scale, 215–217.

## An Empirical Study of the Computational Thinking Learning Game, "Online Robot City"

Tai-Ping HSU & Ting-Chia HSU\*

Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taiwan \*taipinshe@gmail.com; <u>ckhsu@ntnu.edu.tw</u>

**Abstract**: Having strong computational thinking (CT) skills is key for students studying computer science (CS) and STEM-related fields. However, relying too heavily on programming languages can hinder students' understanding and application of CT concepts. Therefore, to effectively promote the learning of CT, this study conducted a quasi-experimental design to integrate two different questioning strategies into the Online Robot City game for teaching purposes. The study aimed to explore the impact of using the 5W1H strategy and the context-based concept mapping strategy for information gathering and problem decomposition on students' achievement in CT and on their learning behavior patterns. The results are expected to respond to the importance of different learning approaches, in particular the context-based thinking method like the context-based concept mapping strategy proposed in this study, which may make relatively more contributions to CT learning achievement and self-efficacy. **Keywords:** Computational thinking, self-efficacy, 5W1H, context-based concept mapping strategy.

#### 1. Introduction

Computational thinking (CT) is a problem-solving skillset essential for the next generation. To help students acquire these skills, CT must be integrated into K-12 education (Avci & Deniz, 2022). CT is a widely used concept in K-12 computer science (CS) education, with different experts giving it diverse meanings. Possessing CT skills is vital for students to excel in CS and STEM fields (Sneider et al., 2014).

To teach CT in K-12, educators must create an environment that fosters students' CT skills while providing appropriate support and resources (Voogt et al., 2015). Programming using "Scratch" is effective in terms of enhancing students' understanding and application of CT concepts (Resnick et al., 2009), but teaching CT should also develop abstract thinking and algorithmic concepts (Voogt et al., 2015). This requires interdisciplinary education and the design of relevant teaching materials and activities. Brennan and Resnick (2012) argued

that CT should be a core discipline, not just a skill. They suggested that students learn computer science principles beyond programming to improve their CT skills. Lye and Koh (2014) cautioned that over-reliance on programming may hinder CT understanding. Many programs claim to teach programming skills, but few explore the thinking process used in CT (Shute et al., 2017). According to Lu and Fletcher (2009), teaching CT should use familiar concepts instead of programming languages to help students learn concepts like abstraction and algorithms.

Studies show that teaching CT through instructional strategies is effective. Kuo and Hsu (2020) suggested using interactive learning strategies to teach CT as they can improve learners' abilities. Hwang et al. (2015) found that teaching CT through game design improves CT abilities and STEM learning. These pedagogical strategies are widely used in K-12 education. Programming may be too difficult for some students, so it is important to develop appropriate ways to teach CT to students at different stages (Bers et al., 2014). Board games can be a suitable alternative.

Scholars have proposed strategies to cultivate students' CT abilities. Papert (1999) developed Logo programming language tools for teaching, advocating that students can develop CT and problem-solving abilities through games. Resnick et al. (2009) developed the Scratch programming language to teach CT through gamification and creativity. They advocated for students to learn by designing and sharing their own games and projects. The Online Robot City board game requires programming concepts include sequences, loops, conditionals, and functions (Garfield, 1994). These studies highlight the importance of teaching CT through games and programming to engage students in learning.

Board games can be enjoyable and useful for teaching, but their suitability varies among students due to their different learning styles and needs (Kebritchi, 2010). Playing board games alone may lack depth and breadth (Li, 2007), and factors like class size and material preparation can affect their effectiveness. Limited student interaction can decrease motivation to learn. To enhance motivation and learning effectiveness, increasing instructional interactivity is recommended (Newlin & Wang, 2002). Despite its advantages such as overcoming spatial limitations and promoting collaboration (Khorsandi et al., 2012), online teaching also has drawbacks such as technological constraints, lack of belongingness, anxiety, and reduced engagement (Xie et al., 2020). Regardless of the format, providing additional scaffolding and effective strategies to support learning CT is essential. Blended learning, combining online and classroom approaches, along with concept maps and the 5W1H strategy, can aid students in understanding and applying learning content, improving outcomes and motivation (Lin & Sekiguchi, 2020).

This study had two groups of young students (average age of 10) use two different questioning strategies. One questioning strategy was 5W1H while the other was context-based concept mapping. The participants adopted different questioning strategies before

997

using Online Robot City, which is an online CT game-based learning system to carry out structural programming logic to solve the tasks in the game. We wanted to explore the following two research questions by conducting an empirical study:

1) Did the students applying different questioning strategies for gathering information and problem decomposition before using structural programming logic to solve the tasks in the game make significant improvement in their CT learning achievement and self-efficacy?

2) Did the students applying different questioning strategies for gathering information and problem decomposition before using structural programming logic to solve the tasks in the game have different behavioral patterns?

Online Robot City provides an alternative and useful tool and method for cultivating the CT skills of students in a physical classroom. By comparing the two teaching strategies through the game, a better solution can be found for interpreting the online CT board game.

#### 2. Literature Review

#### 2.1. Computational Thinking

Learning CT is vital for today's students so they can use computer science concepts to solve problems and understand human behavior (Wing, 2006). Early CT training fosters independent thinking and problem solving. CT's applicability spans subjects and daily life. To adapt to the information society, students need both creative and digital literacy skills and effective CT and technology usage (Hsu et al., 2018). Programming access allows deep thinking and intellectual growth (Papert, 1980). Brennan and Resnick (2012) found that skills like sequences, loops, and conditionals transfer to other languages, and tools facilitating understanding of these concepts can cultivate CT abilities. Bers (2018) stressed visibility in CT tool design, enabling immediate feedback during the design process. Hsu and Chen (2022) proposed a CT board game as an interactive, visible way for novices to learn CT. Erdogan et al. (2022) indicated that board games require verification of each player's moves in order to ensure compliance with the rules and to identify winning strategies. However, when students' actions cannot be verified in real time, it can negatively impact the fairness and smoothness of the game. Therefore, an educational robot board game needed to be developed.

#### 2.2. Educational Robot Board Games

Educational robots (ERs) have gained popularity in classrooms, as they are considered as effective tools for fostering students' CT skills (Hsu et al., 2022). ERs are increasingly recognized as tools to develop CT competences (Chevalier et al., 2022). To apply CT practically, programming is usually used. However, this alone is insufficient for learning. Clear guidance is also necessary (Chevalier et al., 2022). Chevalier et al. (2022) suggested that there are few effective strategies for promoting "CT practice" and "CT perspectives," with most strategies focusing on "CT concepts."

Kuo and Hsu (2020) proposed that Educational Robot Board Games (ERBG) correspond to structural programming, including sequential structure, conditional structure, repetitive structure, and the modeling concept of calling a procedure in programming languages. Therefore, combining ER with board game to propose effective teaching strategies to promote the implementation of CT is worth researching.

#### 2.3. The 5W1H model

ERBG with shared gamified elements make learning fun, and instructors can add their content to the games. This type of instruction uses game rules like earning points, collecting badges, and leaderboards to motivate students, keep them engaged, and encourage communication. This approach leads to better learning results and satisfaction (Hsu et al., 2023). However, students may experience cognitive load when learning via ERBG (Sweller et al., 1990). Therefore, Hamborg et al. (2019) proposed the 5W1H strategy to help students understand the main events, describing who did what, when, where, why, and how. This simplifies and structures the entire event, making it easier to see and analyze. The usage environment, functions, and methods of the product are examined using the 5W1H method, which helps clarify the aim of promoting the potential use of ERBG in education (Yang et al., 2019).

However, even though the 5W1H strategy can effectively help students decompose problems and collect information about the context, functions, and methods of ERBG, parallel messages lack a linear sequence and causal relationships (Jinks & Jinks, 2019). Therefore, more contextual guidance strategies may be needed to deepen learning.

#### 2.4. Concept mapping

Novak and Gowin (1984) introduced concept maps as a visualization tool for knowledge networks. CT involves abstraction and automation (Weiwei et al., 2015), and concept maps can clarify knowledge relationships in teaching (Chen et al., 2021). However, CT often requires problem decomposition and algorithms (Lockwood & Mooney, 2017). Concept maps have limitations without structured patterns (Conceição et al., 2017). "Context-based concept mapping" integrates multiple aspects to enhance concept understanding (Ksibi et al., 2013). Cañas et al. (2012) emphasized that teachers' content and conditions impact students' learning outcomes. Research on context-based concept mapping to cultivate CT abilities and visualize cognitive processes is essential for students to effectively solve tasks in the game.

#### 2.5. Learning Behaviors

The cultivation of CT is a gradual process. Numerous studies have shown that through various courses, CT skills can be significantly improved. However, this cannot represent the true level of students' CT (Xu et al., 2019). CT involves utilizing fundamental concepts of computer science to solve problems, design systems, and understand human

behavior (Weiwei et al., 2015). Examining students' learning performance and comparing their learning behavior in different contexts is crucial to identify potential challenges and provide insights into curriculum implementation and support for instructors (Hsu et al., 2022). Chao (2016) mentioned that students' behavior and strategies of solving computational problems in a visual programming environment may affect their problem-solving performance. Therefore, it is necessary to further explore the behavioral patterns of novice programmers in a visual programming environment, and to investigate the differences in their strategies and performance in solving computational problems among different behavioral patterns. Hence, it is worth exploring various methods, such as learning behavior analysis, to examine CT ability.

#### 3. Method

#### 3.1. Participants

The study included 54 fourth- and fifth-grade students in northern Taiwan, of whom 27 learned CT with Online Robot City and context-based concept mapping, and 27 learned with Online Robot City and 5W1H thinking. A teacher with over 10 years of experience guided the groups. Students worked in pairs during the game.

#### 3.2. Instructional Design

The study used Online Robot City, a board game involving robot city construction. Two groups of students used different CT questioning strategies: context-based concept mapping and the 5W1H approach. The experimental group integrated context-based concept mapping before employing structural programming logic in the game (Figure 1).





Figure 1 context-based concept map.

#### Figure 2 5W1H strategies

The control group used the 5W1H approach for gathering information and problem analysis, followed by structural programming in the game (Figure 2). Before the study, both groups received instructions on control card use and robot operation, including game rules and CT concepts. The game allowed students to practice concepts like repetition and conditional statements using repeat and condition cards. Students had to control their robots to obtain resources on the map, using up to eight control cards at a time to control the

number of steps. Precise landing on correct positions was necessary for scoring. The game encouraged teamwork, objective identification, planning, and logical thinking, promoting CT learning in pairs and during competitions.

#### 3.3. Research Process

A quasi-experimental design with robot teaching was implemented for 1 hour/week over 10 weeks. Students whose average age is 10 completed pretest surveys, then were split into experimental and control groups. Both groups participated in the same teaching sessions by the same teacher in the same classroom. Posttest surveys and behavior analysis followed. Both groups spent the same amount of time.

#### 3.4. Instruments

We designed a CT achievement test based on the Bebras International CT Test (Dagiene & Stupuriene, 2016), using a learning achievement test to assess students' CT abilities. The CT test on sequential concepts, repeated concepts, and selection included 34 multiple-choice questions. A technology education teacher with over 10 years of experience in CT education and programming was invited to verify the CT test. The two scholars ensured the reliability of the test and ensured that the questions containing CT characteristics were related to the learning content of CT board games. To reduce testing effects, the order of items in the pretest and posttest was switched, and additional questions were added to the posttest. Figure 3 is an example of a Bebras competition, involving the concept of flowcharts in structured programming and the traveling robot problem. Figure 4 is an example of the pretest design for this study, covering the learning content of the CT chess game. Figure 5 is an example of the posttest design for this study, covering the learning content of the CT chess game.



Question: How many cards are needed for the car to move from A to B (as shown on the right)? Choice 1: 1 piece. Choice 2: 2 pieces. Choice 3: 3 pieces.

*Figure 3* Bebras International CT Test (Dagiene & Stupuriene, 2016) *Figure 4* Example of a pretest design for this study.



Figure 5 Example of a posttest design for this study.

To evaluate students' self-efficacy in computational thinking, this study used the Computational Thinking Scale (CTS) developed by Korkmaz et al. (2017), using a 5-point

Likert scale (1 = "strongly disagree", 5 = "strongly agree") to assess students' creativity, algorithmic thinking, collaboration, and critical thinking. The Cronbach's alpha value was 0.82, indicating satisfactory reliability.

#### 3.5. Learning Behavior Analysis

This study explores the CT learning behavior of students in ERBG, based on actual operations recorded of student behavior (e.g. moving ER after peer discussion). Through system recording, analysis of learning behavior is conducted and the analysis content is encoded. The encoding is then converted into GSEQ format data and analyzed using Sequence Analysis GSEQ Version 5.0 to explain whether there is significant correlation between behaviors.

#### 4. Proposed Contribution

The teaching design of this study was divided into three stages: sequential, repetitive, and conditional structures. The study aims to observe the impact of different questioning strategies on students' CT learning during the teaching process. The study is expected to echo the findings of Hsu et al. (2022), who found that students made significant progress in CT through teaching-based methods and customized educational robot activities. By analyzing students' learning behavior, the study hopes to observe students' CT behavior and misconceptions, such as the inability to use algorithmic cards. Teachers can use teaching prompts to correct students' misconceptions. Overall, the study aims to transform students' CT learning from focusing on programming language features to logical structures and thinking strategies.

#### Reference

- Avcı, C., & Deniz, M. N. (2022). Computational thinking: early childhood teachers' and prospective teachers' preconceptions and self-efficacy. *Education and Information Technologies*, 27(8), 11689-11713.
- Bakeman, R., Quera, V., & Gnisci, A. (2009). Observer agreement for timed-event sequential data: A comparison of time-based and event-based algorithms. *Behavior Research Methods*, 41, 137-147.
- Bers, M. U. (2018). Coding and computational thinking in early childhood: The impact of ScratchJr in Europe. *European Journal of STEM Education*, *3*(3), 8.
- Bers, M. U., Flannery, L., Kazakoff, E. R., & Sullivan, A. (2014). Computational thinking and tinkering: Exploration of an early childhood robotics curriculum. *Computers & Education*, 72, 145-157.
- Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. Proceedings of the 2012 annual meeting of the American educational research association, Vancouver, Canada,

- Cañas, A. J., Novak, J. D., & Reiska, P. (2012). Freedom vs. restriction of content and structure during concept mapping-possibilities and limitations for construction and assessment.
- Chao, P.-Y. (2016). Exploring students' computational practice, design and performance of problem-solving through a visual programming environment. *Computers & Education*, *95*, 202-215.
- Chen, B.-H., Chen, L.-H., Chen, C.-C., & Chen, C.-H. (2021). Influence of Programming Robot Learning Mode Based on Conceptual Mapping on College Students' Computational Thinking. 2021 IEEE 3rd Eurasia Conference on IOT, Communication and Engineering (ECICE),
- Chevalier, M., Giang, C., El-Hamamsy, L., Bonnet, E., Papaspyros, V., Pellet, J.-P., Audrin, C., Romero, M., Baumberger, B., & Mondada, F. (2022). The role of feedback and guidance as intervention methods to foster computational thinking in educational robotics learning activities for primary school. *Computers & Education*, 180, 104431.
- Conceição, S. C., Samuel, A., & Yelich Biniecki, S. M. (2017). Using concept mapping as a tool for conducting research: An analysis of three approaches. *Cogent Social Sciences*, *3*(1), 1404753.
- Dagiene, V., & Stupuriene, G. (2016). Bebras--A Sustainable Community Building Model for the Concept Based Learning of Informatics and Computational Thinking. *Informatics in education*, *15*(1), 25-44.
- Erdogan, A., Sunay, M. A., & Çevirgen, A. E. (2022). How to Play Board Games? A Framework Proposal for Classroom Settings. *Board Game Studies Journal*, *16*(2), 101-122.

Garfield, R. (1994). RoboRally. In: Wizards of the Coast Renton.

- Hamborg, F., Breitinger, C., & Gipp, B. (2019). Giveme5w1h: A universal system for extracting main events from news articles. *arXiv preprint arXiv:1909.02766*.
- Hsu, T.-C., Chang, C., & Liang, Y.-S. (2023). Sequential Behavior Analysis of Interdisciplinary Activities in Computational Thinking and EFL Learning With Game-Based Learning. *IEEE Transactions on Learning Technologies*, *16*(2), 256-265.
- Hsu, T.-C., Chang, C., Wu, L.-K., & Looi, C.-K. (2022). Effects of a Pair Programming
  Educational Robot-Based Approach on Students' Interdisciplinary Learning of
  Computational Thinking and Language Learning. *Frontiers in Psychology*, 13, 888215.
- Hsu, T.-C., Chang, S.-C., & Hung, Y.-T. (2018). How to learn and how to teach computational thinking: Suggestions based on a review of the literature. *Computers & Education*, *126*, 296-310.
- Hsu, T.-C., & Chen, M.-S. (2022). The engagement of students when learning to use a personal audio classifier to control robot cars in a computational thinking board game. *Research and Practice in Technology Enhanced Learning*, *17*(1), 1-17.

- Hwang, G.-J., Chiu, L.-Y., & Chen, C.-H. (2015). A contextual game-based learning approach to improving students' inquiry-based learning performance in social studies courses. *Computers & Education*, *81*, 13-25.
- Jinks, T., & Jinks, T. (2019). The 5W1H Method. *Psychological Perspectives on Reality, Consciousness and Paranormal Experience*, 41-44.
- Kebritchi, M. (2010). Factors affecting teachers' adoption of educational computer games: A case study. *British Journal of Educational Technology*, *41*(2), 256-270.
- Khorsandi, M., Kobra, A., Ghobadzadeh, M., Kalantari, M., & Seifei, M. (2012). Online vs. traditional teaching evaluation: a cross-sectional study. *Procedia-Social and Behavioral Sciences*, 46, 481-483.
- Korkmaz, Ö., Çakir, R., & Özden, M. Y. (2017). A validity and reliability study of the computational thinking scales (CTS). *Computers in human behavior, 72,* 558-569.
- Ksibi, A., Ammar, A. B., & Amar, C. B. (2013). Enhanced context-based query-to-concept mapping in social image retrieval. 2013 11th International Workshop on Content-Based Multimedia Indexing (CBMI),
- Kuo, W.-C., & Hsu, T.-C. (2020). Learning Computational Thinking Without a Computer: How Computational Participation Happens in a Computational Thinking Board Game. *The Asia-Pacific Education Researcher*, *29*(1), 67-83. https://doi.org/10.1007/s40299-019-00479-9
- Li, M.-H. (2007). Lessons learned from web-enhanced teaching in landscape architecture studios. *International Journal on E-Learning*, *6*(2), 205-212.
- Lin, J., & Sekiguchi, T. (2020). E-learning in entrepreneurship education: A systematic literature review. 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE),
- Lockwood, J., & Mooney, A. (2017). Computational Thinking in Education: Where does it fit? A systematic literary review. *arXiv preprint arXiv:1703.07659*.
- Lu, J. J., & Fletcher, G. H. (2009). Thinking about computational thinking. Proceedings of the 40th ACM technical symposium on Computer science education,
- Lye, S. Y., & Koh, J. H. L. (2014). Review on teaching and learning of computational thinking through programming: What is next for K-12? *Computers in human behavior*, 41, 51-61.
- Newlin, M. H., & Wang, A. Y. (2002). Integrating technology and pedagogy: Web instruction and seven principles of undergraduate education. *Teaching of Psychology*, 29(4), 325-330.
- Novak, J. D., & Gowin, D. B. (1984). Learning how to learn. cambridge University press.
- Papert, S. (1980). Personal computing and its impact on education. *The computer in the school: Tutor, tool, tutee,* 197-202.
- Papert, S. (1999). What is Logo? Who needs it. Logo philosophy and implementation, 4-16.

- Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K.,
  Millner, A., Rosenbaum, E., Silver, J., & Silverman, B. (2009). Scratch: programming for all. *Communications of the ACM*, *52*(11), 60-67.
- Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational research review*, 22, 142-158.
- Sneider, C., Stephenson, C., Schafer, B., & Flick, L. (2014). Computational thinking in high school science classrooms. *The Science Teacher*, *81*(5), 53.
- Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of experimental psychology: general*, 119(2), 176.
- Voogt, J., Fisser, P., Good, J., Mishra, P., & Yadav, A. (2015). Computational thinking in compulsory education: Towards an agenda for research and practice. *Education and Information Technologies*, 20, 715-728.
- Weiwei, C., Zhigang, L., Aihua, B., & Yanqin, T. (2015). Curriculum design for computational thinking training based on concept map. 2015 7th International Conference on Information Technology in Medicine and Education (ITME),

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33-35.

- Xie, X., Siau, K., & Nah, F. F.-H. (2020). COVID-19 pandemic–online education in the new normal and the next normal. *Journal of information technology case and application research*, *22*(3), 175-187.
- Xu, L., Tong, M.-W., Li, B., Meng, J., & Fan, C.-Y. (2019). Application of concept map in the study of computational thinking training. 2019 14th International Conference on Computer Science & Education (ICCSE),
- Yang, C.-M., Liu, T.-H., Zheng, Y.-Y., & Man, H.-T. T. (2019). A Study on Design Thinking Based Creative Product Design Process in a Design Project. *International Journal of Systematic Innovation*, 5(3).

## Integrating Explainable Artificial Intelligence in Active Video Watching

Raul Vincent LUMAPAS<sup>a\*</sup>

<sup>a</sup>Computer Science and Software Engineering Department, University of Canterbury, New Zealand \*raulvincent.lumapas@pg.canterbury.ac.nz

Abstract: The use of videos in learning has increased over the past years. Along with the popularity of video-based learning is the surge in the interest in Artificial Intelligence in education. Previous studies explored the use of Artificial Intelligence technologies in Active Video Watching, a form of video-based learning. A particular case of Artificial Intelligence in Active Video Watching would be Active Video Watching (AVW)-Space, a video-based learning platform developed by the University of Canterbury. The use of AI in AVW-Space, for example, in assessing the guality of comments made by users, has resulted in an increase in student engagement and learning. Student feedback in recent surveys on the use of Active Video Watching showed an interest in explanations of how the system's AI makes decisions. A way to integrate explanations to the system is through Explainable Artificial Intelligence (XAI). Therefore, this research aims to provide additional insights into the use of XAI and explanations in education and professional training through active video watching. This research also aims to explore the potential of XAI as a way to increase user engagement and learning when using Al-supported features of active video watching systems. A second goal is to look at currently implemented AI models and identify potential points of improvement in the AI / ML models used in active video watching.

**Keywords:** Explainable Artificial Intelligence, Active Video Watching, Videobased Learning

#### 1. Introduction

Rapid advancements in the field of Artificial Intelligence (AI) (Mullins & Conati, 2020) have dramatically increased the interest in AI and the use of AI applications and tools in diverse domains over the past decade (Fiok et al., 2022). With this surge of interest in AI come questions related to the Fairness, Accountability, Transparency, and Ethics (FATE) of AI. As AI is also being used to support high-stake and critical human decisions, such as in the medical field. Hence, there is a clear need for FATE in AI systems, mainly as AI models are used to make important predictions and decisions. These concerns propelled the interest in explanations in AI systems, which in turn increased the interest in the field of Explainable Artificial Intelligence (XAI). Explainable Artificial Intelligence uses different methods to "enable end users to understand, appropriate trust, and effectively manage the emerging generation of artificially intelligent partners" (Gunning & Aha, 2019, p.45).

The use of XAI can also be beneficial in providing explanations in educational systems (Khosravi et al., 2022). Despite the increasing use of AI in educational systems (AIED), the integration of XAI has been particularly low in Intelligent Tutoring Systems (ITS) as compared to other areas, such as recommender systems, office assistants, and interactive systems (Conati et al., 2020). The same is true with video-based learning and active video watching.

This need for explanations has become apparent in Active Video Watching (AVW)-Space, an online video-based learning platform developed at the University of Canterbury. AVW-Space supports engagement in video-based learning and is aimed at learning soft skills. In recent surveys done on AVW-Space, a number of users have indicated the need for explanations, particularly in features that supports student engagement through the use of AI. One of the features that use Machine Learning (ML) models assesses the quality of comments students write while watching videos. The use of AI and ML models has increased student engagement (Mohammadhassan et al., 2022). With this increase in engagement, we observed students wondering how AI is used in the system, particularly how AI decides on the comment quality.

This research will focus on three different areas in integrating XAI to active video watching. The first area focuses on determining the XAI desiderata and identifying the various needs, interests, goals, demands, and expectations of different AVW users. The second focus area of the research is on how XAI can be utilized to identify potential improvements or changes in implemented machine learning models in AVW. The third focus area of the research is on integrating XAI into active video watching by designing explanations suitable for end users of AVW. This research aims to answer the following research questions:



Figure 1. Project Framework and Research Questions

#### 2. Literature Review

#### 2.1 Active Video Watching

Video-based learning provides positive effects on students' learning experiences, such as flexibility and opportunities for self-regulated learning, increased student attention, and increased motivation to learn (Chatti et al., 2016; Dimitrova & Mitrovic, 2022; Sablić et al., 2021). The use of video-based learning has also shown promise in learning soft skills, as it allows the contextualization of personal experiences and shows real-life experiences (Cronin & Cronin, 1992; Dimitrova & Mitrovic, 2022). However, despite these benefits and potential, the lack of interaction between students and teachers in video-based learning, which may result in passive learning and a low level of engagement, becomes a challenge. A way to address this challenge is to support student engagement through active video watching here learners conduct activities while watching videos. One example platform that implements active video watching is AVW-Space.

This VBL platform supports engagement in video-based learning and is aimed at learning soft skills. AVW-Space allows teachers to create a space and integrate videos sourced from YouTube. Using the space created by their teacher, students can watch videos, write comments, and anonymously rate their peers' comments. Students can only review or rate their peers' anonymous comments once the teacher enables rating on the platform. AVW-Space contains different features, such as personalized prompts called Nudges and Visualizations. For example, as soon as a student writes a comment, AVW-Space automatically analyses it and displays the comment quality. The comment quality is determined by an AI model developed in previous research (Mohammadhassan, 2022). Student feedback from previous studies shows that there is a demand on explaining how the AI works in AVW-Space, for example how the comment quality is determined. Mohammadhassan et al. also proposed that the AI models used can be improved further and potentially improve users' engagement with the system by explaining the decisions made.

#### 2.2 Explainable Artificial Intelligence in Education (XAI-ED)

Building trust with stakeholders in educational systems is important, given their distinct needs and the nature of data. Furthermore, the need for trust and accountability in educational systems has increased due to challenges arising from the use of AI. Issues related to accountability, learner autonomy and agency, support for learner metacognitive processes and reflective processes, and broader issues related to authentic assessment, credentialing, and academic integrity highlight has highlighted the need for explanations (Khosravi et al., 2022). This shows the necessity and importance of XAI in education systems.

In recent years, there has been an increase in research towards XAI in Education or XAI-ED. Khosravi et al. (2022) provided a framework that characterizes the nature of XAI in education in terms of questions about these six key aspects: (1) the people involved (stakeholders); (2) the benefits to each group; (3) how to deliver the explanation; (4) the widely used classes of models used in education; (5) the human-centered design of the AI and interfaces to support explanation; and (6) the potential pitfalls of providing explanations. This framework can help and guide researchers in AIED in studying, designing, and developing XAI systems for education and answer important questions related to the key aspects of XAI-ED. Some work has been done in the area of XAI-ED. This has been highlighted in the case studies investigated by Khosravi et al. (2022) illustrating the use of XAI-ED in a range of AIED systems. The case analysis showed the potential of XAI in different forms of AIED systems.

Utilizing XAI can help explain AI's behavior and achievements in educational systems (Fiok et al., 2022). However, there is still a great need to explore explanations of AI in education, given the increasing use of AI in the domain. Moreover, Fiok et al. indicate that XAI can also be designed in a way that the system allows to achieve goals essential from the professional training perspective. The use of XAI has been echoed by Khosravi et al. (2022) as well, explaining that the benefits of XAI is not just applicable in education systems but also in human capital development as well. Given the potential of XAI in education and professional development systems, the present research investigates its application in learning soft skills using active video watching.

#### 3. Proposed Work

This research focuses on the integration of XAI in active video watching. AVW-Space will be used as the platform to implement this research. As indicated, there will be three different focus areas for this research with each part having different research activities and studies. The first focus area will be on determining the XAI desiderata for active video watching. Insights from this area will motivate and guide the implementation of studies related to the other two succeeding focus areas. Given the possible complexity of explanations, understandable and timely explanations are needed to help learners understand how the system uses AI and its decisions. As there are different needs and demands by AVW users, it is important to determine the XAI desiderata. An important part of this research is to look at the needs and demands of AVW users, in particular formal education users and professional development and training users.

The second focus area will explore possible improvements to the currently implemented models, potentially integrating explanation elements to pinpoint gaps or issues in the models or features used for the models. The goal is to further improve the currently used models or identify other ways of integrating AI in AVW. In the case of AVW-Space, the models used in assessing the quality of comments were shown to be effective in enhancing engagement and in nudging users to write more high-quality comments (Mohammadhassan, 2022). However, the performance of the models can be enhanced further (Mohammadhassan et al., 2022). Given the complexity of the various features used to assess comment quality, integrating explanations would benefit platform users. Using XAI can potentially help identify ways to improve or redesign the currently implemented ML models or can help in identifying the potential need to integrate and implement new AI methods or models. XAI can be utilized

to to look at currently implemented AI models and identify potential points of improvement in the AI / ML models used in active video watching systems.

The third focus area will be on designing and integrating explanations for users of AVW. Results from studies on the first focus area will serve as a guide in designing explanations in the system. Careful investigation of the explanation designed out of the explanation desiderata and requirements will be investigated. The appropriate XAI techniques and tools will then be used in providing explanations for the AI-features in the system. Furthermore, these explanations will be designed or presented in a way that is understandable by users of the system.

Improvements from the implementing the XAI design and improvements in the ML models will then be integrated into the AVW-Space platform. The improved AVW platform will then be tested with formal education and professional development and training users. This test will evaluate improvements in user engagement, learning, and satisfaction. Other potential metrics will also be used to identify improvements in the use of AVW, such as user trust and confidence in the system.

We did some initial work for the first part of the research, particularly in determining the XAI desiderata for Active Video Watching. We have also integrated some improvement in the ML models in AVW-Space. We conducted a study to help in setting the XAI desiderata for active video watching through a question-driven process for XAI. Users of AVW-Space were asked questions they have regarding the AI features of AVW-Space. A paper on this entitled *"Question-Driven Design Process for Explainable AI in Active Video Watching"* has been submitted as a Work-in-Progress paper in the International Conference on Computers in Education (ICCE). Another study has also been conducted to discuss the changes in the ML models for the Communication Skills space. This paper on the ML models entitled *"Evaluating the Assessment of Comment Quality in Learning Communication Skills in AVW-Space"* is also submitted as an Extended Summary for ICCE 2023.

To summarize, the ultimate goal of the research is for software engineering practitioners to learn soft skills (such as Communication and Presentation skills) using active video watching that utilizes AI features and integrates explainability using XAI. Further work will be done to identify how explanations can be implemented to support the learning of different types of users, in this case students and software engineering professionals, and how XAI can further help in improving active video watching systems.

#### References

- Chatti, M. A., Marinov, M., Sabov, O., Laksono, R., Sofyan, Z., Fahmy Yousef, A. M., & Schroeder, U. (2016). Video annotation and analytics in CourseMapper. Smart Learning Environments, 3(1), 10. https://doi.org/10.1186/s40561-016-0035-1
- Conati, C., Barral, O., Putnam, V., & Rieger, L. (2020). Toward Personalized XAI: A Case Study in Intelligent Tutoring Systems (arXiv:1912.04464). arXiv. http://arxiv.org/abs/1912.04464
- Cronin, M. W., & Cronin, K. A. (1992). Recent empirical studies of the pedagogical effects of interactive video instruction in "soft skill" areas. Journal of Computing in Higher Education, 3(2), 53–85. https://doi.org/10.1007/BF02942356
- Dimitrova, V., & Mitrovic, A. (2022). Choice Architecture for Nudges to Support Constructive Learning in Active Video Watching. International Journal of Artificial Intelligence in Education, 32(4), 892–930. https://doi.org/10.1007/s40593-021-00263-1
- Fiok, K., Farahani, F. V., Karwowski, W., & Ahram, T. (2022). Explainable artificial intelligence for education and training. The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, 19(2), 133–144. https://doi.org/10.1177/15485129211028651
- Gunning, D. & Aha, D. (2019). DARPA's Explainable Artificial Intelligence (XAI) program. AI Magazine, 40(2), 44–58.
- Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable Artificial Intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074. https://doi.org/10.1016/j.caeai.2022.100074
- Mohammadhassan, Negar. Supporting engagement in active video watching using quality nudges and visualisations. [Doctoral Thesis, University of Canterbury] (2022)
- Mohammadhassan, N., Mitrovic, A., & Neshatian, K. (2022). Investigating the effect of nudges for improving comment quality in active video watching. Computers & Education, 176, 104340. https://doi.org/10.1016/j.compedu.2021.104340
- Mullins, R., & Conati, C. (2020). Enabling understanding of AI-Enabled Intelligent Tutoring Systems. In Design Recommendations for Intelligent Tutoring Systems: Volume 8—Data Visualization (pp. 141–148). U.S. Army Combat Capabilities Development Command – Soldier Center.
- Sablić, M., Mirosavljević, A., & Škugor, A. (2021). Video-Based Learning (VBL)—Past, Present and Future: An Overview of the Research Published from 2008 to 2019. *Technology, Knowledge and Learning*, 26(4), 1061–1077. https://doi.org/10.1007/s10758-020-09455-5

### Supporting Interactive Learning in Active Video Watching

#### Ehsan BOJNORDI

University of Canterbury, Christchurch, New Zealand ehsan.bojnordi@pg.canterbury.ac.nz

**Abstract:** The potential of Active Video Watching (AVW) to train soft skills is evident. Despite considerable achievements in learning via AVW, the capacity of interactive learning has yet to be explored. This PhD research aims to investigate and support interactive learning in AVW via (a) interaction among learners, and (b) intelligent and adaptive interaction between the system and learners (comment recommendations). To this end, I initially provided learners with the possibility of responding to other learners' comments on videos and calling attention to them by Like/Dislike rating mechanism. In the second phase, a comment recommendation engine will be developed, based on the learner model and a specific-domain ontology with respect to the target soft skill. I will also augment comment recommendations in hope of enhancing learners' engagement, interactions, and learning.

Keywords: Active video watching, interactive learning, recommender system, ontology

#### 1. Introduction and Literature Review

Video-Based Learning (VBL) is increasingly used in informal training as well as in formal settings (Yousef et al., 2014). Learning through video watching has multiple advantages, such as providing flexible, self-controlled learning opportunities, and raising learners' motivation (Chatti et al., 2016; Sablić et al., 2021). It can also pave the way for learning transferable skills (Anthony & Garner, 2016; Mitrovic et al., 2016). However, merely viewing videos is a passive activity that leads to limited engagement and shallow learning (Chi & Wylie, 2014). Another issue with most VBL systems is that they treat all learners the same way, while it would be more beneficial for learners if they could be adaptive and apply appropriate methods, such as collaborative learning, micro-teaching, and student-centered learning, tailored for diverse learners (Yousef et al., 2014).

According to the ICAP framework (Chi & Wylie, 2014), learners' behaviors are classified into four types of learning modes: *Interactive, Constructive, Active* and *Passive* based on their overt behaviours. In this classification, the learning effectiveness is reduced in the following order: *Interactive> Constructive> Active > Passive*. To learn effectively from videos, learners need to engage with video content. This is possible by incorporating some activities with video, called Active Video Watching (AVW). Numerous studies have been recently conducted to investigate the significance of and improvements in AVW (Dimitrova & Mitrovic, 2022; Gostomski et al., 2019; Mitrovic et al., 2016; Mitrovic et al., 2019; Mohammadhassan et al., 2022; Piotrkowicz et al., 2018).

AVW-Space, as a notable example of AVW platforms, provides a wide range of supports for learners' engagement including commenting on videos, visualization of learners' activities, reviewing and rating comments, proposing micro-scaffolds or aspects, and embedding intelligent nudges to prompt learners to exhibit constructive behaviors, like writing high-quality comments (Dimitrova et al., 2017; Mitrovic et al., 2017; Mitrovic et al., 2016; Mitrovic et al., 2019). However, interactive behaviours like direct interactions among learners have not been employed in AVW. As interactive mode of learning has a great learning effectiveness in general (Chi & Wylie, 2014), the purpose of this research is to investigate and support it in AVW, as described in more detail in the following section.

#### 2. Project Details and Methodologies

This PhD project aims to cover both aspects of interactions - human-human and humancomputer - in AVW. The research question for the first study is as follows.

• **RQ1:** What are the effects of using an interactive form of comment reviewing/rating on learners' engagement and learning in AVW?

To answer this question, I provided learners with an interactive activity of replying to peers' comments in AVW-Space (a platform that implements AVW) so that they could discuss their views. This way, learners can share their views on video content, see opinions and reflections of their peers, and ultimately contribute to *social learning* (the philosophy that people can learn from each other through observation, imitation and modeling (Bandura & Walters, 1977)). To provide direct interaction between students, I implemented a reviewing/rating approach in which students could Like/Dislike comments, but also respond to comments as well as previously written responses in two nested levels (level 1: respond to comment, level 2: respond to response), as shown in Figure 1.

Student actions	
Student Actions » Space Instance: Presentation Skills » Review v	deo comments: TUTORIAL 1: How to Give an Awesome (PowerPoint) Presentation
eview video comments for "TUTORIAL 1: H	ow to Give an Awesome (PowerPoint) Presentation"
(A) How to Give an Awesome (RowerPoint) Presentat	Comments to rate
OWERFUL CONSTRUCTIONS	02:00 Think that this is a good point. Sometimes I find myself planning for the test of the corversation, rather than taking in what the speaker is actually saying
	🛁 30 🤫 0 🖘 Roply Emposes (1)
	02:09 Using hand gestures is something id like to do more of
Notch on Diverse mply Stated	ult 16 🖷8 🗮Reply Impound(t)
	Convention and In my views, having a good table of content or overview page can be quite helpful.
	ado mit mikepty Basponer(d)
	try to use body language in your presentation
	201 ALL 1977

Figure 1. A screenshot from reviewing comments in AVW-Space

In former instantiations of AVW-Space, as for rating comments, choices were limited to pre-defined rating categories and learners were only able to choose the closest one to their thoughts. They were unable to express their views directly. Instead, I added a Like/Dislike rating and the possibility of expressing their views through responding to comments. The number of Likes and Dislikes can imply the majority opinion about proposed ideas. I employed this numerable rating mechanism for both comments and responses to let individuals engage with and feel encouraged to contribute to one another's perspectives. Learners can raise positive and negative points about existing comments, which will bring about separate consequences in discussion (Norris et al., 2010), i.e., negative feedback has a stronger impact than positive feedback on attention, perception, and motivation (Norris, 2021).

Furthermore, learners in previous studies (Mohammadhassan & Mitrovic, 2022a, 2022b) reported that they were given a large number of comments to rate, which are time-consuming and overwhelming to review. All learners received the same list of comments, some of them repetitive or uninformative. These issues indicate that the review phase in AVW is not adaptive to diverse learners. Hence, the main objective of my second study is to develop a new component that can offer adaptive comment recommendations. I aim to design an ontology-based learner model and leverage a comment recommendation engine to cover adaptively the areas of weakness of a given learner. With that in mind, the following research question was formulated:

• **RQ2:** Is there a causal effect of recommending relevant comments on learners' engagement and learning?

To implement a recommender engine, it is necessary to assess learners' knowledge. I will create a learner model based on a comprehensive domain-specific ontology, as employing ontology facilitates content recommendation and navigational sequencing in learner model (Mitrovic et al., 2007; Vesin et al., 2013). Hence, I will initially create an ontology for a soft skill (e.g., empathy skills). Along with the ontology, I will utilize a classifier like Bayesian Classifier (BC) to generate and update the learner model, as BC is highly effective in text categorization (Go et al., 2009; Pedregosa et al., 2011). Before watching videos, the classifier will receive the learner's responses on a survey. Then, it will receive information about the activities the learner performs, such as commenting on videos, responding to/rating comments or responses. Recommendations of comments will be made to cover learners' areas of weakness (the concepts they have not yet written about, or those for which learner shave written only low-quality comments). The learner model will be updated when the learner writes a response to or rates a comment.

The transparency of recommendations (or expressing the rationale behind the recommendations) plays a crucial role in individuals' trust and satisfaction (Barria-Pineda et al., 2021; Kulesza et al., 2013). Therefore, in the third study, I aim to investigate the effects of leveraging explainable comment recommendations on learners' engagement and learning outcomes. Hence, the third research question is:

• **RQ3:** Do explanations of comment recommendations give rise to more engagement, interactions and learning among learners?

Two forms of explanations (verbal and visual) will be provided. I will employ the learner model to foster these two forms of explanations. Explanations reveal the reason(s) why a comment is recommended to the learner. According to the learner model, the areas of weakness will initially be identified, and then a personalized list of comments along with reasons (explanations) will be created. The system will generate short paragraphs for each unit of verbal explanation indicating the target concept and the reason for recommending a given comment. Visual explanations correspond to the main concepts in the ontology. The number of visual elements is the same as the number of concepts of ontology that the learner has not learned. Each recommended comment will be accompanied by these visual explanations (labeled to these parts of ontology) shown with varying intensity of colours according to the number of relevant phrases included in the comment.

All aforementioned approaches will be implemented in AVW-Space and tested in several experiments on different soft skills and learner groups, including university students and ICT professionals, to evaluate the generalizability of my research. Each experiment will contain two phases. Initially, learners will complete Survey 1, consisting of questions on demographic, training, and experiences with a target soft skill, as well as a knowledge test. Then, they can watch and comment on videos in AVW-Space. A number of comments will be selected manually to be shown (in study 1) or recommended by the system (in studies 2 & 3) for learners to review, rate, and respond to. At the end, learners will complete Survey 2, containing a knowledge test plus some open-ended questions about the usefulness of our approaches and modifications.

#### References

Anthony, S., & Garner, B. (2016). Teaching soft skills to business students: An analysis of multiple pedagogical methods. *Business and Professional Communication Quarterly*, 79(3), 360-370.
 Bandura, A., & Walters, R. H. (1977). *Social learning theory* (Vol. 1). Englewood cliffs Prentice Hall.

- Barria-Pineda, J., Akhuseyinoglu, K., Želem-Ćelap, S., Brusilovsky, P., Milicevic, A. K., & Ivanovic, M. (2021). Explainable recommendations in a personalized programming practice system. Artificial Intelligence in Education: 22nd International Conference, AIED 2021, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part I,
- Chatti, M. A., Marinov, M., Sabov, O., Laksono, R., Sofyan, Z., Fahmy Yousef, A. M., & Schroeder, U. (2016). Video annotation and analytics in CourseMapper. *Smart Learning Environments*, *3*(1), 1-21.

- Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist*, 49(4), 219-243.
- Dimitrova, V., & Mitrovic, A. (2022). Choice architecture for nudges to support constructive learning in active video watching. *International Journal of Artificial Intelligence in Education*, *32*(4), 892-930.
- Dimitrova, V., Mitrovic, A., Piotrkowicz, A., Lau, L., & Weerasinghe, A. (2017). Using learning analytics to devise interactive personalised nudges for active video watching. Proceedings of the 25th conference on user modeling, adaptation and personalization,
- Go, A., Huang, L., & Bhayani, R. (2009). Twitter sentiment analysis. *Entropy*, 17, 252.
- Gostomski, P., Gordon, M., Herritsch, A., Dimitrova, V., & Mitrovic, A. (2019). Teaching presentation skills in first year engineering curriculum using active video watching and personalised nudges. In Chemeca 2019 (48th: 2019: Sydney, NSW) (pp. 254-265). Engineers Australia Sydney, NSW.
- Kulesza, T., Stumpf, S., Burnett, M., Yang, S., Kwan, I., & Wong, W.-K. (2013). Too much, too little, or just right? Ways explanations impact end users' mental models. 2013 IEEE Symposium on visual languages and human centric computing,
- Mitrovic, A., Dimitrova, V., Lau, L., Weerasinghe, A., & Mathews, M. (2017). Supporting constructive video-based learning: requirements elicitation from exploratory studies. Artificial Intelligence in Education: 18th International Conference, AIED 2017, Wuhan, China, June 28–July 1, 2017, Proceedings 18,
- Mitrovic, A., Dimitrova, V., Weerasinghe, A., & Lau, L. (2016). Reflective experiential learning: Using active video watching for soft skills training. Proceedings of the 24th international conference on computers in education,
- Mitrovic, A., Gordon, M., Piotrkowicz, A., & Dimitrova, V. (2019). Investigating the effect of adding nudges to increase engagement in active video watching. Artificial Intelligence in Education: 20th International Conference, AIED 2019, Chicago, IL, USA, June 25-29, 2019, Proceedings, Part I 20,
- Mitrovic, A., Martin, B., & Suraweera, P. (2007). Intelligent tutors for all: Constraint-based modeling methodology, systems and authoring.
- Mohammadhassan, N., & Mitrovic, A. (2022a). Discovering differences in learning behaviours during active video watching using epistemic network analysis. Advances in Quantitative Ethnography: Third International Conference, ICQE 2021, Virtual Event, November 6–11, 2021, Proceedings 3,
- Mohammadhassan, N., & Mitrovic, A. (2022b). Investigating the Effectiveness of Visual Learning Analytics in Active Video Watching. International Conference on Artificial Intelligence in Education,
- Mohammadhassan, N., Mitrovic, A., & Neshatian, K. (2022). Investigating the effect of nudges for improving comment quality in active video watching. *Computers & Education*, *176*, 104340.
- Norris, C. J. (2021). The negativity bias, revisited: Evidence from neuroscience measures and an individual differences approach. *Social neuroscience*, *16*(1), 68-82.
- Norris, C. J., Gollan, J., Berntson, G. G., & Cacioppo, J. T. (2010). The current status of research on the structure of evaluative space. *Biological psychology*, *84*(3), 422-436.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, *12*, 2825-2830.
- Piotrkowicz, A., Dimitrova, V., Mitrovic, A., & Lau, L. (2018). Using the explicit user profile to predict user engagement in active video watching. Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization,
- Sablić, M., Mirosavljević, A., & Škugor, A. (2021). Video-based learning (VBL)—past, present and future: An overview of the research published from 2008 to 2019. *Technology, Knowledge and Learning*, *26*(4), 1061-1077.
- Vesin, B., Ivanović, M., Klašnja-Milićević, A., & Budimac, Z. (2013). Ontology-based architecture with recommendation strategy in java tutoring system. *Computer Science and Information Systems*, 10(1), 237-261.
- Yousef, A. M. F., Chatti, M. A., & Schroeder, U. (2014). The state of video-based learning: A review and future perspectives. *International Journal on Advances in Life Sciences*, *6*(3), 122-135.

## The design and use of conversational intelligent tutoring systems and computer simulation for the use of students of technology entrepreneurship

Joseph Benjamin ILAGAN<sup>a\*</sup>

<sup>a</sup>Ateneo de Manila University, Philippines \*jbilagan@ateneo.edu

Abstract: Entrepreneurship is complex and dynamic. It involves continuously pursuing novel or better products and business models amidst constraints, uncertainty, and constant change among ecosystem participants (or "agents"). Entrepreneurship education, therefore, needs to be non-linear. Yet, traditional teaching methods in entrepreneurship came from business management education practices: lectures, case studies, and group discussions-mostly ineffective because entrepreneurship is more dynamic and non-linear. Recent entrepreneurial experiential learning attempts include starting and running a business and using computer simulations to reduce time and cost. There are opportunities to introduce non-linear and more human-like approaches to the learning interface, and these are some of the aims of intelligent tutor systems (ITS). This study proposes using a conversational ITS (CITS) as the learning experience interface for a technology entrepreneurship program to teach students various concepts. Conversations, through natural language, will take advantage of recent developments in large language models (LLMs) and related conversational agents and AI assistants such as ChatGPT. At the heart of the learning tool is a suite of computer simulation environments specifically for technology entrepreneurship, with the choice of technology entrepreneurship forcing novelty and relative market uncertainty in product offerings. The design and selection of technologies will follow evaluation frameworks on the effectiveness of entrepreneurship teaching simulation environments: fidelity, verification, and validity. The expected output will be a simulation environment resulting from multiple design-build-implement iterations. The CITS and the simulation core engine shall interface with a Learning Management System (LMS). The study will also generate insights after simulation sessions with domain experts and students through educational data mining (EDM) of the resulting logs.

**Keywords:** conversational intelligent tutoring systems, CITS, large language models, Ilm, GPT, computer simulation, entrepreneurship education, agent-based modeling, agent-based simulation, ABMS

#### 1. Introduction

Entrepreneurial education is crucial in fostering economic growth, yet it lacks many aspects to make it more effective (Allegra et al., 2013). Entrepreneurship is a complex, chaotic, nonlinear, and dynamic process of vision, change, and creation; therefore, it is more than business creation and management (Neck & Greene, 2011). The typical teaching methods in entrepreneurship classes are those used in business: lectures, case studies, and group discussions–all ineffective when teaching entrepreneurship, which requires more active learning approaches (Tasnim & Yahya, 2013). Entrepreneurship education has expanded with business plans, consultation with practicing entrepreneurs, and computer simulations (Kuratko, 2005). This study focuses on computer simulation with a conversational interface. A *simulation* is a self-contained immersive environment in which the learner interacts within the environment to learn or practice skills or knowledge. Simulations allow cost-effective and less time-intensive training than a real (business) scenario (Kapp et al., 2014). Certain simulations also induce the state of flow (Csikszentmihalyi, 1988) during sessions (Almeida & Buzady, 2019). This study aims to design and observe the usage of a computer simulation environment intended for technology entrepreneurship students. The choice of technology entrepreneurship implies novelty and market uncertainty in product offerings. Agent-based Modeling and Simulation (ABMS) covers interactions among independent agents (Macal & North, 2005). An agent is an autonomous individual that can behave and decide independently. Simulation environments may be facilitated through human teachers or computers. Conversational Intelligent Tutoring Systems (CITS) are e-learning systems that deliver tutorial content through discussion, asking and answering questions, identifying gaps in knowledge, and providing feedback in natural language (Holmes et al., 2017). Generative artificial intelligence (GAI) is a machine learning framework that generates content using probability and statistics based on existing digital content such as text, video, images, and audio through training examples, thus learning their patterns and distribution (Baidoo-Anu & Owusu Ansah, 2023). A large language model (LLM) is a GAI and statistical model of the distribution of tokens in the vast public corpus of human-generated text (Shanahan, 2023). Generative Pre-trained Transformer (GPT), an LLM-based system, generates or statistically predicts sequences of words, code, or other data, starting from a source input called the prompt (Floridi & Chiriatti, 2020). GPT is based on the Transformer (Vaswani et al., 2017) deep neural network architecture. A conversational agent or AI assistant based on LLM is ChatGPT, a conversational interface to GPT (Zhai, 2022). A common pattern nowadays is turning an LLM into a question-answering system by a) embedding it in a larger system and b) using prompt engineering to elicit the expected behavior (Shanahan, 2023). This paper is a major enhancement from the previous submission (Ilagan, 2022). The change is the focus on using an LLM-based CITS approach instead of a non-conversational interface.

The following questions properly establish the significance of the study: RQ1) What features must be present when designing and implementing conversational intelligent tutoring systems and simulation platforms for entrepreneurship education? RQ2) How would the efficacy of conversational intelligent tutoring systems and agent-based simulation modeling be measured based on learning outcomes? RQ3) How may a conversational intelligent tutoring system complement a set of human instructors and mentors for entrepreneurship education? How can intelligent tutoring systems, through advances in conversational support of generative AI, be designed to enhance and spark interest in entrepreneurship education? What are the limitations of an LLM, and how can these gaps be addressed in the CITS that embodies it? The hypotheses are as follows: 1) Incorporating regular feedback and guestion prompting in conversations will lead to improved learning outcomes over traditional methods of teaching entrepreneurship. To test this hypothesis, students' assessment results using a non-conversational front-end vs. a conversational front-end will be administered and analyzed at the end of the learning session. 2) An ABMS incorporating realistic and interactive entrepreneurial scenarios will enhance students' understanding and application of entrepreneurship principles, leading to improved learning outcomes over traditional lectures and case studies, 3) A CITS that complements human instructors and mentors in entrepreneurship education will improve learning outcomes by providing additional support and guidance to students. 4) A simulator will spark interest in learning technology entrepreneurship.

#### 2. Related Work

Neck & Greene (2011) offer an approach involving entrepreneurship as a *method* in contrast to a *process*, which implies predictability. Some methods are starting a business, design-based learning, and reflective practice. Another set of methods uses simulations, multimedia instruction, and interactive activities to compact the business creation process (Tasnim & Yahya, 2013). Hindle (2002) finds that experiential learning is a vital component of entrepreneurship education but also raises the issue involving the timeframe of a business venture exceeding an academic subject's term duration and how simulation can help. Work

involving ABMS involves economic activities such as supply chain recovery (Rahman et al., 2021) and Rogers's (1983) theory of diffusion of innovations (Meade & Islam, 2006). ABMS software like NetLogo (Tisue & Wilensky, 2004) and Python-based MESA (Kazil et al., 2020) are free, open-source, come with example models, are customizable, and are extensible. Instruction Design (ID) and tutorials involving ABMS exist but have been scarce, as pointed out by Kasaie & Kelton (2015), whose work will serve as one of the frameworks and guidelines (including discretion) for developing the simulator of this study. Feinstein & Cannon (2002) propose an evaluation framework for entrepreneurship simulations with three criteria: *fidelity* (realism), *verification* (operates as intended), and *validity* (the model and the natural world reach the same conclusions). This study aims to fill the gaps in work on the following: 1) customizing ABMS software to support additional entrepreneurial concepts, 2) extending the model to illustrate competition, 3) integrating these ABMS platforms with LMS for educational data mining (EDM), 4) using LLMs for CITS, and 5) sustaining interest in entrepreneurship through CITS and simulation.

#### 3. Research and Implementation Method

The target concept for the back-end simulator illustration is the theory of diffusion of innovation. The agent model parameters are based on the work of Schramm et al. (2010) and simplified rules from Bass (1980). Text related to concepts of diffusion of innovation shall be fed as a knowledge base into the Python-based framework LangChain (Chase, 2022) to prepare the LLM conversational environment through a form of prompt engineering. The CITS will follow existing frameworks for conversational flow while allowing free-flowing conversations with the chatbot, given unforeseen and rare situations. During development, five (5) student volunteers from the B.S. Information Technology Entrepreneurship (BS ITE) program in the Ateneo de Manila University (Philippines) will use and provide feedback iteratively. At least 10 student technology startup groups from BS ITE, each with an average of five (5) co-founders, are expected to volunteer to try the simulator environment. At least one person per startup group is expected to participate. Each volunteer will be randomly assigned to the non-conversational or conversational simulator environment. Participants may follow this general path: 1) Complete a pre-session survey including questions about the student's background, interests, and previous entrepreneurship experience, with additional questions on the learning experience (ex., outcomes, sentiments); 2) Engage with the simulation environment to learn concepts related to technology entrepreneurship. 3) Complete a post-session survey to provide feedback on experience and understanding of the concepts taught in the simulation. Students will have reflection papers at the start, middle, and end of the semester, with only the volunteers answering questions on their experience with the CITS, indicating interest levels before and after the sessions, and submitting their insights through the LMS, after which EDM will be done. The study will take at least two cycles, with additional ones enhanced based on new insights. Early prototypes will use NetLogo and then migrate to MESA. Before roll-out, at least two domain experts knowledgeable in diffusion of innovations will provide feedback based on fidelity and validity.

#### Acknowledgements

I thank my doctoral dissertation advisor, Maria Mercedes T. Rodrigo, for the guidance and mentorship throughout the whole process of writing this paper.

#### References

Allegra, M., La Guardia, D., Ottaviano, S., Dal Grande, V., & Gentile, M. (2013). A serious game to promote and facilitate entrepreneurship education for young students. *Proceedings of the 2013 International Conference on Education and Educational Technologies*, 256–263.
- Almeida, F., & Buzady, Z. (2019). Assessment of entrepreneurship competencies through the use of FLIGBY. *Digital Education Review*, 151–169.
- Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Available at SSRN 4337484*.
- Bass, F. M. (1980). The relationship between diffusion rates, experience curves, and demand elasticities for consumer durable technological innovations. *Journal of Business*, S51–S67.
- Chase, H. (2022). *LangChain* [Python]. https://github.com/hwchase17/langchain (Original work published 2022)
- Csikszentmihalyi, M. (1988). Optimal experience. 3–14. https://doi.org/10.1017/cbo9780511621956.001
- Feinstein, A. H., & Cannon, H. M. (2002). Constructs of simulation evaluation. *Simulation & Gaming*, 33(4), 425–440. https://doi.org/10.1177/1046878102238606
- Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30, 681–694.
- Hindle, K. (2002). A grounded theory for teaching entrepreneurship using simulation games. *Simulation & Gaming*, 33(2), 236–241. https://doi.org/10.1177/1046878102332012
- Holmes, M., Latham, A., Crockett, K., & O'Shea, J. D. (2017). Tomorrow's Learning: Involving Everyone. Learning with and about Technologies and Computing, 11th IFIP TC 3 World Conference on Computers in Education, WCCE 2017, Dublin, Ireland, July 3-6, 2017, Revised Selected Papers. *IFIP Advances in Information and Communication Technology*, 515, 251–260. https://doi.org/10.1007/978-3-319-74310-3\_27
- Ilagan, J. B. (2022). The design and use of agent-based modeling computer simulation for teaching technology entrepreneurship.
- Kapp, K. M., Blair, L., & Mesch, R. (2014). *The Gamification of learning and instruction. Ideas into practice*. San Francisco: Wiley.
- Kasaie, P., & Kelton, W. D. (2015). Guidelines for design and analysis in agent-based simulation studies. 2015 Winter Simulation Conference (WSC), 183–193.
- Kazil, J., Masad, D., & Crooks, A. (2020). Utilizing python for agent-based modeling: The mesa framework. Social, Cultural, and Behavioral Modeling: 13th International Conference, SBP-BRiMS 2020, Washington, DC, USA, October 18–21, 2020, Proceedings 13, 308–317.
- Kuratko, D. F. (2005). The Emergence of Entrepreneurship Education: Development, Trends, and Challenges. *Entrepreneurship Theory and Practice*, *29*(5), 577–597. https://doi.org/10.1111/j.1540-6520.2005.00099.x
- Macal, C. M., & North, M. J. (2005). Tutorial on agent-based modeling and simulation. *Proceedings of the Winter Simulation Conference*, 2005, 2–15. https://doi.org/10.1109/wsc.2005.1574234
- Meade, N., & Islam, T. (2006). Modelling and forecasting the diffusion of innovation A 25-year review. *International Journal of Forecasting*, 22(3), 519–545. https://doi.org/10.1016/j.ijforecast.2006.01.005
- Neck, H. M., & Greene, P. G. (2011). Entrepreneurship Education: Known Worlds and New Frontiers. Journal of Small Business Management, 49(1), 55–70. https://doi.org/10.1111/j.1540-627x.2010.00314.x
- Rahman, T., Taghikhah, F., Paul, S. K., Shukla, N., & Agarwal, R. (2021). An agent-based model for supply chain recovery in the wake of the COVID-19 pandemic. *Computers & Industrial Engineering*, 158, 107401. https://doi.org/10.1016/j.cie.2021.107401
- Rogers, E. M. (1983). Diffusion of innovations.
- Schramm, M. E., Trainor, K. J., Shanker, M., & Hu, M. Y. (2010). An agent-based diffusion model with consumer and brand agents. *Decision Support Systems*, *50*(1), 234–242.
- Shanahan, M. (2023). *Talking About Large Language Models* (arXiv:2212.03551). https://doi.org/10.48550/arXiv.2212.03551
- Tasnim, R., & Yahya, S. (2013). Playing entrepreneurship: Can games make a difference. *Entrepreneurial Practice Review*, 2(4), 4–16.
- Tisue, S., & Wilensky, U. (2004). Netlogo: A simple environment for modeling complexity. International Conference on Complex Systems, 21, 16–21.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, \Lukasz, & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*.
- Zhai, X. (2022). ChatGPT user experience: Implications for education. Available at SSRN 4312418.



#### ATENEO DE MANILA UNIVERSITY SCHOOL OF SCIENCE AND ENGINEERING Department of Information Systems and Computer Science

Aug 14, 2023

Alwyn Vwen Yen LEE PC Executive Chair, Doctoral Student Consortium ICCE 2023

Dear Dr. Lee:

In my capacity as Mr. Joseph Benjamin Ilagan's dissertation adviser, I support his application for participation in the ICCE 2023 Doctoral Student Consortium. Mr. Ilagan is a student pursuing a Ph.D. in Computer Science degree at the Ateneo de Manila University in the Philippines.

Mr. Ilagan participated in the 2022 DSC and has made substantial changes to his dissertation proposal since then. He is expected to complete all doctoral degree requirements, including dissertations defense, as early as May 2025.

Sincerely,

Madre

Maria Mercedes T. Jodrigo, Ph.D. Professor, Department of Information Systems and Computer Science Head, Ateneo Laboratory for the Learning Sciences International Program Committee Chair, ICCE 2021

Ateneo Laboratory for the Learning Sciences, Rob 208 Faura Hall, Ateneo de Manila University

## The design and use of conversational intelligent tutoring systems and computer simulation for the use of students of technology entrepreneurship

Author	:	Joseph Benjamin ILAGAN	
Affiliation	:	Ateneo de Manila University, Philippines	
Contact	:	jbilagan@ateneo.edu	
		+63 9178193568	
Advisor	:	Maria Mercedes T. Rodrigo	
		Ateneo de Manila University, Philippines	
		mrodrigo@ateneo.edu	
DSC Theme	:	C1: ICCE Sub-Conference on Artificial Intelligence in Education/Intelligent	
		Tutoring System (AIED/ITS) and Adaptive Learning	

## Learning Habits Mining and Data-driven Support of Building Habits in Education

Chia-Yu HSU<sup>a\*</sup>, Izumi HORIKOSHI<sup>b</sup>, Rwitajit MAJUMDAR<sup>b</sup> & Hiroaki OGATA<sup>b</sup>

<sup>a</sup>Graduate School of Informatics, Kyoto University, Japan <sup>b</sup>Academic Center for Computing and Media Studies, Kyoto University, Japan \*hsu.chiayu.25t@st.kyoto-u.ac.jp

**Abstract:** In terms of the learning logs accumulating from real-world educational activities, techniques of learning analytics help understand the characteristics of learners' behaviors. However, extracting learning habits from log data and data-driven support for building learning habits have not yet attracted much attention. Therefore, this research proposes the approach of "Learning Habits Mining", which aims to extract the types and stages of learning habits from learners' daily learning logs and support learners to build learning habits with data-driven methods. We identify two contributions of this research. First, this research reveals the learning habits of K12 learners and provides an approach to trace the process of building learning habits automatically. Second, this research proposes interventions to the data-driven support for building learning habits so that learners can build learning habits based on evidence derived from learning logs.

Keywords: Learning Habits, Daily Learning Logs, Real-World Data, Learning Analytics

#### 1. Introduction

With the advancement of ICT, data is accumulating from real-world educational activities (Ogata et al., 2018). Take Japan for example. The Ministry of Education, Culture, Sports, Science, and Technology (MEXT) initiated the GIGA School Program (Global Innovation Gateway for All) to ensure "one computer for one student". The program allows learning in schools to be done anywhere and anytime (The Government of Japan, 2021). Students can access different learning contents using mobile technology such as tablet computers.

In terms of the accumulated learning logs, techniques of learning analytics help understand the characteristics of learners' behaviors (Li et al., 2022). For example, we analyzed the log data of various learning tools and extracted learners' self-regulated learning strategies (Hsu et al., 2022). In addition, the insights from analyzing learning logs enable datadriven learning support, which facilitates learners' decision-making based on data (Ogata et al., 2018). However, extracting learning habits from log data and data-driven support for building learning habits have not yet attracted much attention.

When it comes to learning habits, it often suggests the learning patterns that learners tend to learn at a specific time. For example, learners may have a unique habit type of preferred time to work, such as morningness-eveningness preference (Sher et al., 2022). Individual differences in types of learning habits could have an impact on learners' academic achievements (Ricker et al., 2020; Sher et al., 2022). That is, building good learning habits can help learners achieve better academic performance.

In terms of the extraction of learning habits, extracting learning habits from daily learning logs makes it possible to evaluate learning habits not simply at a specific time but automatically trace the process of building learning habits. On the other hand, data-driven support for building learning habits can help learners to build learning habits based on evidence derived from learning logs. Therefore, this research proposes the approach of "Learning Habits Mining", which aims to extract the types and stages of learning habits from learners' daily learning logs and support learners to build learning habits with data-driven methods. Based on this, we answer the following three research questions.

- RQ1: What learning habits can be extracted from learners' daily logs?
- RQ2: How do learners build learning habits?
- RQ3: What interventions can support learners to build learning habits?

#### 2. Research Framework and Preliminary Findings

As introduced above, "Learning Habits Mining" is an approach to extract the types and stages of learning habits from learners' daily learning logs and support learners to build learning habits with data-driven methods. Figure 1 shows the three-step workflow of the approach. We have verified Step 1 and Step 2 with preliminary findings, which informed the implementation of Step 3.



Figure 1. Workflow of Learning Habits Mining.

#### 2.1 Step 1: Extracting types of learning habits from daily learning logs (RQ1)

In this step, we first define the types of learning habits by applying a sliding window to the time series data of learning logs (Figure 2). By the sizes of the sliding windows, we then extract different types of "short-term learning habits" and "long-term learning habits". "Short-term learning habits" indicate the active time slots when learners study within a day. "Long-term learning habits" indicate the patterns of whether learning is continued for a long time or crammed into the final stages.

For "short-term learning habits", we set the window size as one week and extracted types of learning habits, such as Morning on Weekdays, Evening on Weekdays, and Afternoon on Weekends, in terms of the weekly learning for regular math tests (Hsu, Otgonbaatar, et al., 2023). As for "long-term learning habits", we set the window size as three weeks and extracted types of learning habits, such as whether learners studied continuously or crammed to prepare for the math exam (Hsu, Horikoshi, et al., 2023).



Figure 2. Examples of Short-term and Long-term Types of Learning Habits.

#### 2.2 Step 2: Detecting stages of learning habits in the process of behavior change (RQ2)

In this step, we detect the stages of learning habits based on the Transtheoretical Model, which indicates people change their behaviors through five stages: Precontemplation (stage 1), Contemplation (stage 2), Preparation (stage 3), Action (stage 4), and Maintenance (stage 5). We refer to the Transtheoretical Model and operationalize the stages of learning habits as follows:

- Precontemplation (stage 1): Learners do not take action to build learning habits within the next four months.
- Contemplation (stage 2): Learners take action to build learning habits within the next four months.
- Preparation (stage 3): Learners take action to build learning habits within the next thirty days.
- Action (stage 4): Learners take action to build learning habits within the last four months.
- Maintenance (stage 5): Learners take action to build learning habits for more than four months.

Based on the definition, we proposed a data model to extract stages of learning habits. To show the application of the data model, we used learning logs of self-directed extensive reading. Figure 3 reveals an example of the stages a learner (e.g., 1295) underwent during the eleven-month learning activity.



Figure 3. Stages of Self-directed Extensive Reading Habits.

#### 2.3 Step 3: Recommending an intervention to build learning habits (RQ3)

In this step, we recommend feedback supporting learners to build learning habits based on the extracted types and stages of learning habits. Specifically, the pros and cons of different types of learning habits are evaluated and stage-based intervention is provided to support learners to achieve the maintenance stage, which means the formation of habits.

Regarding the pros and cons of different types of learning habits, we found learners who studied continuously had better performance than those who crammed to prepare for the math exam (Hsu, Horikoshi, et al., 2023). On the other hand, we can also provide a stagebased intervention for the learner shown in Figure 3, such as: "Let's find a good way to keep reading!" The objective is to prompt him or her transit from Action (stage 4) to Maintenance (stage 5).

#### 3. Contribution

We identify two contributions of this research from the preceding studies. First, this research reveals the learning habits of K12 learners and provides an approach to trace the process of

building learning habits automatically. Second, this research proposes interventions to the data-driven support for building learning habits.

Regarding the first contribution, past studies on learning habits are often in a self-paced learning context like a MOOC (Maslennikova et al., 2022; Ricker et al., 2020). There is a lack of understanding of tracing habits in daily learning at a school level from learning logs. In addition, the stages of learning habits can be easily assessed by questionnaires with speedy answers. However, the results tend to be arbitrary and subjectively dependent on individual assumptions (Maslennikova et al., 2022). Therefore, extracting the types and stages of learning habits is essential to understand students' learning behaviors at the K12 level.

As for the second contribution, past studies looked at stages of habits to figure out appropriate support for building different types of habits, such as physical activities (Jimmy & Martin, 2005) or diet (Clark et al., 2004). However, there is little focus on the support for building learning habits even though it can help learners achieve better academic performance. Therefore, data-driven support for building learning habits can help learners to make decisions on changing their behaviors based on evidence derived from the analysis of their learning logs.

#### Acknowledgements

This work is partially funded by JSPS KAKENHI 22H03902, NEDO JPNP18013, JPNP20006.

#### References

- Clark, M., Hampson, S., Avery, L., & Simpson, R. (2004). Effects of a brief tailored intervention on the process and predictors of lifestyle behaviour change in patients with type 2 diabetes. *Psychology, health & medicine, 9*(4), 440-449.
- Hsu, C.-Y., Horikoshi, I., Li, H., Majumdar, R., & Ogata, H. (November, 2022). Extracting Students' Self-Regulation Strategies in an Online Extensive Reading Environment using the Experience API (xAPI). In *Proceedings of 30th International Conference on Computers in Education* (pp. 326-331).
- Hsu, C. Y., Horikoshi, I., Li, H., Majumdar, R., & Ogata, H. (2023). Supporting "time awareness" in selfregulated learning: How do students allocate time during exam preparation?. *Smart Learning Environments, 10*(1), 21.
- Hsu, C.-Y., Otgonbaatar, M., Horikoshi, I., Majumdar, R., & Ogata, H. (December, 2023). Chronotypes of Learning Habits in Weekly Math Learning of Junior High School. In *Proceedings of 31st International Conference on Computers in Education*.
- Jimmy, G., & Martin, B. W. (2005). Implementation and effectiveness of a primary care based physical activity counselling scheme. *Patient education and counseling*, *56*(3), 323-331.
- Li, J., Li, H., Majumdar, R., Yang, Y., & Ogata, H. (2022, March). Self-directed extensive reading supported with GOAL system: mining sequential patterns of learning behavior and predicting academic performance. In *LAK22: 12th International Learning Analytics and Knowledge Conference* (pp. 472-477).
- Maslennikova, A., Rotelli, D., & Monreale, A. (2022, July). Visual Analytics for Session-based Time-Windows Identification in Virtual Learning Environments. In 2022 26th International Conference Information Visualisation (IV) (pp. 251-258). IEEE.
- Ogata, H., Majumdar, R., Akcapinar, G., Hasnine, M. N., & Flanagan, B. (2018). Beyond learning analytics: Framework for technology-enhanced evidence-based education and learning. In *26th International Conference on Computers in Education Workshop Proceedings* (pp. 493-496). Asia-Pacific Society for Computers in Education (APSCE).
- Sher, V., Hatala, M., & Gašević, D. (2022). When Do Learners Study? An Analysis of the Time-of-Day and Weekday-Weekend Usage Patterns of Learning Management Systems from Mobile and Computers in Blended Learning. *Journal of Learning Analytics*, *9*(2), 1-23.
- The Government of Japan. (2021). Japan's giga school program equips students for digital society [Online] Available: https://www.japantimes.co.jp/2021/03/22/special-supplements/japans-giga-school-program-equips-students-digital-society/ (March 24, 2022).
- Ricker, G., Koziarski, M., & Walters, A. (2020). Student Clickstream Data: Does Time of Day Matter?. *Journal of Online Learning Research*, 6(2), 155-170.

## Applying Macro- and Micro-scripts to Facilitate Undergraduates' Interdisciplinary Competence

#### Guo SU<sup>a</sup>

<sup>a</sup>National Institute of Education, Nanyang Technological University, Singapore \*nie22.sg4560@e.ntu.edu.sg

**Abstract:** This study aims to develop principles for designing CSCL scripts that can guide interdisciplinary collaboration and knowledge-creation practices to facilitate undergraduates' interdisciplinary competence. Design-based research (DBR) will be conducted in an undergraduate course to investigate if the macro- and micro-scripts have the potential to improve students' performance, and why and how these designs make a difference.

Keywords: Interdisciplinary collaboration, CSCL scripts, interdisciplinary competence

#### 1. Introduction

#### 1.1 Interdisciplinary Competence in Interdisciplinary Collaboration

Fostering collaborative learning is a major educational goal related to the development of lifelong learners in the twenty-first century (Lee et al., 2014). Interdisciplinarity is a particular form of collaboration (Claus, & Wiese, 2019). In interdisciplinary collaboration, group members who come from various backgrounds are frequently formed around a certain subject or question. Interdisciplinary collaboration in our context is expected to elicit collaborative knowledge-creating learning where individuals with expertise in several fields are expected to create new knowledge to solve complex real-world problems (Paavola & Hakkarainen, 2005; Vartiainen et al., 2022). However, there is still a lack of research investigating how interdisciplinary collaboration and knowledge creation could be promoted and practically acted in higher education (Vartiainen et al., 2022). Edmondson and Harvey (2018) have argued that interdisciplinary collaboration has become a key success factor for today's organizations. Researchers and practitioners design meaningful interdisciplinary collaboration to investigate interactive patterns as well as students' experiences, promote students' acquiring new knowledge and higher-order thinking skills, etc. For example, Cowden and Santiago (2016) sought to promote students' critical thinking skills via problembased learning in an interdisciplinary collaboration context. Tan and Vicente (2019) also proposed interdisciplinary collaboration combining Computer Science and Product Design in an undergraduate program to support students' development of 21st-century skills such as problem-solving, creativity, collaboration, etc.

To successfully engage in interdisciplinary collaboration - more precisely, to be able to understand and act in interdisciplinary learning or work situations - students need adequate interdisciplinary competence which is defined as "the capacity to integrate knowledge and modes of thinking in two or more disciplines or established areas of expertise to produce a cognitive advancement - such as explaining a phenomenon, solving a problem, or creating a product - in ways that would have been impossible or unlikely through single disciplinary means" (Boix Mansilla et al., 2000, p. 17)." Interdisciplinary competence includes taking a critical stand on disciplinary limitations, solving complex problems across disciplines, communicating across disciplines, handling interdisciplinary collaboration and teamwork, as well as using integrative potentials to create innovations (Brassler, & Dettmers, 2017). Though interdisciplinary competence highly relates to each of the twenty-first-century skills (Brassler, & Dettmers, 2017), there seems to be a lack of design principles and

pedagogical support to facilitate students' interdisciplinary competence in complex problemsolving when students engage in interdisciplinary collaboration (Vartiainen et al., 2022).

The empirical research on interdisciplinary learning design in higher education in the exploratory stages (Brassler & Dettmers, 2017). Many researchers have called for pedagogies in which students can co-create new knowledge and develop more generic key competencies, such as interdisciplinary competence, for the future (Binkley et al., 2010). Further investigation on effective interdisciplinary teaching and learning design and implementation is needed. Hence, this study aims to develop design principles for effective interdisciplinary collaborative learning and knowledge creation to better support students' development of interdisciplinary competence.

#### 1.2 Scripting Interdisciplinary Collaboration

Research shows that learners often have difficulties spontaneously engaging in beneficial collaborative learning activities (Vogel et al., 2017). Also, learners typically do not engage in these "high-level" collaboration processes without guidance (Fischer et al., 2013). To overcome these problems, scaffolding for collaboration can be provided through CSCL scripts. According to the granularity of the prescribed actions, there are two different scripts named macro- and micro-scripts. Macro-scripts are pedagogical models, i.e., they model a sequence of tasks to be performed by groups, whereas micro-scripts are interaction models that are embedded in the environment and that students are expected to use and progressively internalize (Dillenbourg & Hong, 2008). According to the way learners collaborate with each other, two scripts are commonly used: epistemic scripts specify how learners work on a given task, while social scripts structure how learners interact with each other (Mahardale & Lee, 2013).

The goal of scripting interdisciplinary collaboration can be twofold: first, increasing interaction through knowledge interdependence; second, engaging in interdisciplinary problem-solving (Vogel et al., 2017), which shares the same purpose as interdisciplinary competence development. Jigsaw script, a pedagogical model, will be used as the macroscript in this study. According to studies, jigsaw can create a learning environment that is crucial for the development of problem-solving skills (Sevim, 2015) which is part of interdisciplinary competence and is critical to engaging in interdisciplinary collaboration. In the jigsaw script, each group member is given a certain subtopic. Then, the "expert group" is made up of group members from several groups who were given the same subtopic. Finally, the experts go back to their home groups to teach other group members the specialized subtopics and solve problems together. Under the macro-script, micro-scripts are applied to structure the interaction process that students are expected to follow (Villasclaras-Fernández et al., 2009). An epistemic script and a social script will be designed as micro-scripts in our context. To support the "expert group" co-creating knowledge and understanding the subtopics that they focus on, problem-oriented prompts will be used as epistemic scripts. In addition, when learners share their knowledge and ideas in their home groups, sentence openers for responding to each other will be provided as the social scripts that aim to facilitate interactions within groups.

#### 1.3 Significance of the Study

There seems to be a lack of cases that identify effective pedagogical practices for supporting students in developing competencies needed to address complex challenges across disciplinary boundaries. Therefore, the theoretical contribution is that this research has the potential to fill in the gap by developing pedagogical design principles for effective interdisciplinary collaborative learning activities in higher education settings. Furthermore, effective interdisciplinary collaboration results in more creative solutions by bringing new and diverse views to the problem-solving process, which is the practical contribution of this study.

The guiding research question is how to script effective interdisciplinary collaboration and knowledge creation to improve undergraduates' interdisciplinary competence. The specific sub-research questions are: (1) To what extent do the macro- and micro-scripts enhance students' interdisciplinary competence? (2) How to further improve the script design to better support teaching and learning in interdisciplinary higher education?

#### 2. Methodology

#### 2.1 CSCL Context and Participants

Second-year undergraduate students who come from different backgrounds will be randomly grouped together to form interdisciplinary teams of 5 or 6. In the course, students will systematically analyze major current sustainability challenges from various perspectives. Several different topics will be involved in one semester. During the tutorials, first, each group member chooses a role for themselves, such as an economist. Then, each "expert" goes to his or her own expert group to read materials, discuss the relevant topics and record the results on the online platform with the support of epistemic scripts (about 25 minutes). Finally, each expert goes back to his or her original group to discuss these topics with the social scripts and writes a report on the platform about the final decision of the whole group (about 25 minutes).

#### 2.2 Research Design

Design-based research (DBR), which is a useful method for developing new theories and practices (Barab, 2014), has the potential to be conducted in this context. Dedicated to dealing with problems in the real-world learning context via close collaboration between researchers and practitioners, factors of DBR (design, problem, theory, and the naturalistic context) interact with each other, and each iteration in DBR offers a progressive refinement (Collins, 1992). DBR provides evidence of not only what functions, but of why and how functions. There will be about three cycles of DBR throughout the study. The overall procedure of the research design is as follows (see Figure 1). Our goal is to develop design principles that can guide interdisciplinary collaboration and knowledge-creation practices to facilitate undergraduates' interdisciplinary competence. First, with the guidance of theoretical ideas and principles of scripts from the literature, practitioners and researchers will co-design and implement macro- and micro-scripts embedded in tutorial activities for interdisciplinary collaboration and knowledge co-creation. Then, through analyzing data about students' interdisciplinary competence and perceptions of practices by the lecturer and students, we further improve our pedagogical practices which will be implemented in the next cycle. Finally, implications for refining the design principles will come from all the cycles.

To answer the first research question, content analysis will be conducted to evaluate the quality of groups' written artefacts in terms of interdisciplinary competence. Additionally, a self-reported survey on interdisciplinary competence will be analysed. For the second research question, discussion dialogue, behavioural data recorded by videos and screen recordings will be analyzed to investigate how students use the scripts and identify what scripts can be improved. Also, theme-based coding will be implemented to analyse the facilitator's and students' perceptions of scripts which are gathered through interviews.



Implications for refining the design principles

Figure 1. The Overall Procedure of Design-based Research (Lakkala et al., 2015).

#### Acknowledgements

I would like to appreciate all the support given by my supervisor, co-supervisor, and collaborator.

#### References

- Barab, S. (2014). Design-based research: A methodological toolkit for engineering change. *The Cambridge handbook of the learning sciences*, *2*, 151-170.
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., & Rumble, M. (2010). Defining 21st century skills. In E. Care, P. Griffin, & M. Wilson (Eds.), *Assessment and teaching of 21st century skills* (Issue January, pp. 1–65). Springer.
- Boix Mansilla, V., Miller, W. C., & Gardner, H. (2000). On disciplinary lenses and interdisciplinary work. *Interdisciplinary curriculum: Challenges to implementation*, 17-38.
- Brassler, M., & Dettmers, J. (2017). How to enhance interdisciplinary competence—interdisciplinary problem-based learning versus interdisciplinary project-based learning. *Interdisciplinary Journal of Problem-Based Learning*, *11*(2).
- Claus, A. M., & Wiese, B. S. (2019). Development and test of a model of interdisciplinary competencies. *European Journal of Work and Organizational Psychology*, 28(2), 191-205.
- Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15–22). New York: Springer-Verlag.
- Cowden, C. D., & Santiago, M. F. (2016). Interdisciplinary explorations: promoting critical thinking via problem-based learning in an advanced biochemistry class. *Journal of Chemical Education*, 93(3), 464-469.
- Dillenbourg, P., & Hong, F. (2008). The mechanics of CSCL macro scripts. *International Journal of Computer-Supported Collaborative Learning*, *3*, 5-23.
- Edmondson, A. C., & Harvey, J. F. (2018). Cross-boundary teaming for innovation: Integrating research on teams and knowledge in organizations. *Human Resource Management Review*, *28*(4), 347-360.
- Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational psychologist*, *48*(1), 56-66.
- Lakkala, M., Toom, A., Ilomäki, L., & Muukkonen, H. (2015). Re-designing university courses to support collaborative knowledge creation practices. *Australasian Journal of Educational Technology*, *31*(5).
- Lee, K., Tsai, P. S., Chai, C. S., & Koh, J. H. L. (2014). Students' perceptions of self-directed learning and collaborative learning with and without technology. *Journal of Computer Assisted Learning*, 30(5), 425-437.
- Mahardale, J. W., & Lee, C. B. (2013). Understanding how social and epistemic scripts perpetuate intersubjectivity through patterns of interactions. *Interactive Learning Environments*, *21*(1), 68-88.
- Paavola, S., & Hakkarainen, K. (2005). The knowledge creation metaphor—An emergent epistemological approach to learning. *Science & Education*, *14*, 535–555.
- Sevim, O. (2015). Influence of the Subject Jigsaw Technique on Elementary School Seventh Grade Students' Academic Achievement and On Their Problem Solving Skills. *Education & Science/Egitim ve Bilim*, 40(177).
- Tan, T. A. G., & Vicente, A. J. (2019). An innovative experiential and collaborative learning approach to an undergraduate marketing management course: A case of the Philippines. *The International Journal of Management Education*, *17*(3), 100309.
- Vartiainen, H., Vuojärvi, H., Saramäki, K., Eriksson, M., Ratinen, I., Torssonen, P., ... & Pöllänen, S. (2022). Cross-boundary collaboration and knowledge creation in an online higher education course. *British Journal of Educational Technology*, 53(5), 1304-1320.
- Villasclaras-Fernández, E. D., Isotani, S., Hayashi, Y., & Mizoguchi, R. (2009, June). Looking Into Collaborative Learning: Design from Macro-and Micro-Script Perspectives. In *AIED* (pp. 231-238).
- Vogel, F., Wecker, C., Kollar, I., & Fischer, F. (2017). Socio-cognitive scaffolding with computersupported collaboration scripts: A meta-analysis. *Educational Psychology Review*, 29(3), 477-511.

## APPLYING MACRO- AND MICRO-SCRIPTS TO FACILITATE UNDERGRADUATES' INTERDISCIPLINARY COMPETENCE

by

Guo Su nie22.sg4560@e.ntu.edu.sg

Supervisor: Dr. Wenli Chen wenli.chen@nie.edu.sg

Co-supervisor: Dr. Gaoxia Zhu gaoxia.zhu@nie.edu.sg

Submitted to C2: ICCE Sub-Conference on Computer-supported Collaborative Learning (CSCL) and Learning Sciences

## Exploring Business Students' STEM Literacy for Work Preparedness

De-Qi ZHANG<sup>a\*</sup> & Morris Siu-Yung JONG<sup>a</sup>

<sup>a</sup>Department of Curriculum and Instruction, The Chinese University of Hong Kong, Hong Kong, China

\*deqizhang@link.cuhk.edu.hk

**Abstract:** This paper presents a proposed doctoral study exploring STEM (science, technology, engineering, and mathematics) literacy for business students' work preparedness. The overall research consists of two studies. Study 1 will develop a theoretical model relating business students' STEM literacy for work preparedness by adopting a grounded theory approach. Based on the model, a measurement scale, viz. Business Students' STEM Literacy Scale (BSSLS), will be developed. Study 1 will involve fresh business graduate workers and business students with work experience in STEM-related industries. Study 2, which aims to recruit around 500 participants in mainland China and Hong Kong, will validate the BSSLS and investigate the current situation of final-year business students' STEM literacy. The proposed research will contribute to generating new insights into researching STEM literacy in higher education and incorporating this literacy as a curricular component in business education.

**Keywords:** STEM literacy, business students, work preparedness

#### 1. Research Motivation

Over the past two decades, there has been a rapid evolution in STEM (science, technology, engineering, and mathematics) products, leading to an unprecedented expansion in related industries (Podobnik et al., 2023). As a result of this growth, business workers have become increasingly involved in the business processes of STEM industries (Hallioui et al., 2022). Business work tasks related to STEM industries are interdisciplinary business processes rather than relying on individual disciplines (Fedyk et al., 2017; Ransbotham et al., 2017; Sollosy & McInerney, 2022). Therefore, business students should be equipped with STEM literacy for their work preparedness, which is an interdisciplinary ability to solve STEM-related challenges that cannot be solved by a single discipline (Mohr-Schroeder et al., 2020). However, the current problem is that business students are ill-prepared for working with STEM (Sollosy & McInerney, 2022), which is not yet adequately addressed by higher education. This may cause a gap between the required skills expected by employers and the actual skills possessed by business graduates.

STEM education has been widely regarded as a globally important initiative across primary to higher education (Huang et al., 2022; Lau et al., 2022), but limited research has been conducted on STEM literacy in business education (Bunch, 2020; Sollosy & McInerney, 2022). It is unknown what STEM literacy business students should possess to adapt to the constantly updating business workplace, and whether they currently possess the corresponding literacy. It might be because education researchers overlooked the significance of business employers' perspective and business studies. While some universities have attempted to cultivate business students' STEM literacy (e.g., CUHK, 2023), the effectiveness of the existing curriculum remains to be determined because the theoretical standard is unknown.

This research aims to bridge the gap in researching business students' STEM literacy for work preparedness. The anticipated outcomes seek to enhance the overall understanding and implementation of STEM education in business schools.

#### 2. Research Questions

This research aims to explore business students' STEM literacy for work preparedness and identify the gap between the STEM literacy that business students need to possess and that they currently possess. The objectives of this research consist of four folds: (a) develop a theoretical model of constructs that relate business students' STEM literacy for work preparedness; (b) generate the Business Students' STEM Literacy Scale (BSSLS), a measurement scale based on the theoretical model; (c) use the BSSLS to collect data from a large group of final-year business students' STEM literacy level. To achieve the objectives, the research will address the following research questions:

- What is the theory that explains business students' STEM literacy for work preparedness?
- What is the scale's reliability and validity, as analyzed through exploratory and confirmatory factor analysis?
- Does the scale's factor structure confirm the theoretical model?
- What is the current STEM literacy state of final-year business students?

#### 3. Potential Contributions

This research aims to develop a theoretical model of STEM literacy for work preparedness in business education. While frequent rhetoric promotes STEM literacy is essential for this generation (Kayan-Fadlelmula et al., 2022), there exists a research gap in business education regarding STEM literacy for work-preparedness. Current studies on STEM literacy models or constitutions are mainly conducted in K-12 contexts (e.g., National Research Council, 2014; Huang & Jong, 2020; Jackson et al., 2021), STEM teacher education (e.g., Chai et al., 2020; Jong et al., 2021), and general education (e.g., Zollman, 2012; Mohr-Schroeder et al., 2015; Tang & Williams, 2019). However, limited studies have delved into STEM literacy in business education.

The theoretical model can provide a foundation for developing a measurement scale to examine the STEM literacy that final-year business students currently possess. By precisely identifying the areas of STEM literacy in which final-year business students are lacking, the research can provide insights to optimize curricula design for business students' work preparedness. Higher-quality business curricula can equip business students with enhanced workforce competitiveness, thereby further supporting their future professional development.

Furthermore, the research outputs can also serve as a reference for future studies focusing on STEM literacy in higher education. That is, researchers and educators of other majors can draw inspiration from this research to investigate corresponding STEM literacy models for work preparedness for different professions.

#### 4. Methodology

This research consists of two studies. Study 1, via a grounded theory approach, will develop a theoretical model which delineates the necessitated business students' STEM literacy for work preparedness. Based on this model, the initial items of the BSSLS will be generated. The results of Study 1 will provide a comprehensive understanding of business students' STEM literacy for work preparedness, which has not been depicted in previous research. Study 2 will validate the BSSLS and investigate the current situation of final-year business students' STEM literacy. The gap between current business students' STEM literacy and what they need to possess will be identified. The findings of Study 2 can offer insights for curricula design to improve business students' work preparedness, enhance their workforce competitiveness and support their future professional development.

#### 4.1 Study 1: Business Students' STEM Literacy for Work Preparedness

#### 4.1.1 Sampling and Participants Recruitment

Study 1 will follow the tenets of grounded theory and take a theoretical sampling approach (Strauss & Corbin, 1990). The sampling process intends to be purposive, iterative, and adaptive until saturation. Theoretical saturation appears when no new properties of categories have been discovered, leading to the emergence of theory (Aldiabat & Le Navenec, 2018). In-between data analysis will be conducted to refine the next round of data collection.

Study 1 plans to recruit fresh business graduate workers and business students with work experience in STEM-related industries. Fresh business graduate workers refer to those who have just entered the workforce and have not yet been promoted. To collect rich data, this study will recruit participants by balancing three considerations: working departments, business backgrounds, and gender.

#### 4.1.2 Data Collection and Analysis

Study 1 will employ semi-structured interviews. Participants will be invited to share their views based on their current experiences. The interview questions for grounded theory should strike a balance between being broad enough to encompass various experiences and specific enough to delve into the participants' individual experiences (Charmaz & Belgrave, 2012). Therefore, participants will be first invited to share their stories. Then, the researcher will open the interviews accordingly and ask follow-up questions (ibid.). Besides, because this study plans to conduct several rounds of interviews until saturation, interview questions might be adapted slightly in each successive round.

A grounded theory approach will be employed to analyze the interview data (Glaser et al., 1968) through Nvivo10. The analytical process will involve several rounds of data collection, analysis, and visualization. Open, axial, and selective coding and modeling will be applied (ibid.). Based on qualitative data analysis, this step will generate the initial items for BSSLS.

Study 1 intends to employ two validating strategies to ensure the trustworthiness of the interpretations (Creswell & Báez, 2020). Firstly, STEM literacy and business education experts will be invited to review the research and findings. Additionally, the researcher plans to invite some participants to review the final report and provide feedback on the findings.

#### 4.2 Study 2: BSSLS and Final-year Business Students' STEM Literacy

#### 4.2.1 Sampling and Participants Recruitment

Study 2 will adopt stratified sampling. Sampling should meet the following conditions for typicality and representation: (a) the participants should be final-year business students; (b) balancing the proportion of male and female participants; and (c) conducting the sampling in business schools of different levels. Thus, 200 final-year business students from mainland China and Hong Kong will be recruited for the pilot test. In the following validation test, Study 2 plans to employ another 300 final-year business students.

#### 4.2.2 Data Collection and Analysis

The pilot test aims to refine the initial items proposed for the BSSLS. Approximately 200 and 300 business students (from mainland China and Hong Kong) will be recruited to participate in the pilot test and validation test respectively. At the validation stage, participants will be asked to complete the revised BSSLS. The data collected in the pilot test will be analyzed with descriptive statistical analysis, exploratory factor analysis using principal components analysis with varimax rotation, and principal axis factor analysis. The data collected in the validation test will be analyzed with descriptive statistical analysis.

#### References

- Aldiabat, K. M., & Le Navenec, C. L. (2018). Data saturation: The mysterious step in grounded theory methodology. *The Qualitative Report*, 23(1), 245–261.
- Bunch, K. J. (2020). State of undergraduate business education: A perfect storm or climate change?. *Academy of Management Learning & Education, 19*(1), 81–98.
- Chai, C. S., Jong, M. S. Y., & Yan, Z. M. (2020). Surveying Chinese teachers' technological pedagogical STEM knowledge: A pilot validation of STEM-TPACK survey. *International Journal of Mobile Learning & Organisation, 11*(2), 203–214.

Charmaz, K., & Belgrave, L. (2012). Qualitative interviewing and grounded theory analysis. In J. F. Gubrium, J.s A. Holstein, A. B. Marvasti, & K. D. McKinney (Eds.), *The SAGE handbook of interview research: The complexity of* the craft (2nd Ed.) (pp. 347–365). Sage.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3–21.

Creswell, J. W., & Báez, J. C. (2020). 30 essential skills for the qualitative researcher. Sage.

- CUHK. (2023). CUHK x CUHK(Shenzhen) Interdisciplinary Data Analytics + X Double Major *Programme*. Retrieved June 6, 2023, from: https://admission.cuhk.edu.hk/dmp/tc/
- Fedyk, T., Singer, Z., & Soliman, M. (2017). The sharpest tool in the shed: IPO financial statement management of STEM vs. non-STEM firms. *Review of Accounting Studies*, 22, 1541–1581.
- Glaser, B. G., Strauss, A. L., & Strutzel, E. (1968). The discovery of grounded theory; strategies for qualitative research. *Nursing research*, *17*(4), 364.
- Hallioui, A., Herrou, B., Santos, R. S., Katina, P. F., & Egbue, O. (2022). Systems-based approach to contemporary business management: An enabler of business sustainability in a context of industry 4.0, circular economy, competitiveness and diverse stakeholders. *Journal of Cleaner Production*, 133819.
- Huang, B., & Jong, M. S. Y. (2020). Developing a generic rubric for evaluating students' works in STEM Education. *Proceedings of the 6th International Symposium on Educational Technology (ISET 2020)* (pp. 210–213). IEEE.
- Huang, B., Jong, M. S. Y., King, R. B., Chai, C. S., & Jiang, M. Y. C. (2022). Promoting secondary students' twenty-first century skills and STEM career interests through a crossover program of STEM and community service education. *Frontiers in Psychology*, 13, 903252.
- Jackson, C. D., & Mohr-Schroeder, M. J. (2018). Increasing STEM literacy via an informal learning environment. *Journal of STEM Teacher Education*, *53*(1), 43–52.
- Jong, M. S. Y., Song, Y., Soloway, E., & Norris, C. (2021). Teacher professional development in STEM education. *Educational Technology & Society, 24*(4), 81–85.
- Kayan-Fadlelmula, F., Sellami, A., Abdelkader, N., & Umer, S. (2022). A systematic review of STEM education research in the GCC countries: Trends, gaps and barriers. *International Journal of STEM Education*, 9(1), 1–24.
- Lau, W. W., & Jong, M. S. (2022). Typology of teachers' stages of concern for STEM education. *Research in Science & Technological Education*, 1-19.
- Mohr-Schroeder, M., Bush, S. B., Maiorca, C., & Nickels, M. (2020). Moving toward an equity-based approach for STEM literacy. In C. Johnson, M. J. Mohr-Schroeder, T. Moore, & L. English (Eds.), *Handbook of research on STEM education,* (pp. 29–38). Routledge.
- Mohr-Schroeder, M., Cavalcanti, M., & Blyman, K. (2015). STEM education: understanding the changing landscape. In A. Sahin (Ed.), *A practice-based model of STEM teaching* (pp. 3–14). Brill.
- National Research Council. (2014). STEM integration in K-12 education: Status, prospects, and an agenda for research. National Academies Press.
- Podobnik, B., Dabić, M., Wild, D., & Di Matteo, T. (2023). The impact of STEM on the growth of wealth at varying scales, ranging from individuals to firms and countries: The performance of STEM firms during the pandemic across different markets. *Technology in Society*, *72*, 102148.
- Ransbotham, S., Kiron, D., Gerbert, P., & Reeves, M. (2017). Reshaping business with artificial intelligence: Closing the gap between ambition and action. *MIT Sloan Management Review*. https://sloanreview.mit.edu/projects/reshaping-business-with-artificial-intelligence/
- Sollosy, M., & McInerney, M. (2022). Artificial intelligence and business education: What should be taught. *The International Journal of Management Education, 20*(3), 100720.
- Tang, K. S., & Williams, P. J. (2019). STEM literacy or literacies? Examining the empirical basis of these constructs. *Review of Education*, 7(3), 675–697.
- Zollman, A. (2012). Learning for STEM literacy: STEM literacy for learning. *School Science and Mathematics*, *112*(1), 12–19.

## Using meaningful gamification to redesign simulation-labs for engaging learners in science inquiry practices

#### Archana RANE<sup>a\*</sup>

Doctoral Research Scholar, Indian Institute of Technology, Bombay, India <u>\*archana.rane@iitb.ac.in</u>

Abstract: Practical labs in science are expected to foster disciplinary practices of "doing science" in learners. Simulation-based labs which provide the convenience of conducting practical lab experiments online offer several technological affordances, however, come with their own set of challenges such as steep learning curves, working in isolation, lack of personalization which may adversely affect learner engagement and motivation. OLabs is a set of simulation-based labs for schools widely used in India. The main goal of this work is to re-design simulation-based labs like OLabs to engage learners in science practices and to motivate learners to engage in science practices using gamification. Gamification is known to have a positive impact on learner engagement and motivation. Meaningful gamification, a type of gamification approach focuses on intrinsically motivating learners to find meaning in the given learning context. We have designed a framework GaMINLab, which is the proposed redesign of OLabs and the preliminary design is informed by exploratory study findings and literature recommendations of disciplinary practices and simulation-based labs. Currently GaMINLab framework is built around 2 labs and study is planned to validate the design of GaMINLab.

**Keywords:** Simulation-based labs, meaningful gamification, science disciplinary practices, inquiry labs, scientific inquiry

#### 1. Introduction

Practical labs, are expected to help learners learn science, learn about science, and learn to do science. In recent years, the extensive emergence of simulation-based labs has facilitated learners to do virtual experiments anytime, anywhere, and any number of times. These labs also offer possibilities to illustrate concepts beyond the boundaries of traditional labs, for example magnetic field intensity, lifecycle of mosquito, etc. Despite world-wide proliferation and advantages, there are reported challenges such as learner isolation, steep learning curves, lack of suitable guidance, lack of personalization, cookbook exercises, etc (Moore et al., 2013). The challenges may lead to frustration among learners, resulting in early dropout or loss of motivation and engagement. Thus, it is critical to give adequate attention to engage and motivate learners to do the underlying lab related activities.

The context of this work is OLabs, set of standard simulation-based labs for school level experiments in India. OLabs is a significant initiative by the government of India to address concerns such as lack of basic lab infrastructure, limited practice opportunities, and to supplement physical labs (M Sasikumar, 2016). OLabs has about 310 labs in classes 6-12 and is used by lakhs of students/teachers on a daily basis. In its current form, there are no tasks/problems built-in to engage learners in science disciplinary practices. The goal of this work is to re-design OLabs to a) provide opportunities for engaging in science disciplinary practices b) engage and motivate learners as they engage in these practices. This proposed redesign, GaMINLab, is informed by the literature recommendations for disciplinary practices, addressing challenges in simulation-based labs and meaningful gamification. The design features of GaMINLab are discussed and plan for validation of this design is proposed.

#### 2. Literature and Theory

#### 2.1 Simulation-based labs

Simulation-based labs facilitate conducting experiments online, overcoming the geographical and time limitations of traditional labs. To tackle challenges like steep learning curves, isolation, lack of personalization, design guidelines are recommended in literature a) providing investigation opportunities beyond the classroom, b) posing driving questions to focus learners' exploration, c) implicit scaffolding, d) encouraging reflection on findings, and e) promoting peer interaction and engaging in collaborative work (Moore et al., 2013).

#### 2.2 Disciplinary Practices

Science education is increasingly focusing on involving learners in disciplinary practices, as emphasized in global curriculum recommendations. These practices include asking questions, developing and using models, planning and conducting investigations, analyzing and interpreting data, engaging in evidence-based arguments, and obtaining, evaluating, and communicating information, etc (Jaber et al., 2018). To foster disciplinary practices in a learning environment, literature suggests guidelines such as: a) providing direct experience with phenomena, b) making learner thinking visible to grapple with ideas and c) provide essential procedures/relevant resources for completing investigations. Inquiry learning is recommended for effective science learning, where learners engage in sense-making, discussions, evidence-based explanations, etc (Jaber et al., 2018) (Moore et al., 2013).

#### 2.3 Meaningful Gamification

One of the goals for redesigning OLabs is to motivate learners to engage in lab related activities. Gamification, an established approach to fostering student engagement, uses game attributes to encourage game-like behavior in non-game contexts. Game-based mechanics can effectively engage learners, motivate their actions, and aid problem-solving. Reward-based gamification, a popular type of gamification, relies on extrinsic motivation and may not yield lasting changes, and can be unsatisfactory for some learners. Meaningful gamification, another type, employs game design elements to cultivate intrinsic motivation in non-game settings by creating an enjoyable learning environment where participants can explore and find meaning(Nicholson, 2015). This approach aims for sustained engagement compared to the short-term, reward-driven nature of extrinsic motivation. Meaningful gamification is based on self-determination theory (Ryan & Deci, 2000), which mentions competence, autonomy, and relatedness to form intrinsic motivation. The theory highlights that when these three psychological needs are fulfilled, people find tasks meaningful and continue participating.

#### 3. Goal of Research

The broad goal of this work is to re-design simulation-based labs like OLabs to a) provide opportunities for engaging in science disciplinary practices b) motivate learners to engage in these practices. This re-design of simulation-based labs is currently named GaMINLab (Gamification (Meaningful) in Inquiry Labs. As a proof of concept, GaMINLab will be configured for select set of existing OLabs and guidelines for bringing in similar class of science OLabs under GaMINLab will be published. Design guidelines for meaningfully gamifying simulation-labs using this framework. GaMINLab framework can potentially be used by teachers to augment simulation-based labs like OLabs to engage learners in science practices in fun way.

#### 4. Methodology

#### 4.1 Design of the GaMINLab Framework

The current design of GaMINLab is informed by a) literature recommendatins for fostering disciplinary practices and addressing challenges in simulation-based labs, and b) the theory of meaningful gamification. For a specific lab, learners will get categorized problem-solving challenges in scenarios like park, beach, etc and learners are prompted to attempt problem-solving activities associated with chosen problems. The problem-solving activities are aligned with the problem-solving inquiry framework (Kim & Hannafin, 2011), with tasks such as *Investigate using the Lab, Propose Answers and Predictions, and Share and Discuss your Lab Investigation Report*. The related OLabs simulation is embedded in specific tasks and can be manipulated for investigations. Learners are asked to create artifacts like investigation plans, lab investigation reports, for evaluation by self, peers, or teachers using rubrics.

Instead of predefined sequence for the problem-solving tasks, learners are prompted with options choose their strategy, encouraging flexible problem-solving approaches. System awards corresponding badges for each disciplinary task completion, acknowledging progress and motivating further engagement. Lab onboarding familiarizes learners with UI objects, lab variants, observation by including interactive content like hotspots, quizzes, chat/forums facilitate discussion, and report sharing among peers. Scaffolds assist learners in preparing investigation plans, lab reports and provide guidance for alternate paths, aiding progression or retracking actions. Currently GaMINLab has two class IX labs with one problem in each scenario. Corresponding OLabs simulation is embedded with the system. This allows learners to conduct investigations by manipulating variants and using other available controls. We plan to conduct series of studies to validate and refine the framework.

#### 5. Study Design

A study for class IX students is being planned is to validate the design of GaMINLab. We plan to analyze the pattern of engagement of the learners, motivation, evidence of disciplinary practices, usability of the system, user perception, etc. The physical study will be done in two parts, a pilot (10 students) and main study (60 students). Separate schools will be selected for the pilot and the main study to avoid bias. The research questions for are - RQ-1 "What is the pattern of engagement of learners while using GaMINLab?;RQ-2 "How does GaMINLab contribute to motivation of learners; RQ 4.1 How usable is the system? RQ 4.2 What is the learner perception of system? In RQ-1 we are looking at the behavioral engagement of the learner in GaMINLab. Behavioral engagement (Fredricks, 2004) concerns learner participation and learner interest in academic tasks, in our context it is engaging in disciplinary practices. The study session will last approximately 2-2.5 hours. Participants will take a pre-test on disciplinary practices, followed by lab activities in the GaMINLab. Each student is expected to solve one problem in one lab during the allotted time. After interacting with the intervention, participants will complete a post-test on disciplinary practices, the SUS questionnaire, and a learner perception survey, followed by semi-structured interviews.

#### 5.1 Data Sources and Instruments

In identifying patterns of engagement, we will find attempts/reattempts, badges earned, artefacts submitted, etc. We will be measuring intrinsic motivation using the Intrinsic Motivation Inventory a self-reported questionnaire on interest/usefulness, perceived choice, perceived competence, pressure and tension, relatedness, etc. Interest/usefulness is a direct indicator of motivation so it an appropriate metric to get insights of their motivation as they interact with GaMINLab. We are also interested to measure self-reported measures of perceived choice (autonomy), perceived competence (mastery), relatedness. These are also considered as positive predictors while pressure/tension is considered as negative predictor of intrinsic motivation. Table1 illustrates theoretical constructs, operational constructs and data sources.

We measure the disciplinary practices demonstrated in artifacts i.e. investigation plan, lab reports etc. using a 4-point rubric scale for each of the 4 disciplinary practices. Disciplinary practices manifested in interaction with the system will be identified. by thematic analysis of qualitative and log data for specific disciplinary practice. We will measure the learner perception of usability of the system, challenges faced by them and their experience with the system using SUS and learner perception survey. Thematic analysis of perception survey responses will be done and responses will be triangulated with interview responses to a) evaluate usability b) identify challenges c) confirm design considerations are being working as intended.

	Theoretical	Operational	Data
	Construct	Constructs	Sources/Instru
			ments
RQ-1	Pattern of	Paths taken, task	Screen recording,
	Engagement	attempts/reattempts,	logs, Attempts
		variants used,	report, interview
		badges earned	transcripts, notes
RQ-2	Intrinsic	Interest/Usefulness, perceived	IMI survey
	Motivation	choice, perceived competence	responses, interview
		perceived relatedness	transcripts
RQ-3	Disciplinary	Degree of Disciplinary	Pre-test, Post-test
	Practices	practices: demonstrated in	responses, artifacts
		artifacts, interaction in system	
RQ-4	Learner	Usability, Challenges SUS guestionna	
	Perception		Perception Survey

Table 1.	RQs,	Theoretical	Construct,	Operational	Construct	and Data	a Sources
----------	------	-------------	------------	-------------	-----------	----------	-----------

#### 6. Conclusion

We believe that a framework like GaMINLab built around simulation-based labs would intrinsically motivate learners in "doing" science and cognitively engage them in the related activities. The planned pilot study will give us valuable insights about the design features and what impact these have on the intended goals. After studies, review by researchers in areas of scientific inquiry, learning sciences and researchers in areas of gamification/game-based learning in education is being planned. After analysing the findings from study, and inputs from researchers and teachers, the GaMINLab will be refined. Thereafter a longer duration study will be planned in atleast 2 schools and additional labs.

Within the GaMINLab intervention and related study design, we are seeking input from the community on the following points: a) are there any critical pieces of literature that have been overlooked in the context of our research goal? b) Is selecting gamification elements to address the needs of autonomy, mastery, and relatedness an appropriate strategy to enhance intrinsic motivation in learning activities? d) Is it essential to measure baseline motivation (pre-intervention) within our context? e) Would it be useful to conduct a study focusing on which gamification elements will be relevant and appealing for the chosen target audience and context? This feedback will provide valuable inputs refine the intervention and related studies.

#### References

- Jaber, L. Z., Dini, V., Hammer, D., & Danahy, E. (2018). Targeting disciplinary practices in an online learning environment. *Science Education*, *102*(4), 668–692. https://doi.org/10.1002/sce.21340
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. Review of
- Kim, M. C., & Hannafin, M. J. (2011). Scaffolding 6th graders' problem solving in technology-enhanced science classrooms: A qualitative case study. *Instructional Science*, 39(3), 255–282. https://doi.org/10.1007/s11251-010-9127-4
- Moore, E. B., Herzog, T. A., & Perkins, K. K. (2013). Interactive simulations as implicit support for guided-inquiry. *Chemistry Education Research and Practice*, 14(3), 257–268. https://doi.org/10.1039/c3rp20157k
- M Sasikumar. (2016). *OLabs makes school laboratories accessible anytime, anywhere Open Source For You*. Open Source for You. <u>https://opensourceforu.com/2016/09/olabs/</u>
- Nicholson, S. (2015). A recipe for meaningful gamification. In *Gamification in Education and Business* (pp. 1–20). Springer International Publishing. <u>https://doi.org/10.1007/978-3-319-10208-5\_1</u>
- Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. 55(1), 68–78.

# ES

### Learner Perceptions on Gamifying Active Video Watching Platforms

## Pasan PEIRIS<sup>\*</sup>, Matthias GALSTER, Antonija MITROVIC, Sanna MALINEN & Raul LUMAPAS

University of Canterbury, Christchurch, New Zealand \*pitigalagepasan.peiris@pg.canterbury.ac.nz

**Abstract:** Video-based learning (VBL) provides self-paced and flexible learning. However, VBL is often a passive learning method. Active video watching (AVW) has been proposed as an approach to increase learner engagement. We investigate the motivation and perception of learners towards gamification to further increase engagement in AVW. Results from a survey in New Zealand and the Philippines show a positive perception towards integrating gamification into AVW, with learners preferring a combination of game elements rather than individual elements. Our findings provide foundations for a gamification intervention in AVW.

Keywords: Video-based learning, active video watching, gamification

#### 1. Introduction

Video-based learning (VBL) has become a popular learning method (Giannakos et al., 2016). Although VBL offers self-paced and flexible learning (Pal et al., 2019), it can be a passive activity leading to shallow learning (Seo et al., 2020). Since students learn more when they engage in more activities (Koedinger et al., 2015), there have been attempts to increase engagement in VBL by actively supporting engagement via active video watching (AVW) (Mitrovic et al., 2019). Various platforms support AVW, e.g., AVW-Space, which includes activities like writing comments on the content of watched videos or rating comments made on the videos by other learners (Lau et al., 2016; Mitrovic et al., 2016).

Gamification is the use of game design elements in non-game contexts (Deterding et al., 2011). Gamification increases learner engagement by satisfying the learner's psychological needs, such as autonomy, competence, and relatedness (Suh et al., 2018). Education is among the top sectors where gamification interventions have been researched (Hamari et al., 2014). Badges, leader boards, levels, challenges, customization, and points are the most used game elements in the educational context (Klock et al., 2020).

This paper presents a preliminary study to understand learners' perception of gamification in the context of AVW. The following research questions are addressed in this paper:

- 1. What is the perception of learners on how gamification impacts motivation?
- 2. What is the perception of integrating gamification into active video watching?
- 3. What gaming elements do potential learners consider most desirable in active video watching?

#### 2. Methodology

We designed a survey with 17 questions on demographics, motivation, experience with gamification, perception of gamifying active video watching, and game elements (the full questionnaire as well as the responses are available online:

<u>https://doi.org/10.5281/zenodo.8176246</u>). Table 1 shows the game elements we included in the survey (based on the most common game elements identified in the literature (Majuri et al., 2018)).

#### Table 1. *Game elements*

Game element	Description
Point	Positive feedback or a reward (score)
Challenge	A task that requires effort to complete
Badge	Virtual representation of an accomplishment
Leaderboard	An element that orders users according to a criterion

Following the human research ethics approval (HREC 2023/16/LR-PS), participants were recruited via advertisements on university channels. Participants' demographics are in Table 2 and Table 3.

Table 2. Gender and Institute

Institute	Gender	Count
University of Canterbury, New Zealand	Male	26
	Female	21
	Non-binary	2
University of Ateneo de Davao, Philippines	Male	38
	Female	15
	Non-binary	1
	Total	103

#### Table 3. Age distribution

Age range	Count
18-23	87
24-29	5
30-35	9
36-41	2

The survey was integrated into AVW-Space, a platform for active video watching (Lau et al., 2016). Participants were given 30 minutes to go through the platform to get an idea of AVW before completing the survey. A short introductory video on gamification was also provided at the beginning of the survey. The data analysis was conducted as a descriptive statistical analysis.

#### 2.1 Results

Eighty-three percent of respondents (n=85) would feel motivated if they saw their progress (regardless of the actual game element), while 54% (n=56) would feel motivated if they saw their peers' progress as well. Eighty-eight percent (n=91) mentioned they would feel motivated if they received rewards for completed activities (again, regardless of the actual game element).

Sixty-nine percent (n=71) would like game elements integrated into AVW platforms. Fifty-six percent (n=58) stated that gamification would improve their learning on AVW platforms.

Regarding actual game elements, 29% (n=30) selected the combination of Points, Challenges, Badges, and leaderboards as gaming elements that would most effectively motivate them to undertake activities (see Figure 1).



Figure 1. Game elements/combinations that would motivate the learners

When considering individual game elements, Points were the most popular element (n=77), followed by Challenges (n=68), Badges (n= 61), and Leader boards (n=58). However, in terms of motivation to undertake specific learning activities, participants stated Points (15%) or a combined integration of Points and Challenges (15%) as the most liked element(s) when watching videos, and Points for both activities of commenting on videos (32%) and rating comments of videos (23%) (Table 4).

Activity	Most preferred game element(s)
Watching Videos	Points - 15%
	Points & Challenges - 15%
Commenting on Videos	Points - 31%
Rating Videos	Points - 23%

To further explore whether perceptions or preferences depend on respondents' background, chi-square tests were conducted. We found that participants with previous gamification knowledge responded positively when asked about being motivated to interact with the platform if they saw their peers' progress on the platform, while participants without previous knowledge of gamification responded neutral (p = .028). This may be self-explanatory since participants who already know about gamification might have had positive experiences.

#### 3. Conclusions

Research shows that gamification effectively increases engagement (Stanculescu et al., 2016). We analyzed learners' perceptions of integrating gamification with AVW as the first step before implementing gamification in AVW. This study provides evidence for the positive user perception towards gamification that will help us design a gamification intervention for AVW-Space.

Overall, there was a positive perception of gamification impacting the learners' motivation regardless of the experience with AVW or gamification. Also, the perception towards gamification improving learning from the platform was positive, with integrating gaming elements in AVW being the same. Participants preferred having multiple game elements instead of integrating individual elements.

A threat to validity of this research is the dependence of the responses on the overall time the learners spent on AVW-Space to get an idea of what AVW is. New learners answered the survey based on their limited experience with AVW-Space. Other threats to validity are the small number and types (e.g., students with little practical experience) of respondents from two universities.

#### References

- Giannakos, M. N., Sampson, D. G., & Kidziński, Ł. (2016). Introduction to smart learning analytics: Foundations and developments in video-based learning. Smart Learning Environments, 3(1), 12.
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? A Literature Review of Empirical Studies on Gamification. 2014 47th Hawaii International Conference on System Sciences.
- Klock, A.C.T., Gasparini, I., Pimenta, M. S., & Hamari, J. (2020). Tailored gamification: A review of literature, International Journal of Human-Computer Studies, 144, 1071-5819.
- Koedinger, K. R., Kim, J., Jia, J. Z., McLaughlin, E. A., & Bier, N. L. (2015). Learning is Not a Spectator Sport: Doing is Better than Watching for Learning from a MOOC. Proceedings of the Second (2015) ACM Conference on Learning @ Scale, 111–120.
- Lau, L., Mitrovic, A., Weerasinghe, A., & Dimitrova, V. (2016). Usability of an active video watching system for soft skills training. In Proc. 1st Int. Workshop on Intelligent Mentoring Systems, ITS.
- Majuri, J., Koivisto, J., & Hamari, J. (2018). Gamification of education and learning: A review of empirical literature. In Proceedings of the 2nd international GamiFIN conference, GamiFIN 2018. CEUR-WS
- Mitrovic, A., Dimitrova, V., Lau, L., Weerasinghe, A., & Mathews, M. (2017). Supporting Constructive Video-Based Learning: Requirements Elicitation from Exploratory Studies. In E. André, R. Baker, X. Hu, Ma. M. T. Rodrigo, & B. du Boulay (Eds.), Artificial Intelligence in Education (pp. 224–237). Springer International Publishing.
- Mitrovic, A., Dimitrova, V., Weerasinghe, A., & Lau, L. (2016). Reflective Experiential Learning: Using Active Video Watching for Soft Skills Training. Proceedings of the 24th International Conference on Computers in Education, 192–201.
- Mitrovic, A., Gordon, M., Piotrkowicz, A., & Dimitrova, V. (2019). Investigating the Effect of Adding Nudges to Increase Engagement in Active Video Watching. In S. Isotani, E. Millán, A. Ogan, P. Hastings, B. McLaren, & R. Luckin (Eds.), Artificial Intelligence in Education (pp. 320–332). Springer International Publishing.
- Pal, S., Pramanik, P. K. D., Majumdar, T., & Choudhury, P. (2019). A semi-automatic metadata extraction model and method for video-based e-learning contents. Education and Information Technologies, 24(6), 3243–3268.
- Seo, K., Fels, S., Yoon, D., Roll, I., Dodson, S., & Fong, M. (2020). Artificial Intelligence for Videobased Learning at Scale. Proceedings of the Seventh ACM Conference on Learning @ Scale, 215–217.
- Stanculescu, L. C., Bozzon, A., Sips, R.-J., & Houben, G.-J. (2016). Work and Play: An Experiment in Enterprise Gamification. Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, 346–358.
- Suh, A., Wagner, C., & Liu, L. (2018). Enhancing User Engagement through Gamification. Journal of Computer Information Systems, 58(3), 204–213.

## Adding Interactive Mode to Active Video Watching

#### Ehsan BOJNORDI<sup>\*</sup>, Antonija MITROVIC, Matthias GALSTER, Sanna MALINEN & Jay HOLLAND

University of Canterbury, Christchurch, New Zealand \*ehsan.bojnordi@pg.canterbury.ac.nz

**Abstract:** Presentation skills are crucial for tertiary students and graduates but are difficult to teach. We augmented active video watching (AVW) approach with the possibility of interactions among students, and conducted an experiment with AVW-Space, an online platform which supports video-based learning. The participants watched and commented on videos first. In the second phase, the participants reviewed, rated, and responded to their peers' comments. We found that students who interacted with other students and responded to their comments increased their conceptual understanding of presentation skills.

Keywords: Active video watching, interactive learning, presentation skills

#### 1. Introduction

Presentation skills are essential for tertiary students as they play a crucial role in their further education and careers (Wats & Wats, 2009). However, they are challenging to teach and teaching such competencies demands a substantial investment of time, needs extensive practice, and requires continuous feedback from instructors (Anthony & Garner, 2016).

Video watching is one of the common methods of teaching transferrable skills (Conkey et al., 2013; Mitrovic et al., 2016). Although learning from videos has many advantages, such as providing flexible, self-controlled learning opportunities, and raising students' motivation (Sablić et al., 2021), the main disadvantage of Video-Based Learning (VBL) is the lack of interaction between the system and the learner, or between learners (Yousef et al., 2014). Lack of interaction results in low-quality learning due to limited engagement (Mitrovic et al., 2019).

Active Video Watching (AVW) was proposed recently as a means to promote selfreflective learning (Dimitrova & Mitrovic, 2022; Mitrovic et al., 2017). AVW-Space is a controlled VBL-environment, which allows teachers to integrate YouTube videos into their teaching spaces. Students can watch and comment on videos using teacher-specified micro scaffolds for reflection. Students are also able to review and rate the comments written by their peers in the system's social space in order to reinforce their learning. In this paper, we investigate and show the importance of supporting interactions between students in AVW. To provide direct interaction between students, we modified the reviewing/rating phase: Students can Like/Dislike comments, but also respond to comments, as well as previously written responses in two nested levels (level 1: respond to comment, level 2: respond to response). We present some preliminary results of the study in the next section.

#### 2. Methodology and Results

Following ethics approval, we collected data from a mandatory first-year course for all engineering students at the University of Canterbury in 2023. Participants had two weeks to complete Survey 1, watch eight videos about presentation skills and make comments on them. Survey 1 consisted of questions on demographic, training, and experiences with presentation

skills; they were also asked to write as many phrases as they could on giving presentations in terms of *structure*, *delivery and speech*, and *visual aids* (one minute for each question). At the end of this phase, we selected 238 (anonymized) comments which were provided to the whole class to review, rate, and respond. At the end of the study, students were invited to complete Survey 2, which contained the same questions on giving presentations, and additionally open-ended questions about the usefulness of reviewing, rating, and responding to others' comments.

Out of 746 students who completed Survey 1, we excluded responses of participants under 18 since they were unable to give informed consent. Also, we eliminated records of students who did not watch any videos. Of the remaining 647 participants, 69.55% were male, 29.68% were female, and 0.77% did not specify. Most participants (96.75%) were in the 18-23 age group, and (86.55%) were native English speakers. We received 493 valid responses for Survey 2. We classified the participants post-hoc based on their observable learning behaviours using the ICAP framework proposed by Chi and Wylie (2014). ICAP identifies four learning modes with increasing levels of engagement: Passive, Active, Constructive, and Interactive. In this paper, Passive students were those who have not written any comments or responses. The participants who wrote some comments are classified as Constructive if they wrote at least two high-quality comments (i.e., comments which show reflection on past experience, and planning for future presentations), while the remaining ones are classified as Active (Mohammadhassan et al., 2020). We distinguished interactive students as those who wrote at least one response to comments/responses. Table 1 presents the summary of students' activities in AVW-Space.

	Passive	Active	Constructive	Interactive	Total
	n=12	n=289	n=125	n=67	n=493
Days	1.67 (.89)	3.03 (1.88)	<b>4.23</b> (3.00)	3.16 (2.29)	3.32 (2.32)
Sessions	2.25 (1.48)	3.67 (2.53)	<b>5.22</b> (4.66)	4.04 (3.15)	4.08 (3.33)
Videos	8 (3.59)	<b>11.40</b> (4.01)	10.21 (3.17)	10.48 (3.78)	10.89 (3.83)
Comments	0	10.31 (5.49)	<b>21.25</b> (12.29)	15.90 (10.16)	13.59 (9.80)
Ratings	5.25 (9.62)	15.28 (16.40)	27.07 (27.97)	14.54 (12.84)	17.92 (20.24)
Responses				<b>5.07</b> (5.97)	
Review pages	2.58 (3.42)	8.96 (2.50)	9.09 (2.54)	8.84 (2.22)	8.82 (2.68)
CK1	12.75 (3.60)	13.65 (5.75)	14.87 (5.96)	14.37 (6.10)	14.04 (5.82)
CK2	13.25 (3.93)	13.72 (6.12)	15.86 (6.68)	<b>17.30</b> (8.73)	14.74 (6.75)
CK2-CK1	+ 0.50	+ 0.07	+ 0.99	+ 2.93	+ 0.70
T-test, α=.05	t=40, p=.35	t=19, p=.42	t=-1.58, p=.06	t=-2.75, p=.004	t=-1.73, p=.04

Table 1. Summary of Activities for Students Completing both Surveys – Mean (Std)

Bold-styled numbers show the highest values in each measure. ANOVA with the Bonferroni correction identified significant differences among different categories of students in all activities (p<.05). Constructive students used the system on significantly more days and wrote/rated significantly more comments compared to others. In terms of watching videos, Active students were significantly different from Constructive ones. Also, there were significant differences among Passive students and the others in all measures. All groups acted significantly differently in terms of writing comments (df=3, F=160.47, p<.001).

Students' responses to the conceptual knowledge questions from Surveys 1 and 2 (CK1 and CK2 respectively) were analyzed using the ontology developed by (Dimitrova & Mitrovic, 2022). The scores represent the number of ontology concepts mentioned in responses. Using the paired t-test on CK1 and CK2 for all student categories, we found that the only significant increase from CK1 to CK2 scores was for Interactive students, which indicated a considerable growth in CK score compared to other groups. ANCOVA analysis (Dependent Variable=CK2, Covariate=CK1, Fixed Factor=student groups) reported that CK1 (F=128.05, p<.001) and student groups (F=6.10, p<.001) have significant effects on CK2.

In Survey 2, we asked students about the usefulness of replying to the comments made by other students. Out of 490 responses, 390 (79.59%) were positive (e.g., "it gives others' ideas that they might not consider"), 36 (7.35%) cases were neutral, and the remaining 64 (13.06%) were negative ("Not useful").

#### 3. Conclusions

In this study, we adopted the AVW approach, and added interactive components in social space for students to acquire presentation skills. Having watched and commented on videos, students reviewed, rated, and responded to their peers' comments. The obtained results show significant differences in the amount of improvement among different student groups. Students who took advantage of the interactive mode of learning in AVW, interacting with other students, considerably enhanced their conceptual knowledge of presentation skills.

#### References

- Anthony, S., & Garner, B. (2016). Teaching soft skills to business students: An analysis of multiple pedagogical methods. *Business and Professional Communication Quarterly*, 79(3), 360-370.
- Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist*, *49*(4), 219-243.
- Conkey, C. A., Bowers, C., Cannon-Bowers, J., & Sanchez, A. (2013). Machinima and video-based soft-skills training for frontline healthcare workers. *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, *2*(1), 39-43.
- Dimitrova, V., & Mitrovic, A. (2022). Choice architecture for nudges to support constructive learning in active video watching. *International Journal of Artificial Intelligence in Education*, *32*(4), 892-930.
- Mitrovic, A., Dimitrova, V., Lau, L., Weerasinghe, A., & Mathews, M. (2017). Supporting constructive video-based learning: requirements elicitation from exploratory studies. Artificial Intelligence in Education: 18th International Conference, AIED 2017, Wuhan, China, June 28–July 1, 2017, Proceedings 18,
- Mitrovic, A., Dimitrova, V., Weerasinghe, A., & Lau, L. (2016). Reflective experiential learning: Using active video watching for soft skills training. Proceedings of the 24th international conference on computers in education,
- Mitrovic, A., Gordon, M., Piotrkowicz, A., & Dimitrova, V. (2019). Investigating the effect of adding nudges to increase engagement in active video watching. Artificial Intelligence in Education: 20th International Conference, AIED 2019, Chicago, IL, USA, June 25-29, 2019, Proceedings, Part I 20,
- Mohammadhassan, N., Mitrovic, A., Neshatian, K., & Dunn, J. (2020). Automatic assessment of comment quality in active video watching.
- Sablić, M., Mirosavljević, A., & Škugor, A. (2021). Video-based learning (VBL)—past, present and future: An overview of the research published from 2008 to 2019. *Technology, Knowledge and Learning*, *26*(4), 1061-1077.
- Wats, M., & Wats, R. K. (2009). Developing soft skills in students. *International Journal of Learning*, *15*(12).
- Yousef, A. M. F., Chatti, M. A., & Schroeder, U. (2014). The state of video-based learning: A review and future perspectives. *International Journal on Advances in Life Sciences*, *6*(3), 122-135.

## Evaluating the Assessment of Comment Quality in Learning Communication Skills using Active Video Watching

Raul Vincent LUMAPAS<sup>a\*</sup>, Antonija MITROVIC<sup>a</sup>, Matthias GALSTER<sup>a</sup>, Sanna MALINEN<sup>a</sup>, Jay HOLLAND<sup>a</sup>, & Negar MOHAMMADHASSAN<sup>a</sup>

<sup>a</sup>*University of Canterbury, New Zealand* \*raulvincent.lumapas@pg.canterbury.ac.nz

Abstract: Supporting student engagement remains one of the key challenges in videobased learning. This challenge is addressed by active video watching (AVW), a learning approach that supports engagement through different interventions, such as note-taking in the form of comments that learners submit while watching videos. One platform to support AVW is AVW-Space. Previous studies on AVW-Space detail improvements in the system, such as the integration of Artificial Intelligence and Machine Learning (ML) models in the comment feature of the system. This study investigates two machine learning models used to automatically assess the quality of comments when learning communication skills via AVW. One model is generated based on a large set of comments created by students when engaging with videos about presentation skills. For this study, a new model is developed from comments that students submitted when engaging with videos about communication skills. Results show that the new model, which was created from data on communication skills, performed better when assessing comments for communication skills compared to the model generated from comments for another skill. This has been demonstrated by the higher value of inter-rater agreement with the comment quality assessment made by human coders.

Keywords: Video-based Learning, AVW-Space, Comment Quality

#### 1. Introduction

Video-based learning increases students' motivation to learn, provides flexibility and opportunities for self-regulated learning (Chatti et al., 2016; Dimitrova & Mitrovic, 2022). However, there are still challenges in video-based learning, such as the lack of interaction between learners and teachers, resulting in passive learning (Yousef et al., 2014) and a low level of engagement (Dimitrova & Mitrovic, 2022). This need to support student engagement is addressed in active video watching (AVW). A platform to support AVW is AVW-Space developed at the University of Canterbury (Mitrovic et al., 2016; Mitrovic, Dimitrova, et al., 2017). In AVW-Space, teachers create spaces based on topics containing videos and activities. Student watch videos and write comments on the videos uploaded in the platform. AVW-Space provides a multitude of features that help in learning, such as personalized prompts called Nudges and Visualizations.

To increase engagement, Mohammadhassan et al. (2020) developed nudges as personalized reminders to students who write no comments or write only low-quality comments. This particular improvement uses Artificial Intelligence (AI) in assessing the quality of comments. Comments were classified into five categories: (1) Affirmative, negative, or off-topic, (2) Repeating, (3) Critical and analytical, (4) Self-reflective and (5) Self-regulating comments. Category 1 and 2 comments are pedagogically undesirable and are classified as low-quality comments as these comments do not convey deep thinking. On the other hand, categories 3 - 5 are considered high-quality comments as learners elaborate on the video, reflect, and relate the video content to previous experience. In the same study by

Mohammadhassan et al., a random-forest classifier was trained to assess comment quality. After selecting the best machine learning model for assessing the comment quality, the best-performing combination of the comment categories was identified, retaining categories 1 and 2, and combining categories 3-5 into category 3. This improvement enhanced student engagement, resulting in an increase in comments and the number of high-quality comments. Despite these improvements, Mohammadhassan et al. suggested that there is still room for improving the automatic assessment of comment quality. This suggestion is then explored in our study.

#### 2. Methodology and Results

The above-mentioned machine learning model was trained using comments submitted by students in ENGR101, a first-year general introductory engineering course at the University of Canterbury (UC). This course used the online resources in AVW-Space for training presentation skills. The model was trained using the comments collected from the 2017-2019 ENGR101 cohorts. The same model was then also used for the automatic assessment of comment quality in another course and for another skill: SENG202, a course in the second year for Software Engineering students, that uses AVW to teach communication skills.

As the current ML model was generated using comments submitted when learning presentation skills, there is a need to assess its performance when assessing comments submitted when leaning communication skills. To evaluate this, we analyzed comments from the most recent instances of SENG202 (from 2020 to 2022) and the classifications generated by the current model. Using the comments collected during this period (159, 691, and 688 comments for 2020, 2021, and 2022 respectively), we trained a new machine learning model specifically for communication skills. We used the testing set to analyze the performance of the current model used in AVW-Space and the new model trained using the data set generated from comments on communication skills. Two human coders have also manually assessed the quality of the comments in the testing dataset.

Comment Quality	Rater 1	Rater 2	Current Model	New Model
Category 1	3	5	1	2
Category 2	148	118	240	176
Category 3 – 5	157	185	67	130

Table 1. Distribution of Individual Ratings by the models and the human coders

Of the total of 1,529 comments, 80% (1,231 comments) were used to train the model. The remaining 20% (308 comments) were used as the test set. The number of comments used for training and testing the new model is smaller than what was used in the current model (Total: 2,343. Training: 1796 & Testing: 447). As shown in Table 1, it is noticeable that both models classified most comments as Repeating comments (Category 2), a form of low-quality comment. Both human coders had a higher rate for Category 3 as compared to the machine learning models.

Table 2. Results of Analyzing Inter-rater Agreement between the models and the human coders

Results	Current Model	New Model
Krippendorff's Alpha	0.442	0.623
Average Pairwise Cohen's Kappa	0.478	0.627
Pairwise Cohen's Kappa - System & Rater 1	0.295	0.557
Pairwise Cohen's Kappa - System & Rater 2	0.407	0.593
Pairwise Cohen's Kappa - Rater 1 & Rater 2	0.732	0.732

Table 2 shows the inter-rater agreement between the human coders and the ML models. It can be observed that human coders agree more often, having a Cohen's Kappa of 0.732, a substantial level of agreement. The results show that the human coders do not often agree with the current model. However, there is an increase in the agreement between the human coders and the new model. Despite falling short of the acceptable minimum Krippendorff's Alpha value ( $\alpha$ >0.66) (Krippendorff, 2010), the new model still significantly outperforms the current model in classifying comments, even though that the new model was trained using a smaller number of comments.

#### 3. Conclusions

In this research, we assessed two machine learning models to assess the quality of comments submitted when engaging with videos to learn communication skills. The new model was trained with comments from SENG202 classes from 2020 to 2022. The classifications generated by the current and new models were compared to those produced by two human coders. The result shows higher agreement between the human coders and the new model.

Despite the improvement in the performance of the model, there is room for further improvement, particularly in classifying lower-quality comments. With more comments, the new model generated in this study can be improved further, possibly increasing the result in the inter-rater agreement with the human coders. Creating a more generalizable model for assessing comment quality regardless of the concept or skill can also be explored to make the model applicable beyond presentation and communication skills.

#### 4. References

- Chatti, M. A., Marinov, M., Sabov, O., Laksono, R., Sofyan, Z., Fahmy Yousef, A. M., & Schroeder, U. (2016). Video annotation and analytics in CourseMapper. Smart Learning Environments, 3(1), 10. https://doi.org/10.1186/s40561-016-0035-1
- Dimitrova, V., & Mitrovic, A. (2022). Choice Architecture for Nudges to Support Constructive Learning in Active Video Watching. Artificial Intelligence in Education, 32(4), 892–930.
- Krippendorff's Alpha. (2010). In N. Salkind, Encyclopedia of Research Design (p. 670). SAGE Publications, Inc. https://doi.org/10.4135/9781412961288.n206
- Mitrovic, A., Dimitrova, V., Weerasinghe, A., Lau, L. (2016) Reflexive experiential learning: using active video watching for soft skills training. In: Chen, W. et al. (Eds.) Proc. 24th Int. Conf. Computers in Education, pp. 192-201. Mumbai, India, Nov 28 Dec 2, 2016. Asia-Pacific Society for Computers in Education (ASPCE)
- Mitrovic, A., Dimitrova, V., Lau, L., Weerasinghe, A., & Mathews, M. (2017). Supporting Constructive Video-Based Learning: Requirements Elicitation from Exploratory Studies. In E. André, R. Baker, X. Hu, Ma. M. T. Rodrigo, & B. du Boulay (Eds.), Artificial Intelligence in Education (Vol. 10331, pp. 224–237). Springer International Publishing. https://doi.org/10.1007/978-3-319-61425-0\_19
- Mohammadhassan, N., Mitrovic, A., Neshatian, K., Dunn, J. (2020) Automatic quality assessment of comments in active video watching using machine learning techniques. In: So, H.J. et al. (Eds.) Proc. 28<sup>th</sup> International Conference on Computers in Education, pp. 1-10. Asia-Pacific Society for Computers in Education. ISBN978-986-97214-5-5.
- Yousef, A. M. F., Chatti, M., & Schroeder, U. (2014). The State of Video-Based Learning: A Review and Future Perspectives. International Journal on Advances in Life Sciences, 6, 122–135.

## Preservice Teachers' Video-Based Reflection Supported by the Teacher Dashboard: An Epistemic Network Analysis

#### Huiying CAI<sup>a</sup>, Linmeng LU<sup>b</sup>, Xiaoqing GU<sup>b</sup>

<sup>a</sup> Department of educational technology, Jiangnan University, China <sup>b</sup> Department of educational technology, East China Normal University, China

**Abstract:** To evaluate how video-based reflection supported by teacher dashboard influence teachers' abilities to reflect, 48 pre-service teachers from the same university in Eastern China were recruited to participating the comparative quasi-experiment to view the same classrooms video and make some discussion how to make the instructional design better in QQ. Twenty-four pre-service teachers were assigned in group in the experimental condition adopting video-based reflection with the teacher dashboard. Twenty-four pre-service teachers were the control condition adopting video-based reflection only. After analyze the teachers' collaborative discourse in conjunction with descriptive analysis alongside ENA, the results showed that video-based reflection supported by the teacher dashboard can improve the quality of reflection discussion and engage in deep and critical reflection to generate reason-based decisions with more teaching improvement.

**Keywords:** Video-based reflection, Teacher dashboard, Teacher collaborative discourse, Epistemic network analysis

#### 1. Introduction

In teacher education, teachers' reflection is considered as an effective method of teacher learning. Reflection is "deliberate thinking about action with a view to its improvement" (Hatton & Smith, 1995, p. 40). Through their critical reflections, teachers can examine their past experiences and current practice in order to gain insights into their future teaching and continuing professional development (KleinKnecht & Schneider, 2013). Since the 1960s, teacher video-based reflection has been concerned as an effective means to promote teachers' reflection in the field of teacher education (Harford, MacRuairc & McCartan, 2010; Kleinknecht & Schneider, 2013). Teacher video-based reflection refers to the method that teachers watch, analyze and discuss teaching videos, and promote teachers to deeply reflect on teaching practice, so as to improve their professional ability (Hamel & Viau-Guay, 2019). Although classroom videos play a positive role in promoting teachers' reflection, the challenge of overwhelming data inherent in classroom video would be not beneficial for pre-service reflection due to the long and excessive content of classroom videos (Chen, 2020).

To address the problem, the technological approach of the teacher dashboard which visualizing the analyzing information of teacher and student discourse in classroom video by the form of indicator and graph was an alternative solution to facilitate teacher's video-based reflection. Nevertheless, research exploring how the teacher dashboard influence teachers' reflection process remains scares. Consider a growing body of epistemic network analysis (ENA) has been conducted to understand teachers' reflective thinking (e.g., Shaffer & Ruis, 2017). Therefore, in this study we intend to explore the effectiveness of the teacher dashboard on teacher's video-based reflection by focusing on analyzing teachers' online collaborative discourse by ENA.

#### 2. Experiment design

#### 2.1 Participant and research context

In order to evaluate the effectiveness of video-based reflection supported by the teacher dashboard, the quasi-experiment was conducted in a university in Eastern China. Twenty-four pre-service teachers form in eight groups to make video-based reflection with the teacher dashboard in the experimental condition (EC). Twenty-four preservice teachers form in eight groups to make video-based in the control condition (CC).



Figure 1. Research procedure

In this study, groups in the EC and CC were organized to completed the same video-based reflection task on different days but in the same computer laboratory using the same procedures. The research procedure was shown in figure 1.

#### 2.2 Coding framework

According to sociocultural theory, data from the group discourse recorded in QQ will act as a lens to investigate how pre-service teachers in groups engage into reflection (Wegerif et al., 2010). Therefore, in order to explore the different characteristics of preservice teachers' reflective talk in both conditions, we adopted a coding framework based on Mercer's (2000) original three types of talk, namely cumulative talk, disputational talk and exploratory talk. Each dimension contains some sub-dimension to describe the teacher's talk, which was shown in Table 1.

Type of talk	Category	Description		
Cumulative talk	Information description (C.ID)	Teachers exchange and share information related to collaborative teaching and research tasks.		
	Information inquiry (C.II)	Teachers ask questions about facts or task-related issues that do not involve reasoning.		
	Suggestion (C.SU)	Teachers offer their own suggestions for completing the task, such as suggesting ways to complete the task efficiently.		
	Coordination (C.CO)	Teachers monitors the progress of this collaborative task, including mobilizing teacher participation, planning the division of labor, etc.		
Disputational talk	Common understand (D.CU)	Teachers simply echo each other's ideas.		
	Evaluation without justification (D.EV)	Teacher expresses agreement or disagreement with other members' views or opinions without detailing the reasons.		
	Objection (D.OB)	Teachers' rebuttals with simple evidence, but no suggestions for improvement.		
Exploratory talk	Higher order question (E.HO)	Teachers ask questions to provoke others to think others' ideas, and the answers to these questions involve teachers' reasoning.		
	Agreement with justification (E.AJ)	Teachers justify their ideas with reasons or provide additional arguments in response to others' follow-up questions.		
	Disagreement with justification (E.DJ)	Teachers raise objections and give explanations.		
	Alternative plan (E.AP)	The teacher's ideas for optimizing the original teaching involved, and the teacher's optimization of others' ideas.		

Table 1. The coding framework

Table 2. Example of encoded data

Teerberr	T	Code	
Teacher	1 urn of taik	Code	
3	First of all, 1. What are the teaching objectives of this lesson?	Cumulative talk-	
		Information inquiry	
3	You can also look at the materials in your hands first.	Cumulative talk-	
		Suggestion	
3	I think the teaching objectives of this lesson can be in three aspects.	Cumulative talk-	
		Information	
		description	
27	I think the teaching objectives given in this lesson plan are the teaching objectives of this lesson. The teacher lists the knowledge skills with more obvious behavioral verbs such as say out, distinguish, name, and also the degree of performance, while these behaviors are observed.	Exploratory talk - Agreement with justification	
35	The factual knowledge and procedural knowledge in the course knowledge have a large overlapping part, and I think it is enough to integrate them into one aspect.	Exploratory talk - Agreement with justification	

Messages sent within the QQ were segmented into turns before coding in Chinese. Acting as the unit of coding, a turn of talk can be coded in one category based on the above-mentioned three-part typology and descriptions of teachers' group talk in video-based reflection. The example was shown in Table 2.

A total of 1595 messages were collected for independent coding in this study, 740 messages from EC and 855 messages from CC. The coding was done by three coders. The coding consistency coefficient was 0.91 for coders 1 and 2, and 0.89 for coders 1 and 3. This proves that the coding results are reliable. For the coding results with different opinions, the coders negotiated and finally unified their opinions.

#### 2.3 Data analysis

We used a mixed methods approach to analysis (Creswell & Clark, 2017). First, we used a discourse analytic methodology to qualitatively code pre-service teacher' talk (Hardy et al., 2004). Then we applied ENA and statistical analysis to compare pre-service teacher' talk across treatments.

#### 3. Result

#### 3.1 The overall different of collaborative discourse in CC and EC

The categories and frequency distribution of teachers' talk were shown in Table 3. As can be seen from Table 3, the proportions of codes related to cumulative talk are higher in CC (53.10%) than that in EC (36.22%). While, the proportions of codes related to exploratory talk are lower in CC (23.86%) than that in EC (40.00%). The proportions of codes related to disputational talk in CC and EC are similar, 23.04% and 23.78% respectively.

Type of talk	Category	CC	EC
	C.ID	217 (25.38%)	107 (14.46%)
	C.II	102 (11.93%)	40 (5.41%)
Cumulative talk	C.SU	51 (5.96%)	57 (7.70%)
	C.CO	84 (9.82%)	64 (8.65%)
		454 (53.10%)	268 (36.22%)
	D.CU	119 (13.92%)	119 (13.92%)
	D.EV	102 (13.78%)	102 (13.78%)
Disputational talk	D.OB	70 (8.19%)	70 (8.19%)
		197 (23.04%)	176 (23.78%)
	E.HO	52 (7.03%)	52 (7.03%)
	E.AJ	8 (0.94%)	8 (0.94%)
Exploratory talk	E.DJ	22 (2.97%)	22 (2.97%)
	E.AP	45 (5.26%)	45 (5.26%)
		204 (23.86%)	296 (40.00%)

Table 3. Coding frequency statistics table

The chi-square tests we conducted to determine if these differences were statistically significant indicated that they were, in four cases, i.e.: all ( $X^2$ =58.037, p<0.001), cumulative talk ( $X^2$ =20.924, p<0.001), disputational talk ( $X^2$ =15.187, p<0.001), and exploratory talk( $X^2$ =68.009, p<0.001). It indicated that in the video-based reflection supported by the teacher dashboard, teachers' talk focus more on exploratory talk and less on cumulative talk.

#### 3.2 The overall different ENA networks in CC and EC

The epistemic networks of teachers in the EC and CC during the process of online discourse were shown in Figure 2.



*Fig 2.* The Comparison of Epistemic Networks between the CC (left) and the EC (middle) and the network means of the two groups

The black dots in the ENA network shown in Figure 2 represent the positions of each code (e.g. Cumulative talk- information description, C.ID), and the connecting lines between the nodes represent the associations between the cognitive elements. The strength of the connections is reflected in the width of these lines. The thicker this line became and the higher the coefficient of connection between these two types of teacher talk. As shown in Figure 2, the mean of group teachers' reflective talk types in the CC was close to the right side of the X-axis, while the center of mass of the structure of group teachers' reflective talk types in the EC was close to the left side of the X-axis. The t-test showed that there was a significant difference between the group teachers' reflective talk types in the CC and EC on the X-axis (p=0.00<0.05). it means that there are significant different of cognitive network shown in the online discourse in both conditions.

Connection coefficients		EC
C. Coordination - C. Information description (C.CO-C.ID)	0.26	0.12
C. Information inquiry - C. Information description (C.II-C.ID)	0.28	0.11
D. Common understand - C. Information description (D.CU-C.ID)		0.17
E. Agreement and justification - C. Information description (E.AJ-C.ID)	0.18	0.29
E. Alternative plan - C. Information description (E.AP-C.ID)		0.25
E. Alternative plan - E. Agreement and justification (E.AP-E.AJ)		0.36

 Table 4. Connection coefficients of the overall ENA networks in CC and EC

As shown in Table 4, we found that the strength of the connections between "information description" and "common understand" (C.ID-D.CU, 0.29), between "information inquiry" and "information description" (C.II-C.ID, 0.28)," "information description" and "evaluation without argumentation" (C.ID-D.EV, 0.26)" are all stronger in the CC than in its EC. Conversely, the strength of the connections between "alternative plan" and "agreement and justification" (E.AP-E.AJ, 0.36), between "information description" and "agreement and justification" (C.ID-E.AJ, 0.29)," "alternative plan" and "information description" (E.AP-C.ID, 0.25)" are more stronger in the EC than in its CC. This indicates that with help of the video-based reflection with the teacher dashboard, group teachers have higher quality collaborative reflection discourse. During the group teachers' talk in the video-based reflection, teacher would like to description information about classroom video for common understanding, and ask questions about facts or task-related issues that do not involve reasoning. While in the video-based reflection about classroom video for common understanding, and ask questions about facts or task-related issues that do not involve reasoning. While in the video-based reflection about classroom video for common understanding, and ask questions about facts or task-related issues that do not involve reasoning. While in the video-based reflection about classroom video for common understanding, and ask questions about facts or task-related issues that do not involve reasoning. While in the video-based reflection about classroom video for common understanding, and ask questions about classroom video for optimizing others' ideas to form alternative plan, justify their ideas with

reasons or provide additional arguments in response to others' follow-up questions. Teacher will conduct higher level teaching reflection discussion, and generate teaching decisions with more teaching improvement.

#### 4. Discussion and conclusion

The use of classroom videos to reflect on and analyze teaching practices is a valuable tool in teacher education (Sun & van Es, 2015). The results of this study build on this existing knowledge and show that video-based reflection with the teacher dashboard has positive effect on teachers' collaborative discourse, which can contribute to teachers' ability of reflection. Our findings from the pilot study provide important implications for designing systems and for future long-term research and practice aiming at improving the approach of video-based reflection on classroom video for teacher education and profession development supported by the teacher dashboard.

#### Acknowledgements

We would like to thank all the people who prepared and revised previous versions of this document. This research was funded by National Social Science Fund of China (CCA220318).

#### References

Chen, G., Chan, C. K., Chan, K. K., Clarke, S. N., & Resnick, L. B. (2020). Efficacy of video-based teacher professional development for increasing classroom discourse and student learning. Journal of the Learning Sciences, 29(4-5), 642-680.

Creswell, J. W., & Clark, V. L. P. (2017). Designing and conducting mixed methods research. Sage publications.

Hamel, C., & Viau-Guay, A. (2019). Using video to support teachers' reflective practice: A literature review. Cogent Education, 6(1), 1673689.

Hardy, J., Hall, C. R., & Hardy, L. (2004). A note on athletes' use of self-talk. Journal of Applied Sport Psychology, 16(3), 251-257.

Harford, J., MacRuairc, G., & McCartan, D. (2010). 'Lights, camera, reflection': using peer video to promote reflective dialogue among student teachers. Teacher development, 14(1), 57-68.

Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. Teaching and teacher education, 11(1), 33-49.

Kleinknecht, M., & Schneider, J. (2013). What do teachers think and feel when analyzing videos of themselves and other teachers teaching?. Teaching and teacher education, 33, 13-23.

Mercer, N. (2002). Words and minds: How we use language to think together. Routledge.

Shaffer, D., & Ruis, A. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. Handbook of learning analytics.

Sun, J., & Van Es, E. A. (2015). An exploratory study of the influence that analyzing teaching has on preservice teachers' classroom practice. Journal of teacher education, 66(3), 201-214.

Wegerif, R. (2010). Mind expanding: teaching for thinking and creativity in primary education: Teaching for Thinking and Creativity in Primary Education. McGraw-Hill Education (UK).
# Extraction of Characteristic Answering Behavior Using Handwritten Log Data

# Junya ATAKE<sup>a</sup>\*, Taito KANO<sup>a</sup>, Kohei NAKAMURA<sup>a</sup>, Chia-Yu HSU<sup>a</sup>, Izumi HORIKOSHI<sup>b</sup> & Hiroaki OGATA<sup>b</sup>

<sup>a</sup> Graduate School of Informatics, Kyoto University, Japan
 <sup>b</sup> Academic Center for Computing and Media Studies, Kyoto University, Japan
 \*atake.junya.86t@st.kyoto-u.ac.jp

**Abstract**: We extracted learners' characteristic answering behaviors from handwritten process log data and investigated whether learners' situations could be inferred based on these characteristics. The result showed we were able to extract several characteristic answering behaviors, such as stopped pen stroke and late start. Furthermore, we examined the learners' situation for each feature in the actual answering process. The results revealed that several characteristic answering behaviors such as learners' stumbling or giving up. These results imply that handwritten process log data can allow teachers to capture learners' situations and support teachers' interventions.

Keywords: Handwriting, log data, answering behavior, learning analytics

#### 1. Introduction

In Japan, under the Global and Innovation Gateway for All (GIGA) school project promoted by the Ministry of Education, Culture, Sports, Science, and Technology, a digital device was provided to each learner, and educational log data were accumulated. A typical digital device used in the classroom is a handwriting tool that allows learners to write directly on the devices. One advantage of using such tools is that learning logs, which represent the writing process, answer time, and handwriting, can be collected and analyzed. It is important for teachers to analyze the learners' answers to provide appropriate instructions using the handwritten process log data. However, it is difficult for teachers to analyze the learners' answers and understand the situation within a limited time.

Several studies have been conducted on learners' handwriting processes, such as the detection of answer stuck points from character recognition and the pen-stroke, the time interval of one stroke in handwriting (liyama et al., 2017), and the detection of deleted or revised strokes using support vector machines (Maeda et al., 2015). These methods make it easier for teachers to understand learners' stuck points, deletions, and corrections. However, both studies focused on their answering contents and did not focus on their answering behaviors or situations.

This study extracted the learners' characteristic answering behaviors from the handwritten process log data, and investigated whether inferring the learners' situation is possible. "Characteristic answering behaviors" refer to "Interval of stopping Pen-stroke," and "Eraser usage," and so forth, which might reflect that learners spend time reflecting on and correct their mistakes. In addition, we attempted to identify the learners' situations based on the extracted characteristic behaviors. We set the following research question for this study.

RQ: Can we infer the learners' situations by extracting their characteristic answering behaviors from the handwritten process log data?

We expect that the inference of learners' situations from the features extracted in this study can contribute to automatically detecting learners' situations from the data, helping teachers understand their situations, and providing them with appropriate support.

# 2. Methods

# 2.1 Learning and Evidence Analytics Framework (LEAF)

We used handwritten process log data collected from the LEAF system for analyses. The LEAF system is a learning analytics platform that supports teaching and learning by analyzing accumulated educational data (Ogata et al., 2018). It consists of three main sub-tools: BookRoll, Moodle, and LogPalette, and log data are stored in the Learning Record Store (LRS). BookRoll is an e-book reader through which teachers upload learning materials to their learners. Moodle is a Learning Management System (LMS) used to manage courses and materials. LogPalette is a learning analytics dashboard that visualizes the interactions of learners. Learners can use the handwriting tools to answer the learning materials that the teacher has registered on BookRoll, and the teacher can replay the learners' answering processes on LogPalette.

#### 2.2 Data and Analyses

This study's handwritten process log data were from the context in which a teacher asked learners to answer questions in a Japanese junior high school mathematics class. The learners started answering the questions simultaneously when instructed by the teacher, and answered three questions. Handwritten process log data includes the time (Timestamp) when each stroke was added and its position (x-coordinate, y-coordinate). We conducted two analyses of the data, as shown in Figure 1.



Figure 1. Analysis Flow.

In Analysis 1, we visualized the log data as time-series graphs and extracted the characteristic answering behaviors. In these graphs, the x-axis shows the elapsed time, and the y-axis shows the number of strokes and eraser usage per 10 seconds.

In Analysis 2, we examined what kinds of situations could be observed in each characteristic answering behavior extracted from Analysis 1. We observed the actual answering process using the replay, and then examined whether there was a relationship between each characteristic and the learners' situations.

#### 3. Results and Discussion

3.1 Analysis 1: Visualizing Log Data and Extracting Characteristic Answering Behaviors

First, we plotted the average number of strokes (Figure 2. (a)) and erasers used by all learners (Figure 2. (b)). We were able to ascertain the trends in the average number of strokes and the amount of eraser usage for each learner. Three peaks were observed in Figure 2. (a) for the average stroke. These peaks imply the culmination of the answers to the three questions.



(b) The Average Number of Using Erasers for All Learners.

Figure 2. Average Number of Strokes and Erasers.

Next, we plotted the answering behavior of all the learners (Figure 3) and extracted the characteristic answering behavior. Six characteristic features were identified (Figure 4).

Figure 4. (a), (b), (c), and (d) shows the behaviors characterized by how much time the strokes were stopped when the learners answered the questions. Specifically, Behavior A indicates the learners stopped the strokes for more than 50% of the answering time, Behavior B indicates the learners stopped for more than 80%, and Behavior C indicates the learners stopped for less than 20%. In addition, Behavior D indicates the learners paused and continued several times when answering. Every time they paused, the strokes were stopped for more than five minutes.

On the other hand, Figure 4. (e) shows the behavior characterized by the amount of using erasers. Behavior E indicates the learners used erasers with the amount above the average. Finally, Figure 4. (f) shows the behavior (Behavior F) that the learners waited for more than five minutes and started to answer after the teacher gave the instruction. It is noted that these features were not mutually exclusive and might overlap.

In this manner, we were able to extract some characteristic answering behaviors from the results. These features may reflect the learners spent time thinking and corrected mistakes. However, it was difficult to identify the learners' situation from these results alone, and it proved necessary to observe the state of the actual answers. Therefore, a more detailed analysis was performed in Analysis 2.



Figure 3. Answering Behaviors for Each Learner.





Figure 4. Characteristic Answering Behaviors.

# 3.2 Analysis 2: Grasping the Situation of the Learner in Each Characteristics from the State of the Actual Answers

In Analysis 2, we replayed the actual handwriting process and observed what kinds of learners' situations could be understood from the features extracted from Analysis 1. We were able to identify the learners' situations, as shown in Table 1.

Behavior	Feature (The number of learners)	Learner situation (The number of learners who matched to the situation)		
А	Stopped pen stroke over the half (n = 5)	Give up on an answer halfway (n = 5)		
В	Mostly stopped pen stroke (n = 6)	Give up on an answer halfway (n = 6)		
С	Few stopped pen strokes (n = 6)	<ul><li>(1) Keep hands moving (n = 6)</li><li>(2) Tend to get correct (n = 6)</li></ul>		
D	Few stopped pen strokes (n = 5)	Answer only certain questions (n = 5) Give up on an answer halfway (n = 5)		
Е	A Large amount of using erasers (n = 5)	Give up on an answer halfway (n = 5)		
F	Late start (n = 2)	No ideas (n = 1) Have trouble opening materials (n = 1)		

Table 1. Characteristic Answering Behaviors and Learners' Situations

The results indicate that some features can suggest situations such as learners' stumbling or giving up. These features have the potential to help the teachers understand learners' situations and provide support interventions.

# 4. Conclusion & Future Work

In this study, we set the research question as "Can we infer the learners' situation by extracting the learners' characteristic answering behaviors from the handwritten process log data?" To answer this question, we visualized time-series data to extract characteristic answering behaviors, and analyzed the answering process for each of them. From the results, we identified some features such as stopped pen strokes and eraser use. Additionally, some features suggest unique situations for learners. Although further research is required, we expect that these results are not confined to mathematics courses. In future research, it would be conceivable to support teacher interventions using these features. If the learner's situation is automatically identified from the data, it is possible to assist the teacher in understanding the learner's situation during the class.

# Acknowledgements

This study was partially supported by NEDO JPNP20006, JSPS KAKENHI JP22K20246, and JP23H00505.

#### References

- liyama , M., Nakatsuka, C., Morimura , Y., Hashimoto, A., Murakami, M., & Mino, M. (2017). Detecting answer stuck point using time intervals of pen strokes. *Transactions of Japanese Society for Information and Systems in Education, 34* (2), 166-171.
- Maeda, Y., Miura, M., & Kobayashi, Y. (2015, August). Detecting erase strokes for understanding learning states from online handwritten notes. In *Proceedings of Informatic Education Symposium* (pp. 109-116).

Ogata, H., Majumdar, R., Akçapınar , G., Hasnine , M. N., & Flanagan, B. (2018). Beyond learning analytics: Framework for technology-enhanced evidence-based education and learning. In *International Conference on Computers in Education (ICCE 2018)* (pp. 486-489).

# Interaction Patterns between Learners and AI Tools for English Writing

# Yun Shu XIE, Jin Ho JANG, Su Yeon KIM & Young Hoan CHO\*

Seoul National University, South Korea \*yhcho95@snu.ac.kr

**Abstract:** The advancement of artificial intelligence (AI) has drawn educators' attention to its educational potential. However, the efficacy of AI depends on learners' effective interaction, avoiding either neglect or excessive dependence. Despite the significance of learner-AI interaction, research on interaction patterns remains limited. This study investigates how EFL (English as a foreign language) learners interact with an AI tool for English writing and explores the effects of different interaction patterns on their writing performance. Through an experiment involving 29 EFL undergraduates, three distinct interaction patterns emerged, exhibiting significant differences in their engagement with the AI tool. The comparisons of the three clusters indicate that different AI interaction patterns lead to varied interaction approaches, and not all learners equally benefit from AI's potential. In order to promote productive learner-AI interaction in educational environments, instructors should provide personalized support and feedback.

Keywords: Artificial intelligence, Interaction, English writing, EFL

#### 1. Introduction

Technological advancement has AI playing an increasingly vital role in the education field, particularly in enhancing English learning. The AI tools have been demonstrated to enhance learners' writing skills by systematically providing feedback about incorrect grammar, spelling and so on (Liu et al., 2021). As a tutor, a companion, or a fellow student, AI has the potential to exert cognitive and affective influences on learners (Engwall & Lopes, 2020). However, it should be noted that not all learner-AI interactions are uniformly effective, the success of these interactions depends on how they are conducted (Wang et al., 2023). Limited understanding exists regarding the process of learning through learner-AI interactions, and the variations in interactions among different learners, although such research could contribute to the design of AIED by illuminating effective ways of learner-AI interaction. The purpose of this study is to explore learner-AI interaction patterns during English writing tasks and the effects of interaction patterns on learners' performance.

# 2. Literature Review

Al has been reported to positively contribute to enhancing learners' English writing skills. EFL learners provided with Al-based semantic and syntactic feedback outperformed those without Al support (Hwang et al., 2023). Learners of different competence levels may interpret the same learning activity differently, resulting in diverse approaches to interacting with Al. Wang et al. (2023) identified four learner-Al interaction clusters based on system usage data in EFL learning: effective learners, passive learners, well-balanced learners, and inefficient learners. They observed that the primary beneficiaries were those utilizing deep learning methods, engaging in the critical reception of Al feedback rather than following it mechanically. Previous research on English writing with an Al translator also categorized learner-Al interaction into Al-dependent, limited, and collaborative interaction (Kim et al., 2023). Collaborative interaction

exhibits the most effective patterns as learners go beyond copying the AI's responses, actively modifying inputs to meet their specific needs and fostering two-way interactions.

#### 3. Method

This study involved 29 EFL undergraduates from diverse academic majors in South Korea. QuillBot (https://quillbot.com), a user-friendly AI system, was used to support English essay writing. Participants had no prior experience with QuillBot and received instructions on the AI tool. There was a practice session before carrying out English essay writing in two conditions: (1) independent writing and (2) AI-supported writing conditions. Participants conducted think-aloud while carrying out the writing tasks in a laboratory, and all activities were recorded with video. Three researchers independently analyzed videos using a coding scheme of learner-AI interaction and the inter-rater reliability was high (Cohen's Kappa at .96). All disagreements were resolved through discussions. English writing performance was assessed by two high school teachers using a rubric.

#### 4. Results

This study found that the AI tool, QuillBot, was helpful for EFL learners who might lack English writing skills. Wilcoxon Signed-Rank Test showed that learners achieved significantly higher scores in the AI-supported writing condition (M = 7.37, SD = 1.50) than in the independent writing condition (M = 6.55, SD = 1.72, Z = 3.01, p < .01).

This study explored learner-AI interaction patterns by investigating the number of clusters with hierarchical cluster analysis and carrying out k-means cluster analysis. Three types of learner-Al interaction patterns were identified: limited, Al-dependent and collaborative interactions (see Table 1). Cluster 1, limited interaction patterns, showed high independence with minimal AI usage. Cluster 2, AI-dependent interaction patterns, frequently relied on the Al tool to accomplish tasks and seldom modified Al recommendations, consistently monitoring the progress of English writing. It is noteworthy that Cluster 2 identified the benefits of the AI tool more favorably than the others. Cluster 3, collaborative interaction patterns, valued feedback and recommendation from the AI tool and engaged in critical thinking to revise the recommended words and sentences rather than mechanically following it. Furthermore, according to the result of Kruskal Wallis H test, there were significant differences between mean rank of three clusters in task strategy (H = 17.06, p < .01), individual work (H = 22.45, p < .01), revision-oriented interaction (H = 13.77, p < .01), acceptance-oriented interaction (H = 15.36, p < .01), performance monitoring (H = 9.91, p < .01), interaction monitoring (H = 12.34, p < .01), positive evaluation of AI (H = 11.13, p < .01), and negative evaluation of AI (H = 13.08, p < .01).

An ANCOVA was conducted to examine the influence of learner-Al interaction patterns on English writing performance in the Al-supported writing condition, using English writing performance in the independent writing condition as a covariate. Although the independent writing performance significantly influenced the Al-supported writing performance (F = 23.32, p < .01), there was no significant influence of learner-Al interaction patterns (F = .17, p = .849).

	Sub Codes	Clus	ter 1	Clus	ter 2	Clus	ter 3	· · · ·	
Codes		(n = 15, 52%)		(n = 6, 21%)		(n = 8, 27%)		н	p
		М	SD	М	SD	М	SD	7	-
Planning	Task analysis	2.87	4.30	1.04	1.76	3.30	5.07	.84	.658
	Task strategy	.00	.00	1.26	1.39	.00	.00	17.06	<.01
	Al function exploration	0.737	1.59	0.65	1.36	2.48	3.56	2.20	.332
	Individual work	73.18	7.42	38.11	6.56	52.17	6.52	22.45	<.01
Task	Revision- oriented interaction	2.31	3.36	0.45	1.11	9.51	7.34	13.77	<.01
Performance	Acceptance- oriented interaction	10.72	6.37	26.71	7.40	21.42	7.02	15.36	<.01
	Rejection- oriented interaction	4.49	4.78	7.09	3.20	3.86	2.44	3.81	.149
	Performance monitoring	5.33	5.71	15.35	3.97	5.26	5.11	9.91	<.01
	Interaction monitoring	.00	.00	1.63	2.67	.00	.00	12.34	<.01
Monitoring	Positive evaluation of Al	0.31	0.61	5.67	5.48	1.99	2.11	11.13	<.01
	Negative evaluation of Al	0.05	0.20	2.05	2.28	.00	.00	13.08	<.01

Table 1. Types of interaction between learners and the AI tool

# 5. Conclusion

This study found three types of learner-Al interaction patterns (i.e., limited, Al-dependent and collaborative interactions), each of which showed a distinct way of interacting with Al for English writing. This result is consistent with Kim et al. (2023). Compared to Kim et al. (2023), the Al tool in this study provided continuous feedback based on input, and engaged in a more flexible, personalized and active interaction with learners. Consequently, a wider range of learner-Al interactions were analyzed. Due to the nature of the task, differences of interaction patterns might have no statistically significant effect on English writing performance. This study explored the influence of learner-Al interaction patterns on an Al-enhanced task, not an Al-enabled task, which could not be conducted without Al tools. The influence of learner-Al interactions would be enlarged in the Al-enabled task like English conversation with Al. This study implies that instructors should provide scaffoldings to enhance collaborative interaction with Al tools and encourage learners to reflect on their interaction patterns in or after using Al tools for learning activities.

#### References

- Engwall, O., & Lopes, J. (2022). Interaction and collaboration in robot-assisted language learning for adults. *Computer Assisted Language Learning*, *35*(5-6), 1273-1309. https://doi.org/10.1080/09588221.2020.1799821
- Hwang, W. Y., Nurtantyana, R., Purba, S. W. D., Hariyanti, U., Indrihapsari, Y., & Surjono, H. D. (2023). Al and recognition technologies to facilitate English as Foreign Language writing for supporting personalization and contextualization in authentic contexts. *Journal of Educational Computing Research*, 61(5), 1008–1035. https://doi.org/10.1177/07356331221137253
- Kim, H., Cho, Y. H., & Park, S. (2023). Exploring the interaction patterns between learners and Al translator in English writing. *The Journal of Educational Information and Media*, 29(1), 201-228. https://doi.org/10.15833/KAFEIAM.29.1.201
- Liu, C., Hou, J., Tu, Y. F., Wang, Y., & Hwang, G. J. (2021). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments,* 1-19. https://doi.org/10.1080/10494820.2021.2012812
- Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in Al-supported learning: A study of human-Al interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education, 194*, 104703. https://doi.org/10.1016/j.compedu.2022.104703.

# A holistic visualisation solution to understanding multimodal data in an educational metaverse platform – Learningverse

Yanjie SONG<sup>a\*</sup>, Jiaxin CAO<sup>b</sup>, Lei TAO<sup>c</sup>, Dragan Gašević<sup>d</sup>

<sup>abc</sup>Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR, China <sup>d</sup>Faculty of Information Technology, Monash University, Australia <u>\*ysong@eduhk.hk</u>

**Abstract:** Traditional digital learning environments faced challenges in obtaining comprehensive user interaction data, often yielding fragmented insights without a cohesive visual representation. The emergence of metaverse platforms has enriched this landscape, enabling detailed user activity representation with multimodal data through avatars. However, how to understand the multimodal data related to teaching, social and cognitive presences underpinned by the 'Community of Inquiry' theoretical framework in the metaverse is a big challenge for educators. This study introduces a holistic visualisation solution to bridge this gap, ensuring a better understanding of avatars' behaviours in an educational metaverse platform – Learningverse developed by our research team. The solution captures a range of multimodal data in Learningverse, such as avatar location, behaviours, emotions, and conversation. Key visualisation elements include heatmaps, points, and arrows, each with distinct informational value. In the future, integrating the solution with multimodal learning analytics is our next step work to understand teaching, social and cognitive presences.

**Keywords:** Metaverse, Learningverse, holistic visualisation solution, Community of Inquiry

# 1. Background

The evolution of digital learning environments has ushered in new opportunities to observe and analyse user activities and behaviours (Diwanji et al., 2020; Knox et al., 2020). In traditional virtual learning environments, capturing comprehensive and detailed data of user interactions remains challenging, often resulting in a fragmented understanding of the activities and behaviours (Willermark & Gellerstedt, 2022). The introduction of the metaverse provides an immersive platform for social interactions and collaboration, with avatars representing users, demonstrating their activities (Zallio & Clarkson, 2022). The need for a holistic visualisation solution to understanding multimodal data in the metaverse becomes paramount, paving the way for a deeper, and more comprehensive understanding of users' activities and behaviours.

Nevertheless, the avatars in existing metaverse platforms such as Spatial, Verbela, and Second Life cannot mirror the real users, making users' multimodal data such as facial expressions, gestures, poses and conversations harder to capture, and learning process less engaging and interactive (Song et al., 2023). Against this background, our research team developed the metaverse platform – Learningverse, which is introduced in the next section.

# 2. An Educational Metaverse Platform – Learningverse

Learningverse is a 3D educational metaverse platform developed specifically for learning and training purposes (Song et al., 2023). The platform has three unique features: (1) it is underpinned by the 'Community of Inquiry' framework (Garrison et al., 2010) which supports varied pedagogical social-constructivist designs and promotes deep and meaningful learning experience; (2) users can act as avatars mirroring themselves in the metaverse simply on a common computer with a webcam, thus lowering the threshold for users' accessibility; and (3) users can interact with the virtual objects and other avatars, and are able to use custom collaborative tools and 3D models to create their own virtual learning spaces or activities. Owning to these features, avatars' activities and behaviours in a multimodal form in Learningverse can be captured. In the next section, we present our proposed holistic visualisation solution to understanding multimodal data in the metaverse platform.

# 3. Design of a Holistic Visualisation Solution

#### 3.1 Multimodal Data in Learningverse

In Learningverse, multimodal data can be logged. These include avatar location, tracked by the 3D axis positioning of the avatar, behaviours such as upper-body posture and gestures, emotions inferred from facial expressions, activity records on tasks set by educators, and conversations captured in chat rooms or transcribed from oral communications (See Table 1).

Multimodal data	Raw data	Туре			
Avatar location	Avatar 3D axis in Learningverse	Distribution			
Behaviour	Upper-body posture	Density			
	Gesture	Density			
Emotion	Facial expression	Density			
Task completion	Activity records on tasks	Trajectory			
Conversation	Chat room	Occurrence			
	Oral speaking to text	Occurrence			

 Table 1. Properties of Multimodal Data in Learningverse

#### 3.2 A Holistic Visualisation Solution to Understanding Multimodal Data

The holistic visualisation solution incorporates three primary visualisation elements: (1) Heatmap: illustrates the concentration and range of avatar activity within a designated area with a deeper colour indicating heightened activity or emotional intensity; (2) Point: Serves as an indicator, marking instances of conversation; and (3) Arrow: Represents the trajectory of avatars as they navigate through various learning tasks (see Figure 1).

The holistic visualisation solution is grounded in representing the activity locations of avatars in Learningverse, predominantly using a heatmap format (See Figure 2). This approach visualises the intensity and scope of their spatial activity (Pokojski et al., 2021). Data is captured at 30-second intervals to ensure the accumulation of meaningful information. The choice of this interval, influenced by research from Emerson et al. (2020), ensures sufficient feature changes across modalities.

Such an approach strikes a balance, allowing for granular data acquisition without overwhelming the system or the educators. Subsequently, the system matches this data with avatars' specific locations and their respective timestamps, aligning these with the heatmap on a thumbnail view. The rich multimodal data in Learningverse can be presented through the holistic visualisation solution, offering a flexible, interactive, and selectable mode. This design allows educators and learners to reflect upon and review their entire learning process and outcomes.



Figure 1. Design of the Holistic Visualisation Solution for Multimodal Data



Figure 2. Demonstration of the Holistic Visualisation Solution

# 4. Future work

The current visualisation solution predominantly presents standard activity records, with less emphasis on pedagogical insights and actionable takeaways. The solution aims to offer a foundational understanding of multimodal data in the metaverse environment, contributing to the literature in this regard. The next step in our work is to integrate the solution with multimodal learning analytics to understand teaching, social and cognitive presences using the 'Community of Inquiry' as a framework. This integration aims to provide educators and learners with visualisation to guide pedagogical decision-making and refinement.

# Acknowledgements

This study was supported by the Central Reserve Allocation Committee (Ref: 03ABN), the Education University of Hong Kong.

#### References

- Diwanji, V., Reed, A., Ferchaud, A., Seibert, J., Weinbrecht, V., & Sellers, N. (2020). Don't just watch, join in: Exploring information behavior and copresence on Twitch. *Computers in Human Behavior*, *105*, 106221.
- Emerson, A., Henderson, N., Rowe, J., Min, W., Lee, S., Minogue, J., & Lester, J. (2020). Early prediction of visitor engagement in science museums with multimodal learning analytics. In *Proceedings of the 2020 International Conference on Multimodal Interaction* (pp. 107-116).
- Garrison, D. R., Anderson, T., & Archer, W. (2010). The first decade of the community of inquiry framework: A retrospective. *The internet and higher education*, *13*(1-2), 5-9.
- Knox, J., Williamson, B., & Bayne, S. (2020). Machine behaviourism: Future visions of 'learnification'and 'datafication'across humans and digital technologies. *Learning, Media and Technology*, *45*(1), 31-45.
- Pokojski, W., Panecki, T., & Słomska-Przech, K. (2021). Cartographic visualization of density: exploring the opportunities and constraints of Heat Maps. *Polish Cartographical Review*, *53*, 21-36.
- Song, Y., Cao, J., Wu, K., Yu, P. L. H., & Lee, J. C.-K. (2023). Developing 'Learningverse'-a 3D Metaverse Platform to Support Teaching, Social and Cognitive Presences. *IEEE Transactions* on Learning Technologies.
- Willermark, S., & Gellerstedt, M. (2022). Facing radical digitalization: capturing teachers' transition to virtual classrooms through ideal type experiences. *Journal of Educational Computing Research*, 60(6), 1351-1372.
- Zallio, M., & Clarkson, P. J. (2022). Designing the metaverse: A study on inclusion, diversity, equity, accessibility and safety for digital immersive environments. *Telematics and Informatics*, 75, 101909.

# Embracing Synchronicity in Distance Education: Unraveling the Paradox

# Kumiko Aoki

The Open University of Japan, Japan kaoki@ouj.ac.jp

**Abstract:** In this paper, I describe the challenges of introducing synchronicity in distance education in which the flexibility for students to learn anywhere and anytime has been most valued. With the wide diffusion of web conferencing systems such as Zoom, Microsoft Teams and Google Meet during the COVID-19 pandemic, it has become natural not only for traditional on-campus institutions, but also for distance education institutions to utilize synchronous communication tools for instructional delivery. There are advantages to utilize such tools while there are certain disadvantages especially for distance education institutions. This paper presents the results of synchronous web class trials at the Open University of Japan.

**Keywords:** synchronicity, online education, distance education, live web classes, open education

# 1. Introduction

Distance education predates the pandemic, shattering time and space constraints to offer accessible learning. As technology advances, distance education has evolved, enabling learners to engage with content, collaborate, and interact remotely. Yet, beneath this convenience lies the paradox of introducing synchronicity in this context. Synchronicity, fundamental in traditional classrooms, involves real-time educator-student and student-student interaction. Replicating this in distance education is challenging.

This paper explores synchronicity's interplay with distance education as integrating real-time elements into asynchronous learning is innovative yet challenging. Benefits of real-time interaction, like community building and immediate problem-solving, are significant. However, challenges like time constraints, connectivity, and overwhelming learners must be addressed.

This paper examines the Open University of Japan's (OUJ) case, initially adopting synchronicity due to COVID-19, and then realizing the significance of synchronicity. It also presents survey results focusing on the factors that hinder students from taking synchronous online classes. The paradox of introducing synchronicity in distance education embodies both an opportunity and a challenge. As we strive to embrace synchronous elements, we must carefully weigh the benefits against the practical constraints and design innovative solutions that harness technology's potential without sacrificing the flexibility that makes distance education so empowering. By unraveling the complexities of this paradox, educators can pave the way for a more holistic and enriched distance learning experience, poised to meet the diverse needs of learners in an ever-changing educational landscape.

# 2. The Open University of Japan

# 2.1 Pre-pandemic Era

The Open University of Japan (OUJ) has been a stand-alone distance education university and the only university in Japan (and probably in the world) with a fully operational broadcasting station with dedicated television channels. All of its broadcast programs are produced and pre-recorded at the studios at OUJ before being broadcasted and OUJ has never had a live broadcasting facility. In 2007 some of the radio broadcast programs were first made available on the Internet for the registered students, and in the following year some of the television broadcast programs were made available on the Internet for the students. Although those broadcast instructional programs were made available on the Internet for the students, only in 2016 some courses started to be offered online using the learning management system, Moodle.

Therefore, besides those face-to-face classes offered at the study centers across the country before the COVID-19 pandemic, OUJ offered no synchronicity in its instruction and all its distance education courses were offered asynchronously allowing students to take courses anywhere and anytime. For the broadcast courses, the instructional video or audio programs and text materials are all created prior to the start of the courses, and for the online courses, all the learning materials and assignments are created before the beginning of the course, which are made available to students at the beginning of the semester after a thorough checking. No real-time interaction had occurred between teachers and students or among students remotely. The synchronous interaction with teachers was possible only for those who registered for senior thesis credits in which students could seek real-time advising and instruction from their teachers.

#### 2.2 During the Pandemic

In April 2020, it was decided that all the face-to-face classes at the study centers had to be cancelled due to the pandemic. For substituting the cancelled face-to-face classes, 70 out of 92 full-time faculty members at OUJ volunteered to offer a total of 80 online synchronous classes. As for the students, 2,704 (17%) out of 15,900 students who had registered for the face-to-face classes registered for the emergency online synchronous classes. For those students who registered for the face-to-face classes, but decided not to take emergency online synchronous classes, the refunds of the fees were made. For many professors at OUJ, it was the first time to offer online synchronous classes and there was some confusion and difficulties in materializing what they planned to teach. According to the survey conducted to the professors who taught the emergency online classes at OUJ, 16 out of 69 respondents said they would not want to do it again and 26 were neutral though 44 of them said such online synchronous classes could be an important educational mode to be considered for the future of the university.

#### 2.3 From Pilot to Permanent

After the experience of offering emergency online synchronous classes described above, the need of online synchronicity at OUJ started to be recognized especially as the on-campus universities also started to offer such online synchronous classes during the pandemic. The taskforce to discuss the measures for implementing online synchronous classes permanently was set up in May 2021 in which the author was assigned as the leader.

To further identify the issues that need to be considered in providing online synchronous classes on a permanent basis, in the second semester of 2021 and the first semester of 2022, the online synchronous classes were offered as trial bases. For the trial offerings of online synchronous classes, volunteers were sought among the full-time faculty members at OUJ. Nine faculty members (six of them were the taskforce members) volunteered to offer synchronous online classes in the second semester of 2021 and 16 faculty members did in the first semester of 2022. The low volunteer rate of faculty members has raised the issue of the teaching load as teaching online synchronous classes was an additional burden, which was not part of the main job descriptions of the faculty members at OUJ.

# 3. Surveys

#### 3.1 Methods

To answer the research question surrounding the significance of online synchronicity within the educational framework of OUJ, a series of online surveys were conducted to the students who took the online synchronous classes as well as those who have never taken such classes. This paper focuses on the results of the last survey conducted to those who haven't taken online synchronous classes, providing valuable insights into the perceptions and preferences of students who had not participated in online synchronous classes, and shedding light on the factors influencing their decisions not to take such classes.

The methodology involved the careful selection of participants from the university's registered undergraduate students, with a total population of 76,251. By excluding 639 students who had already taken online synchronous classes, a representative sample of 7,500 students was randomly chosen from the remaining 75,642. This approach ensured a diverse range of perspectives, enhancing the generalizability of the findings. With a response rate of 5.6%, as reflected by the 422 completed survey questionnaires, the study achieved a reasonable level of engagement from the selected participants. The responses revealed insightful patterns concerning the factors inhibiting students from enrolling in online synchronous classes.

# 3.2 Factors for Not Taking Online Synchronous Classes

Notably, the primary reasons for abstaining from these classes were consistent. A significant percentage cited the inconvenience of class timings (36.7%), underscoring the challenge of accommodating diverse schedules in synchronous learning environments. Another substantial group expressed a lack of interest in available class topics (36.7%), highlighting the need to align course offerings with the varied preferences of learners. Moreover, apprehensions about adapting to this novel style of learning were prominent, with 35.3% of respondents feeling unsure about the new method.

However, the future outlook for online synchronous classes remains optimistic, as 46% of respondents expressed intent to explore these classes, while 37.4% remained undecided. The reluctance to commit to a specific timeframe emerged as a predominant deterrent (54.8%), indicating the value placed on the flexibility inherent in asynchronous learning. Additionally, a considerable portion (35.1%) remained uncertain about the instructional approach, emphasizing the need for clear communication and guidance regarding this educational modality.

# 4. Conclusion

In closing, this research has unveiled valuable insights into the paradoxical nature of introducing synchronicity in distance education. The study's findings underscore the importance of addressing scheduling flexibility, diversifying course offerings, and transparently communicating the benefits of synchronous learning to enhance student engagement and participation. As distance education continues to evolve, bridging the gap between synchronous and asynchronous modalities presents both challenges and opportunities that demand innovative strategies to craft a truly enriched learning experience.

# A systematic review on the competences of human-AI collaboration

Youngjin YOO, Young Hoan CHO<sup>\*</sup> & Jeewon CHOI Seoul National University. South Korea

\*yhcho95@snu.ac.kr

**Abstract:** Research on AI literacy has been conducted in the context of the essential skills needed for humans in a future society coexisting with AI. From a post-humanism perspective, existing AI literacy has been insufficient in presenting human competences for communication and collaboration with AI. The purpose of this study is to systematically derive the areas and competences essential for human-AI collaboration through a literature review. The related documents were collected from online academic databases, and a total of 32 papers were selected based on inclusion-exclusion criteria and quality assessment of the studies. Upon reviewing the selected literature, four main areas of human-AI collaboration competences were identified: 1) Understanding and utilizing AI, 2) Communication with AI, 3) Collaborative task regulation, and 4) AI ethics and socio-cultural values. This study offers theoretical implications for the development of AI curriculums, educational methods, and assessments to enhance human-AI collaboration competences.

Keywords: Human-AI Collaboration, Competence, Systematic review

# 1. Introduction

Globally, various efforts led by governments and international organizations are underway to advance AI-based education. In Korea, AI-based Intelligent Tutoring Systems (ITS) have been created and spread for individualized learning, guided by the Ministry of Education and various educational offices. This aligns with the global trend as AI becomes significant in education. Al literacy is considered an essential skill for humans to enhance. Previous studies on AI literacy have variation on the definition of AI literacy including know and understand AI, use AI, evaluate, and create AI (Ng et al., 2021). These studies commonly emphasized coding education, viewing the crucial skill for humans to understand and use AI (Kim et al., 2020; Ng et al., 2021). This perspective on AI literacy recognizes AI as a media or tool that humans handle. From a post-humanism perspective, AI is viewed as an actor interacting with humans in an equal relationship. Humans learn and thrive through communication and collaboration with AI (Kim et al., 2022; Park, 2020). In the Human-Computer Interaction (HCI) domain, research is also being conducted to develop AI systems that collaborate well with humans, reflecting a shift in understanding AI not merely as a tool but as a collaborative partner. If humans perceive AI as a collaborative partner, the competences required for humans would differ from the instrumental view of AI. Human-AI collaboration competences should be systematically enhanced through AI-related learning experience in AI education. In this regard, it is necessary to investigate the human competences for collaboration with AI. Therefore, this study aims to systematically review the literature on human-AI collaboration to identify the main areas and competences and delve into the educational implications for enhancing these competences.

# 2. Theoretical Background

The previous studies on human-AI collaboration competence can be broadly categorized into studies on AI literacy and AI collaboration. Firstly, studies on AI literacy emphasized cognitive competences related to understanding and utilizing AI. Common components

identified in AI literacy include understanding AI, utilizing AI, evaluating, and creating AI, and AI ethics. These studies highlighted the cognitive area of acquiring knowledge about AI and its application (Ng et al., 2021). Secondly, studies on AI collaboration include both cognitive and social competences in the collaboration process. For effective collaboration with AI, competences such as understanding AI, communication, positive attitudes towards AI, task regulation and reflection are essential. These competences included social competences that encompass interactions required in the collaboration with AI. In this study, human-AI collaboration competence is defined as the ability in an interactive process where multiple humans and AI systems reciprocally engage in joint activities aimed at achieving shared goals (Cañas, 2022; Markauskaite et al., 2022; Song & Cho, 2023; Terveen, 1995). While various competences related to human-AI collaboration were presented, comprehensive definitions and descriptions including both cognitive and social competences were not provided. It is necessary to explore the specific areas and sub-competences of human-AI collaboration competence to enhance these competences through AI education.

# 3. Method

Following Newman and Gough's (2020) framework for qualitative systematic literature analysis, literature was selected from online databases using keywords such as 'human-AI collaboration', 'competence', and 'AI literacy'. A total of 32 articles were selected based on inclusion-exclusion criteria and quality assessment of the studies. The researchers conducted coding and thematic analysis, constructing an initial outline with overarching themes. This outline was continuously revised throughout the analysis process. Through discussions among researchers, main areas and competence related to human-AI collaboration were derived, along with their respective definitions and meanings.

# 4. Findings

Human-AI collaboration competence derived from the systematic literature review contains 4 main areas and 11competences.





# 4.1 Understanding and Utilizing AI

Understanding and Utilizing AI involves understanding foundational knowledge about AI, continuously utilizing AI tools, autonomously learning and evaluating these tools, and solving everyday challenges. Understanding AI is about comprehending AI technologies, their operational principles, and actively learning new AI tools. Utilizing

Al Tools is the process of selecting the right Al tools for specific tasks and actively employing them to solve problems. Evaluating Al Tools is the critical assessment of the performance and limitations of these tools, understanding their suitability for certain tasks, and recognizing their beneficial impact on daily life.

#### 4.2 Communication with AI

Communication with AI involves conveying intent in a suitable manner for AI, critically interpreting and evaluating AI output, and adjusting trust towards AI. Expressing Intent is the ability to clearly express one's intentions to AI by inputting natural language or manipulating information in a proper manner for AI. Critical Thinking involves actively interpreting and evaluating the validity and potential inaccuracies of AI outputs. Trust is about adjusting between relying on AI and human judgment, recognizing the importance of human-AI collaboration, and treating AI as a valuable partner in the process.

#### 4.3 Collaborative Task Regulation

Collaborative task regulation is about developing plans for task resolution, regularly monitoring, and adjusting the collaboration process, and co-performing. Planning is about allocating roles and responsibilities between humans and AI and creating plans for task resolution. Monitoring and Adjusting is the ability to proactively observe collaboration with AI, monitor the actions of both humans and AI for shared goals, and adjust plans and strategies accordingly. Co-performance is the ability to integrate AI outputs with one's own knowledge and collaboratively explore, develop, and execute optimal solutions with AI.

#### 4.4 AI Ethics and Sociocultural Values

Al ethics and sociocultural values is to recognize and practice the necessary ethics and sociocultural values required in the process of collaborating with Al. Ethical Awareness is the ability to recognize ethical issues related to Al, take responsibility for the outcomes, and adhere to information security and ethics. Practicing Sociocultural Values is the ability to recognize diverse human-centered sociocultural values, evaluate the social impact of Al, and put those values into practice.

#### 5. Conclusion

Human-AI collaboration competences are a fundamental ability for humans to live in the era of coexistence with AI. These competences not only enhance human learning and development by promoting effective collaboration with AI, leading to benefits such as improved problem-solving, increased efficiency, and better achievements in co-performing tasks but also foster a coexistence by trusting AI as significant actors in the evolving society. The findings of this study implied that Human-AI collaboration competences should be improved through educational supports such as AI-integrated curriculums, activity-based instruction with AI, and assessment of these competences with data from AI-related activity. It is necessary that the empirical research in educational setting should be conducted to implement these educational supports.

#### References

Cañas, J. (2022). Al and ethics when human beings collaborate with Al agents. *Frontiers in Psychology, 13*, 1-9.

- Kim, J., Lee, H., & Cho, Y. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies*, 27(5), 6069-6104.
- Kim, S., Kim, S., Lee, M., & Kim, H. (2020). A study on artificial intelligence education for K-12 students and teachers. *Journal of Computer Education*, 23(4), 1-11.

Newman, M., & Gough, D. (2020). Systematic reviews in educational research: Methodology, Perspectives and Application. Springer VS, Wiesbaden.

- Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martinez-Maldonado, R., Howard, S., ... & Siemens, G. (2022). Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI. *Computers and Education: Artificial Intelligence*, *3*, 100056.
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041.
- Park, H. (2020). Posthuman literacy Its concept, categories, theoretical basis, and the directions of education. *The Korean Journal of Literacy Research*, *11*(1), 11-55
- Song, H., Cho, Y. (2023). A Developmental study on design principles of activity-based instruction for improving human-AI collaboration competency, *The Journal of Educational Information and Media.* 29(1), 145-173.
- Terveen, L. (1995). Overview of human-computer collaboration. *Knowledge-based Systems, 8*(2), 67-81.

# Influence of Robot Roles on Self-Review

#### Shunsuke SADA\* & Akihiro KASHIHARA

The University of Electro Communications, Japan 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585 \*s2330055@gl.cc.uec.ac.jp

**Abstract:** In recent years, a considerable number of studies have been made on learning with social robots as partner for learners. In these studies, robots are often assigned their roles in advance such as tutor, peer, or learners themselves. On the other hand, the roles assigned to the robots could give learners a cognitive bias to their own thoughts, emotions, attitudes, interaction ways, etc., related to learning. However, such role influence has not been sufficiently explored so far. In this work, we investigate this in the context where learners self-review their own presentation with a social robot. We have conducted a case study with the presentation robot we developed in our previous work. The robot reproduced learners' presentation with two roles that were their peer and themselves. The results suggest that the role of robot as peer brings about more effective self-review in regard to reducing uncomfortableness in self-review and promoting awareness of points to be improved in presentation than the role as learners themselves.

Keywords: social robot, cognitive bias, robot role, self-review, presentation behavior

#### 1. Introduction

In related work on social robotics for learning, robots ordinarily play their own roles such as tutor, peer, and learners themselves, which are assigned in advance for facilitating communication with learners. In communication between human, the role one (communicator) plays often gives the other (communicatee) a cognitive bias towards his/her own thoughts, emotions, attitudes, interaction ways, etc. When a communicator explains something, for example, a communicate tends to accept the contents. In case a communicator explains the same contents as peer, on the other hand, the communicate tends to listen to the contents more carefully with some doubt/questions. In this way, the communicator (Rod D. Roscoe and Michelene T. H. Chi, 2008).

Since robots as learning media possess physical embodiment and anthropomorphic tendency compared to other media such as PC, tablets, etc., they could look like human (S. Kiesler, A. Powers, S. R. Fussell, and C. Torrey, 2008). Even in human (learner)-robot communication, such cognitive bias would be accordingly expected to arise depending on the roles of robot as their learning partner. Understanding the influence of robot roles also makes it possible to design more instructive robots. However, it has not been sufficiently explored so far.

In this paper, we investigate whether robot roles have an influence on learners' thoughts and engagement in the context where learners self-review their own presentation behavior by means of a social robot reproducing it.

#### 2. Self-Review with Presentation Robot

Self-review of presentation involves learners reviewing their presentation by themselves to identify points to be improved. Learners must consider not only what to present but also how to present with non-verbal behavior (presentation behavior), which

includes gesture, gaze, and paralanguage. Presentation behavior is particularly important for communicating the presentation contents.

The general way for learners to self-review is to record a video of their presentation and then to check it out. However, it would be too uncomfortable for them due to their appearance and voice to conduct an in-depth review. In addition, novices have insufficient knowledge about what to review. Towards resolving these problems, we implemented a presentation robot, which stands in for learners and reproduces their presentation including presentation behavior (K. Inazawa, A. Kashihara, 2022).

Figure 1 shows a presentation made by the presentation robot. The robot reproduces learners' presentation as the role of learners themselves. We also designed a presentation behavior model to prepare a checklist for review, which allows learners to become aware of points to be improved. Refer to (K. Inazawa, A. Kashihara, 2022) for the model and checklist in detail.

The results of a case study we conducted suggest that the robot could reduce learners' uncomfortableness to promote their engagement on self-review, and also bring about more awareness of points to be improved in their presentation behavior (K. Inazawa, A. Kashihara, 2022).

Although the presentation robot plays a role of learners themselves in our previous work, we expect it to bring about more effective self-review when it plays a role as peer of learners. That is the main topic of this paper.

# 3. Influence of Presentation Robot Roles

We have conducted a case study whose purpose was to investigate the influence of presentation robot roles on self-review. We also used three criteria: engagement in reviewing, uncomfortable feeling, and awareness of points to be improved.

The hypotheses we set up in this study were as follows:

- H1: The robot as peer promotes engagement on reviewing compared with the robot as learners themselves.
- H2: The robot as peer reduces uncomfortable feeling compared with the robot as learners themselves.
- H3: The robot as peer enhances awareness of points to be improved compared with the robot as learners themselves.



**Figure 1: Presentation Robot** 



**Figure 2: Experimental Procedure** 

The participants were 14 undergraduate and graduate students in informatics and engineering. We set two conditions: self-review with the robot as participant himself/herself (control condition), and self-review with the robot as peer of participant (experimental condition). Figure 2 shows the procedure of this study, in which the participants were divided into two groups (Group A and B). Each participant reviewed his/her own presentation twice as within-subject design. The robot reproduced his/her presentation in both reviews, but the robot roles, the color, and the voice pitch were different in each review. As for the robot role, we particularly informed him/her in advance that the robot demonstrates peer's presentation in the control condition. In the experimental condition, on the other hand, he/she was informed that the robot demonstrates your own presentation.

The post-questionnaire consisted of 10 5-point Likert scale questions, and the overall-questionnaire consisted of 8 questions that asked which condition allowed the participants to engage/reduce uncomfortableness in reviewing. During self-review, the participants were also required to write down points to be improved for their presentation behavior with the checklist.

# 4. Results and Discussion



As for the results of the questionnaires about engagement, there were no significant

Figure 3: Average Numbers of Points to be Improved



Figure 4: Average Numbers of Points to be Improved in Group A and B

differences between the two conditions, and H1 was not supported. However, the postquestionnaire results showed high average scares in both conditions. This suggests that the use of the robot increases engagement on self-review regardless of the robot roles.

As for uncomfortable feeling, there were significant differences between the two conditions, which suggest the robot as peer reduces uncomfortable feeling on self-review compared to the robot as learner himself/herself. This supports H2.

As for awareness of point to be improved about presentation behavior, Figure 3 shows the average number of improved points found out in each condition. There was no significant difference between the conditions, and H3 was not supported. We will take this into consideration below.

Since the participants conducted two reviews of their presentation, the second review seems to be easier for them to become aware of points to be improved than the first review. This order effect is confirmed in Figure 4 (a), which shows the average number of improved points in Group A. As shown in Figure 4 (b) for Group B, on the other hand, the second review under the control condition brought about a fewer number of improved points than the first review under the experimental condition. This suggests the possibility that the role as peer counteracts the order effect to bring about more awareness of improved points even in the first review.

#### 5. Conclusion

In this paper, we investigated the influence of robot roles in the context where learners self-review their own presentation with a social robot. The results suggest that the robot roles have an influence on uncomfortableness and cognitive awareness in reviewing presentation.

In future, we need to ascertain the influence of robot roles in other learning contexts.

#### Acknowledgements

This work was supported in part by JSPS KAKENHI Grant Numbers 20H04294 and 23H03505.

#### References

Rod D. Roscoe & Michelene T. H. Chi: Tutor learning: the role of explaining and responding to questions, Instructional Science volume 36, p.321-350 (2008)

- S. Kiesler, A. Powers, S. R. Fussell, and C. Torrey: Anthropomorphic Interactions with a Robot and Robot-like Agent, Social Cognition, Vol.26, pp.169-181 (2008)
- K. Inazawa, A. Kashihara: Designing and Evaluating Presentation Avatar for Promoting Self-Review, IEICE Transactions on Information and Systems, Vol.E105-D, No.9, pp.1546-1556 (2022)

# Metaverse and Education: A Bibliometric Analysis Based on the Past Twenty Years

Chien-Liang Lin<sup>a</sup>, Zeren Zhu<sup>b\*</sup>, Yushun Su<sup>b</sup>, Juan Zhou<sup>c</sup> & Yu-Sheng Su<sup>d</sup>

<sup>a</sup>Department of Information Management, Ming Chuan University, Taiwan <sup>b</sup>Musashi Urawa Japanese Language Institute, Saitama-ken, Japan <sup>c</sup>Tokyo Institute of Technology, Japan <sup>d</sup>Computer Science and Information Engineering, National Chung Cheng University \*ccucssu@cs.ccu.edu.tw

**Abstract:** This study explores the integration of metaverse-associated technologies such as VR, AR, MR in education, employing bibliometric techniques and VOSviewer software to analyze the field. Drawing from the Science Citation Index (SCI) and Social Science Citation Index (SSCI), 937 papers were preliminarily obtained to identify prevailing themes and emerging trends. Five main research clusters were identified: challenges and attitudes in implementing technology, virtual reality in skill training, college students' self-efficacy and perceptions, interactive educational settings for children, and specific disciplinary knowledge applications. The study contributes to the understanding of the practical application of metaverse technologies in education and may guide future research and development in this rapidly evolving field.

Keywords: metaverse, education, bibliographic

# 1. Introduction

In 1992, acclaimed American author Neal Stephenson introduced the idea of the 'metaverse' in his cyberpunk science fiction novel 'Snow Crash'. In this vision, he portrayed a 3D space that mirrors the real world where individuals could interact through digital avatars. With the advent of technologies like Virtual Reality (VR), Augmented Reality (AR), Artificial Intelligence (AI), and blockchain, coupled with societal events such as the COVID-19 pandemic which pushed for a transition to virtual interactions, our digital footprint has deepened (Zhang et al., 2022). The metaverse's boundaries are seemingly limited only by our imagination, making it an alluring concept to many (Hwang & Chien, 2022). However, the emerging metaverse is not without its challenges, particularly regarding user privacy (AI-Ghaili et al., 2022). As pointed out by Mystakidis (2022), the battle among dominant corporations to shape this realm will be pivotal in determining future privacy standards for users.

Different interest groups related to the metaverse shape its portrayal based on their distinct perspectives, leading to a varied representation of the metaverse and its associated technologies (Dolata & Schwabe, 2023). Despite its years of evolution, the metaverse remains an elusive concept without a universally accepted definition. Broadly speaking, its development can be categorized into three progressive stages: (i) digital twins, (ii) digital natives, and (iii) surreality. This progression begins with replicating real-world entities, evolves into innovations grounded in reality, and culminates in an independent realm separate from our reality (Wang et al., 2022). Mystakidis (2022) posits that the metaverse relies on technologies that support multisensory interactions within virtual spaces, among digital entities and people. Extended or Cross Reality (XR)—comprising virtual reality (VR), augmented reality (AR), and mixed reality (MR)—is seen as a vehicle meeting the metaverse's interaction demands. Contrasting traditional VR or AR, Hwang and Chien (2022) emphasize three unique traits of virtual universes: they are 'shared,' 'persistent,' and 'decentralized.' Barrera and Shah (2023), synthesizing literature and industry insights, define the contemporary metaverse as "a technology-driven nexus of scalable and potentially interoperable extended reality environments. These spaces bridge physical and virtual realms, offering experiences marked by their immersion level, environmental authenticity, and social engagement." This characterization distinguishes the metaverse from earlier, simpler virtual worlds. Dolata and Schwabe (2023) attribute the current phase of the metaverse to the rise of VR, AR, NFTs, AI, and more, dubbing it the third wave. They observe that today's metaverse carries hallmarks of formal acknowledgment and platform-independent entitlements related to ownership, authorship, responsibility, and autonomy. These traits pave the way for a socioeconomic structure deeply linked to the global economy. In essence, XR technologies stand central to metaverse discourse, acting as a bridge between avatars in the virtual domain and users in the tangible world (Park & Kim, 2022).

Technical support in teaching plays a significant role in education and development (Weisberg, 2011). However, the application of VR, AR, MR, and other simulation devices related to the metaverse in actual teaching lacks comprehensive exploration. This study employs bibliometric methods and analyzes the literature in this field using VOSviewer software to understand the hot topics in research related to the application of metaverse-related simulation technologies in education.

# 2. Data Set and Research Methodology

All the articles in this study were retrieved from the citation databases of the Science Citation Index (SCI) and Social Science Citation Index (SSCI) and were obtained from the Web of Science (WOS) created by the Institute of Science Information (ISI). In selecting keywords, this study focused primarily on educational journals. Therefore, the following retrieval methods were used for the keywords: TS=("metaverse" OR ("virtual reality" OR "augmented reality" OR "mixed reality" OR "extended reality") AND "education"). Some articles that had no relevance to VR, AR, or the metaverse were excluded, as were review articles. Finally, 937 papers were initially obtained.

#### 3. Results

Based on the network map, it is evident that research in this field primarily forms five clusters. The red section mainly discusses challenges and attitudes toward the application of technology in education. Penn and Ramnarain (2019) used virtual environment simulations to provide visualizations of chemical concepts, improving students' understanding of chemistry concepts and their attitudes toward chemistry learning. Fidan and Tuncel (2019) positively affected learning in physics subjects through Augmented Reality (AR)-assisted Problem-Based Learning (PBL) while enhancing the perceptual abilities of students with disabilities. The yellow section primarily discusses topics related to virtual reality hands-on simulation skill training. Çakiroğlu and Gökoğlu (2019) employed a Virtual Reality Behavior Skills Training (VR-BST) approach to teach basic firefighting skills, incorporating behavioral modeling and on-site training into the learning environment to improve students' behavioral skills. The blue section mainly discusses self-efficacy, impressions, and overcoming issues in higher education, such as Cai et al. (2019), who discuss how AR helps students understand abstract mathematical concepts and enables high self-perceived performance students to apply more advanced strategies when learning mathematics. The green section discusses interactive children's educational environments and framework design topics. Degli Innocenti et al. (2019) developed new VR software systems that help patients better understand medical knowledge and treatment information through VR simulations. The purple section discusses disciplinary knowledge and applications in subject-specific contexts.



Figure 1. Keyword Map of Terms Occurring 20 Times or More

#### 4. Discussion

Based on the data results, it can be observed that research related to VR places a stronger emphasis on simulating learning through practical skills, while AR stands out more in mobile learning. When designing experiments using VR and AR devices in educational research, most studies focus on individual learning, with limited research discussing the simultaneous use of devices by multiple users or the impact of cooperative learning and multi-user device usage (Back, Tinga, & Louwerse, 2021). Additionally, despite the close relationship between the metaverse and VR/AR technologies, few studies directly link the two, resulting in a lack of research that combines the latest technologies like AI and blockchain with VR/AR (Chatterjee et al., 2021; Abdinejad et al., 2021). This study provides the first analysis and categorization of research on the use of metaverse-based interactive media and XR technologies in education. At its current stage, research in this field is still relatively nascent, primarily focusing on building educational platforms based on VR and AR devices. While some studies have explored the effects of cooperative learning and group learning, discussions regarding the integration of AI, blockchain, and other technologies are lacking.

#### References

- Abdinejad, M., Ferrag, C., Qorbani, H. S., & Dalili, S. (2021). Developing a simple and costeffective markerless augmented reality tool for chemistry education.
- Al-Ghaili, A. M., Kasim, H., Al-Hada, N. M., Hassan, Z., Othman, M., Hussain, T. J., ... & Shayea, I. (2022). A review of metaverse's definitions, architecture, applications, challenges, issues, solutions, and future trends. IEEE Access.
- Barrera, K. G., & Shah, D. (2023). Marketing in the Metaverse: Conceptual understanding, framework, and research agenda. Journal of Business Research, 155, 113420.
- Cai, S., Liu, E., Yang, Y., & Liang, J. C. (2019). Tablet-based AR technology: Impacts on students' conceptions and approaches to learning mathematics according to their selfefficacy. British Journal of Educational Technology, 50(1), 248-263.
- Çakiroğlu, Ü., & Gökoğlu, S. (2019). Development of fire safety behavioral skills via virtual reality. Computers & Education, 133, 56-68.
- Chatterjee, S., Moon, S., Rowlands, A., Chin, F., Seeberger, P. H., Merbouh, N., & Gilmore, K. (2021). Click, Zoom, Explore: Interactive 3D (i-3D) Figures in Standard Teaching Materials (PDFs). Journal of Chemical Education, 98(11), 3470-3475.
- de Back, T. T., Tinga, A. M., & Louwerse, M. M. (2021). CAVE-based immersive learning in undergraduate courses: examining the effect of group size and time of application. International Journal of Educational Technology in Higher Education, 18, 1-18.
- Degli Innocenti, E., Geronazzo, M., Vescovi, D., Nordahl, R., Serafin, S., Ludovico, L. A., & Avanzini, F. (2019). Mobile virtual reality for musical genre learning in primary education. Computers & Education, 139, 102-117.
- Dolata, M., & Schwabe, G. (2023). What is the Metaverse and who seeks to define it? Mapping the site of social construction. Journal of Information Technology, 02683962231159927.

- Fidan, M., & Tuncel, M. (2019). Integrating augmented reality into problem based learning: The effects on learning achievement and attitude in physics education. Computers & Education, 142, 103635.
- Hwang, G. J., & Chien, S. Y. (2022). Definition, roles, and potential research issues of the metaverse in education: An artificial intelligence perspective. Computers and Education: Artificial Intelligence, 3, 100082.

Mystakidis, S. (2022). Metaverse. Encyclopedia, 2(1), 486-497.

- Park, S. M., & Kim, Y. G. (2022). A metaverse: Taxonomy, components, applications, and open challenges. IEEE access, 10, 4209-4251.
- Penn, M., & Ramnarain, U. (2019). South African university students' attitudes towards chemistry learning in a virtually simulated learning environment. Chemistry education research and practice, 20(4), 699-709.
- Wang, Y., Su, Z., Zhang, N., Xing, R., Liu, D., Luan, T. H., & Shen, X. (2022). A survey on metaverse: Fundamentals, security, and privacy. IEEE Communications Surveys & Tutorials.
- Weisberg, M. (2011). Student attitudes and behaviors towards digital textbooks. Publishing research quarterly, 27(2), 188-196.
- Zhang, X., Chen, Y., Hu, L., & Wang, Y. (2022). The metaverse in education: Definition, framework, features, potential applications, challenges, and future research topics. Frontiers in Psychology, 13, 6063.

# An Interactive Robot Lecture System for Attention and Understanding Recovery

#### Toshiyuki SHIMAZAKI<sup>a\*</sup>, Akihiro KASHIHARA<sup>a</sup>

<sup>a</sup>Graduate School of Informatics and Engineering, The University of Electro-Communications, Japan \*shimazaki@uec.ac.jp

Abstract: In lecturing, lecturers need to control the attention of learners to maintain while monitoring their learning states. We have claimed that a robot as lecturer can properly conduct nonverbal lecture behavior to control learners' attention and enhance their engagement compared with human lecturers. Although attention control is possible in shorter lectures, it would be more difficult for the robot to maintain learners' attention in a longer lecture. This prevents them from becoming aware of an important part of the lecture contents to understand. In this work, we have been developing an interactive robot lecture system that interacts with learners to recover their attention and understanding when they overlook the lecture contents. To recover learners' attention, the system first estimates their states of attention and understanding from their posture and gaze during lectures. Second, the system reconstructs a lecture scenario to generate an interactive lecture behavior, which combines non-verbal behavior about pause, walk, and repeat with paralanguage. The interactive behavior is implemented using a humanoid robot NAO. Third, the system provides questions about the slide contents for recovering their understanding when they lose their attention in spite of attention recovery.

Keywords: Interactive Robot Lecture, Attention Control, Understanding Recovery

#### 1. Introduction

In e-learning and face-to-face lectures, lecturers need to control the attention of learners to attract their interests and to preserve their engagement. This also requires lecturers to properly use their nonverbal behavior involving gaze, gesture, and paralanguage. In longer lectures, in particular, lecturers should interact with learners using nonverbal behavior according to their learning states. Such interactive behavior allows lecturers to maintain learners' attention and engagement.

However, it is not easy even for experienced human lecturers to make appropriate use of nonverbal actions during the lecture or to keep track of learners' states. It is also difficult for learners to become aware of what they must pay attention to, thereby making them overlook an important part of the lecture contents. They often fail to follow the lecture and to understand the contents.

Towards resolving these problems, we have focused on nonverbal behavior for e-Learning lecture, and developed a robot lecture system in which the robot reconstructs nonverbal behavior of human lecturer, and presents the lecture contents with the reconstructed behavior (Ishino et al., 2022). The results of the case study with the system suggest that the robot lecture is more effective in attracting learners' attention and giving an understanding of the lecture contents rather than the video lecture and the lecture reproduction with the robot.

The robot lecture, on the other hand, could not recover learners' attention and understanding when they lose their attention and understanding (Kashihara et al., 2019 & Shimazaki et al., 2022). In this paper, we propose an interactive robot lecture system that recovers learners' attention and understanding during lecture.

# 2. Framework of Interactive Robot Lecture

Let us first explain the framework of the interactive robot lecture system we have been developing. In this work, we use a humanoid robot, NAO manufactured by SoftBank Robotics. NAO can walk, and perform nonverbal behavior such as gaze, pointing gesture, etc., and paralanguage such as voice volume, speed, intension, pause, etc., which are similar to those of human lecturers.

Figure 1 shows the framework of the system. The system conducts interactive robot lectures using three sub systems:

- Recoding system: Recording of lecture made by human lecturer,

- Sensing system: Learning states estimation, and
- Interaction system: Lecture scenario reconstruction and attention/understanding recovery.

A lecture scenario is defined as the contents of each slide, associated nonverbal behavior, verbal explanation, and the slide sequence that the lecturer considers when designing a lecture. The slide sequence is divided into several segments, each of which also consists of several slides, and represents a section/sub-section in the lecture contents. Each segment includes in-slide and between-slides questions used for recovering understanding of the corresponding slide contents.



Figure 1. Framework of Interactive Robot Lecture.



(a) Robot Walking for Attention Recovery.



(b)Understanding Recovery with an In-slide Question.

Figure 2. Interactive Lecture for Attention and Understanding Recovery.

The recoding system records a human lecturer's lecture with Kinect to diagnose and reconstruct improper nonverbal behavior by means of a lecture behavior model we designed (Ishino et al., 2022). It then generates a lecture scenario. The segments of the lecture sequence and in-slide/between-slide questions are prepared in advance by the human lecturer.

The sensing system estimates learner's states of attention to and understanding of each slide in the lecture sequence from their posture and gaze using Kinect and tobii pro nano.

The interaction system reconstructs the lecture scenario according to the estimated states to generate interactive lecture behavior for recovering their attention. In case the attention recovery fires more than once in one segment (in other words, their attention could not be recovered), in addition, the system provides them with in-slide/between-slides questions for recovering their understanding of the corresponding slide or the segment. The scenario reconstruction is done with the scenario control model and procedures for attention and understanding recovery (Shimazaki et al., 2023).

# 3. Sample Interaction

Figure 2 (a) demonstrates the robot walking to a learner. This shows an example of the attention recovery interaction. When the system estimates a learner looks away not to concentrate on a slide, in this example, NAO interrupts the lecture, and then walks towards him/her to ask if he/she would like it to repeat explanation of the slide contents. If he/she says yes, NAO explains the slide contents once again. Such interaction allows the robot to recover his/her attention.

Figure 2 (b) demonstrates the robot providing a learner with an in-slide question for recovering an understanding of a slide. This shows an example of the understanding recovery interaction. The understanding recovery fires when the attention recovery is consecutively failed in one segment of the lecture sequence. The number of times of the consecutive fails is defined by a human lecturer. If the number is defined as 2, NAO will fire the understanding recovery interaction after the attention recovery consecutively failed twice. In the understanding recovery interaction, NAO provides the learner with an in-slide or between-slides questions. These questions are 4-choice questions asking the contents in or the relationships between the slides where attention recovery has failed. If he/she answers all the questions correctly, NAO assumes that he/she understands the lecture contents, and restarts the lecture. If he/she makes a mistake, NAO explains the corresponding slide once again. Such interaction allows the robot to recover their understanding of the lecture.

# 4. Conclusion

In this work, we have been developing an interactive robot lecture system for learner's attention and understanding recovery. A preliminary case study has been conducted with high school participants to evaluate the accuracy of sensing data for learning state estimation and the interactive lecture with NAO. The results suggest that the learning state estimation using posture and gaze sensing data is similar to that of a human lecturer, and that the interaction with NAO contributes to maintaining learner's attention and recovering their understanding of the lecture contents.

In future, we will conduct a more detailed evaluation of the interactive robot lecture system with university students.

# Acknowledgements

This work was supported in part by JSPS KAKENHI Grant Numbers 20H04294 and 23H03505.

# References

- Ishino, T., Goto, M., Kashihara, A. (2022). Robot Lecture for Enhancing Presentation in Lecture. Research and Practice in Technology Enhanced Learning (RPTEL) 17.1, 1-26.
- Kashihara, A., Ishino, T., Goto, M. Robot (2019). Robot Lecture for Enhancing Non-Verbal Behavior in Lecture.Proc. of the 20th International Conference on Artificial Intelligence, 128-132. Chicago.
- Shimazaki, T., Sugawara, A., Goto, M., Kashihara, A. (2022). An Interactive Robot Lecture System Embedding Lecture Behavior Model. The 24th International Conference on Human-Computer Interaction, 224–236.
- Shimazaki, T., Kashihara, A. (2023). Evaluating Case Studies and Designing Interaction Model for Interactive Robot Lecture. Japanese Society for Information and Systems in Education Research Report, Vol.38, No.2, 65-72. in Japanese.

# Posthumanizing Creativity and Material Histories

#### Alekh V<sup>\*</sup>, Chandan DASGUPTA

Indian Institute of Technology, Bombay, India \*alekhv@iitb.ac.in

**Abstract:** With this article, we attempt to shift the conceptualization of creativity from a human centered perspective to a wider perspective of posthumanism. We use the scenario of collaborative making context, and illustrate how material history evolved rooting on the posthumanist standpoint. We see that the humans and nonhuman entities are enmeshed in creative acts where agency is spread across the various actors leading to learning opportunities and innovative outcomes.

Keywords: Creativity, Constructionism, Posthumanism, Materiality

#### 1. Introduction

Posthumanism looks at decentralizing humans from the sole source and center of actions and unsettling the concepts related to material neutrality (Barad, 2003, 2007; Pickering, 1993). It can bring the vitality of the matter, unpredictability, unfolding ontology, and signifies material turns as humans encounter non-human matter (Barad, 2003; Haraway, 1985). The approach shifts the attention from who to how all the agents are entangled in the creative phenomemon. Here, agency is neither rested solely on human nor nonhuman elements but through becoming of emergent manifestations, effects through certain configurations of situated entities (Suchman, 1987), and unfolds in practice (Pickering, 1993). To bring in meaning to this standpoint, we consider the practice of "making", which is a materially grounded chain of actions where makers create personally meaningful and creative outcomes. Making is featured by iterative nature, encountering failures, gathering feedback from fellow makers and experts resulting in unexpected outcomes. While engaging in design which is an inherent part of making, makers engages with humans and non-human entities like material. Collaborative making involves the coming together of makers, sharing resources in the situated environment, influencing each other's ideas and design processes. Creativity has been conceived as an individual entity, with less connections to social networks or material, technological association (Glaveanu, 2014). This ideology neglects the influence of other individuals, tools, technologies and other non-human entities while analyzing the creative processes.

Based on these prior literature, we look to posthumanize creative actions emerging out of collaborative making where novice makers engage in design problem solving and understand how material histories are generated as material qualities evolve. We position the creative aspects of the making process as emergent (Sawyer & Dezutter, 2009; Tangaard, 2013).

# 2. Methods

For addressing the research goal of the article we look into a maker activity-centered workshop which was part of a maker workshop series organized at a leading engineering institute in India. The participants of were eight second-year mechanical engineering undergraduate students who completed a maker workshop based on Lego Mindstorms EV3 robotic kit and responded to an open invitation. A total of four teams were formed in the order of their response, with each team consisting of two members. After the introductory session on digital
fabrication, participants were given a design challenge: *Conceptualize an assembly line/production line that is semi-automated with static and dynamic robots. Model and build the setup with resources available in the makerspace.* The making sessions lasted for a total We followed the making activities of teams- Team P which consisted of one female student (G1) and one male student (B1). A facilitator was present during the making sessions to support the teams.

We conducted the analysis with the data sources as video, photographs, design artifacts, field notes, and observations during making. We followed theories of new materialism and adopted the methodological process of thinking with theory (Jackson & Mazzei, 2012) to illuminate the creative entanglements. We viewed the assemblage of data sources simultaneously and iteratively. Here, we see ourselves entangled and becoming with the unfurling of research course. The research entanglement involved the encounter with data sources as video, photographs, design artifacts, field notes, and observations during making.

## 3. Findings

We find that the makers, materials and the facilitator are constitutively entangled in the making sessions leading to creative outcomes.

The team started with the idea of parallel robot, with the rope and pulley mechanism for functioning the fork. Here the makers use a variety of materials including, Lego motors, leg kit parts, chopsticks, earphone wires, play-doh, and tapes. The following excerpt shows how makers and materials come together and open up spots for creative actions.

Maker B1:	Do we have strings? can we use those filaments as strings? [referring to 3D printing filaments]
Facilitator:	Yeah you can try. There are some rope pieces in that box [pointing to the box with discarded items]
Maker G1:	[Checks the box] there are some wires too
Maker B1:	Aah we can use it but might slide over.
Maker G1: Maker B1: Facilitator:	Or maybe we can cut the ropes. This earphone can we take that. Yes, those are junk ones.
Maker G1:	[Untangles the earphones]
Maker B1:	[Checks the earphones] the lower part looks ok, we can cut that.

In the prototyping session, we observed that the makers trying to blend the Lego robotic kit components with mundane materials like chopsticks, earphone wires, play-doh, tapes, etc. The team members start by building the rope and pulley mechanism for the fork. The makers asked for strings to build the mechanism. The makers and facilitator checks the space for the same, but could not find the required kind of thread. But they found a box consisting of used cables, ropes, wires, paper, defective electronic devices like calculators, and earphones, and the facilitator suggested using the rope for the purpose. One of the makers identifies the possibility of using a defective earphone for the same and which is then cut and joined with Lego rods for making the pulley assembly as seen in Figure 1. Here we see that the makers encounter a wide variety of materials to address the problem in hand. Considering the encounter of the defective earphone and maker, maker-material relation changed and the history evolved. The meaning of the earphone was that of a connecting component as it had the history of being defect, and hence the maker approached the material. The maker and material enter into transforming states as defective earphone is modified to form "rope" undergoing changes in the moment to moment maker-material histories. From an earphone to defective earphone to rope and then to be part of rope and pulley mechanism in relation with makers and constituting materials, the material history evolved.



*Figure 1.* The earphone wire – lego components combination for supporting the gripper movement.

## 4. Discussion & Conclusion

In this article, we have presented a posthumainst view of creativity by looking into a representative episode form a collaborative making context. The findings point to the relational dynamics among humans and non-human elements of the situated making context. The material histories evolved over time with dynamic changes and shifts in emergent quality. We see that the constructivist and cognitivist approach as reductionist research approach where materials are seen as only means to explore problem and solution space of the making-with human exceptionalism. Posthumanism opens ways to disrupt linear ways of knowing, and challenge anthropocentric analysis of maker-material encounters. It can be operationalized to find further meanings of collaboration and emergent creativity coming out of it. This flattened and relational ontological view can reveal how makers and materials shape and create opportunities for learning and creative movements.

## References

- Barad, K. (2003). Posthumanist performativity: Toward an understanding of how matter comes to matter. *Signs: Journal of women in culture and society*, *28*(3), 801-831.
- Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning.* duke university Press.
- Glăveanu, V. P. (2014). *Distributed creativity: Thinking outside the box of the creative individual*. Cham/Heidelberger: Springer International Publishing.
- Haraway, D. (1985). A Manifesto for Cyborgs: Science, Technology, and socialist feminism in the 1980s. *Socialist Review*, *80*(1985), 65–108.
- Jackson, A. Y., & Mazzei, L. A. (2012). Thinking with theory in qualitative research: Viewing data across multiple perspectives/Youngblood. *Imprint Abingdon*.
- Pickering, A. (1993). The mangle of practice: Agency and emergence in the sociology of science. *American journal of sociology*, *99*(3), 559-589.
- Sawyer, R. K., & DeZutter, S. (2009). Distributed creativity: How collective creations emerge from collaboration. *Psychology of aesthetics, creativity, and the arts, 3*(2), 81.
- Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge university press.
- Tanggaard, L. (2013). The sociomateriality of creativity in everyday life. *Culture & Psychology*, *19*(1), 20-32.



