

Simon K. S. Cheung · Fu Lee Wang ·
Naraphorn Paoprasert ·
Peerayuth Charnsethikul ·
Kam Cheong Li ·
Kongkiti Phusavat (Eds.)

Communications in Computer and Information Science

1974

Technology in Education


Innovative Practices for the New Normal

6th International Conference on Technology in Education, ICTE 2023
Hong Kong, China, December 19–21, 2023
Proceedings


Communications in Computer and Information Science

1974

Editorial Board Members

Joaquim Filipe , *Polytechnic Institute of Setúbal, Setúbal, Portugal*

Ashish Ghosh , *Indian Statistical Institute, Kolkata, India*

Raquel Oliveira Prates , *Federal University of Minas Gerais (UFMG),
Belo Horizonte, Brazil*

Lizhu Zhou, *Tsinghua University, Beijing, China*

Rationale

The CCIS series is devoted to the publication of proceedings of computer science conferences. Its aim is to efficiently disseminate original research results in informatics in printed and electronic form. While the focus is on publication of peer-reviewed full papers presenting mature work, inclusion of reviewed short papers reporting on work in progress is welcome, too. Besides globally relevant meetings with internationally representative program committees guaranteeing a strict peer-reviewing and paper selection process, conferences run by societies or of high regional or national relevance are also considered for publication.

Topics

The topical scope of CCIS spans the entire spectrum of informatics ranging from foundational topics in the theory of computing to information and communications science and technology and a broad variety of interdisciplinary application fields.

Information for Volume Editors and Authors

Publication in CCIS is free of charge. No royalties are paid, however, we offer registered conference participants temporary free access to the online version of the conference proceedings on SpringerLink (<http://link.springer.com>) by means of an http referrer from the conference website and/or a number of complimentary printed copies, as specified in the official acceptance email of the event.

CCIS proceedings can be published in time for distribution at conferences or as post-proceedings, and delivered in the form of printed books and/or electronically as USBs and/or e-content licenses for accessing proceedings at SpringerLink. Furthermore, CCIS proceedings are included in the CCIS electronic book series hosted in the SpringerLink digital library at <http://link.springer.com/bookseries/7899>. Conferences publishing in CCIS are allowed to use Online Conference Service (OCS) for managing the whole proceedings lifecycle (from submission and reviewing to preparing for publication) free of charge.

Publication process

The language of publication is exclusively English. Authors publishing in CCIS have to sign the Springer CCIS copyright transfer form, however, they are free to use their material published in CCIS for substantially changed, more elaborate subsequent publications elsewhere. For the preparation of the camera-ready papers/files, authors have to strictly adhere to the Springer CCIS Authors' Instructions and are strongly encouraged to use the CCIS LaTeX style files or templates.

Abstracting/Indexing

CCIS is abstracted/indexed in DBLP, Google Scholar, EI-Compendex, Mathematical Reviews, SCImago, Scopus. CCIS volumes are also submitted for the inclusion in ISI Proceedings.

How to start

To start the evaluation of your proposal for inclusion in the CCIS series, please send an e-mail to ccis@springer.com.

Simon K. S. Cheung · Fu Lee Wang ·
Naraphorn Paoprasert · Peerayuth Charnsethikul ·
Kam Cheong Li · Kongkiti Phusavat
Editors

Technology in Education

Innovative Practices for the New Normal

6th International Conference on Technology in Education, ICTE 2023
Hong Kong, China, December 19–21, 2023
Proceedings

Editors

Simon K. S. Cheung 
Hong Kong Metropolitan University
Hong Kong, China

Fu Lee Wang 
Hong Kong Metropolitan University
Hong Kong, China

Naraphorn Paoprasert
Kasetsart University
Bangkok, Thailand

Peerayuth Charnsethikul
Kasetsart University
Bangkok, Thailand

Kam Cheong Li
Hong Kong Metropolitan University
Hong Kong, China

Kongkiti Phusavat
Kasetsart University
Bangkok, Thailand

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-981-99-8254-7 ISBN 978-981-99-8255-4 (eBook)
<https://doi.org/10.1007/978-981-99-8255-4>

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Singapore Pte Ltd. 2024

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Paper in this product is recyclable.

Preface

This edited volume consists of papers selected from the 6th International Conference on Technology in Education (ICTE 2023), which was held during 19–21 December 2023.

Technology has become an integral part in virtually all aspects of education, broadly covering curriculum planning, content development and delivery, communication among learners, instructors and institution, assessment and programme evaluation. Enabled by the latest technological advances, new and innovative measures are derived to improve teaching and learning effectiveness for education at all levels.

The recent outbreak of the COVID-19 pandemic has radically changed teaching and learning practices with innovative uses of technologies. Many changes are irreversible in the sense that these new practices continue even after the pandemic. Taking the theme “Innovative Practices for the New Normal”, ICTE 2023 serves to provide a platform for relevant academic exchanges across educational institutions, focusing especially on sharing of new research findings and good practices in this new normal era.

ICTE 2023 attracted a total of 75 submissions from authors on different continents. After a careful paper review process, where at least two single-blind reviews were conducted for each submission, 30 papers were selected for inclusion in this volume. These papers are organized in six sections: (1) keynote papers, (2) online and innovative learning, (3) personalized and individualized learning, (4) smart learning environment, (5) artificial intelligence in education, (6) institutional strategies and practices.

Our sincere thanks go to the organizing committee for their effective administration and unfailing support to make ICTE 2023 successful. Our thanks also go to the international programme committee. The high quality of the papers could not have been maintained without their professional comments and advice in the paper review process.

December 2023

Simon K. S. Cheung
Fu Lee Wang
Naraphorn Paoprasert
Peerayuth Charnsethikul
Kam Cheong Li
Kongkiti Phusavat

Organization

Organizing Committee

Conference Chairs

Kam Cheong Li
Kongkiti Phusavat

Hong Kong Metropolitan University, China
Kasetsart University, Thailand

Programme Chairs

Simon K. S. Cheung
Fu Lee Wang
Naraphorn Paoprasert
Peerayuth Charnsethikul

Hong Kong Metropolitan University, China
Hong Kong Metropolitan University, China
Kasetsart University, Thailand
Kasetsart University, Thailand

Organizing Chairs

Lap-Kei Lee
Suwitchaporn Witchakul
Praewpran Prayadsab
Pornthep Anussornnitisarn

Hong Kong Metropolitan University, China
Kasetsart University, Thailand
Kasetsart University, Thailand
Kasetsart University, Thailand

Financial Chairs

Kwan-Keung Ng
Ramidayu Yousuk

Ming-Ai (London) Institute, UK
Kasetsart University, Thailand

Registration Chairs

Chung Keung Poon
Phatcharee Tokaew Thongrattana

Hang Seng University of Hong Kong, China
Kasetsart University, Thailand

Webmasters

Leung-Pun Wong
Kwok Tai Chui

Tung Wah College, China
Hong Kong Metropolitan University, China

International Programme Committee

Chairs

Simon K. S. Cheung	Hong Kong Metropolitan University, China
Fu Lee Wang	Hong Kong Metropolitan University, China
Naraphorn Paoprasert	Kasetsart University, Thailand
Peerayuth Charnsethikul	Kasetsart University, Thailand

Members

Pornthep Anussornnitisarn	Kasetsart University, Thailand
Jinnong Cao	Hong Kong Polytechnic University, China
C. S. Chai	Chinese University of Hong Kong, China
Keith Chan	Hong Kong Polytechnic University, China
Peter Chan	Brigham Young University, USA
Chih-Tsan Chang	National Chung Cheng University, Taiwan RoC
Gary Cheng	Education University of Hong Kong, China
Guo-Li Chiou	National Taiwan Normal University, Taiwan RoC
Thomas Chiu	University of Hong Kong, China
Mike Chui	Education University of Hong Kong, China
Supaporn Erjongmanee	Kasetsart University, Thailand
Jittat Fakcharoenphol	Kasetsart University, Thailand
Chaiporn Jaikaeo	Kasetsart University, Thailand
Yoko Hirata	Hokkai-Gakuen University, Japan
Yoshihiro Hirata	Hokkai-Gakuen University, Japan
Jenq-Muh Hsu	National Chiayi University, Taiwan RoC
Ronghuia Huang	Beijing Normal University, China
John Hui	Vocational Training Council of Hong Kong, China
Josef Hynek	University of Hradec Králové, Czech Republic
Sasarose Jaijit	Kasetsart University, Thailand
Geng Jie	Chinese University of Hong Kong, China
Morris Jong	Chinese University of Hong Kong, China
Blanka Klimova	University of Hradec Králové, Czech Republic
Reggie Kwan	Hong Kong Metropolitan University, China
Ivan Lai	City University of Macau, China
Keith Lee	Hong Kong Metropolitan University, China
Kam Cheong Li	Hong Kong Metropolitan University, China
Martina Manenova	University of Hradec Kralove, Czech Republic
Jaroslava Mikulecká	University of Hradec Králové, Czech Republic

Anan Mungwattana	Kasetsart University, Thailand
Jitti Niramitranon	Kasetsart University, Thailand
Somchai Numprasertchai	Kasetsart University, Thailand
Praewpran Payadsab	Kasetsart University, Thailand
Kongkiti Phusavat	Kasetsart University, Thailand
Punpiti Piamsa-nga	Kasetsart University, Thailand
Juta Pichitlamken	Kasetsart University, Thailand
Chung Keung Poon	Hang Seng University of Hong Kong, China
Linda Pospíšilová	University of Pardubice, Czech Republic
Petra Poulova	University of Hradec Králové, Czech Republic
Praewpran Prayadsab	Kasetsart University, Thailand
Lucie Rohlikova	University of West Bohemia, Czech Republic
Junjie Shang	Peking University, China
Yinghui Shi	Central China Normal University, China
Ju Ling Shih	National University of Tainan, Taiwan RoC
Ivana Simonova	University of Ostrava, Czech Republic
Chansiri Singhtaun	Kasetsart University, Thailand
Niwat Srisawasdi	Khon Kaen University, Thailand
Liana Stanescu	University of Craiova, Romania
Cai Su	Beijing Normal University, China
Kobkul Sunphakitjumnong	Kasetsart University, Thailand
Jay Tashiro	Ontario Tech University, Canada
Shikui Tu	Shanghai Jiao Tong University, China
Olga Vrastilova	University of Hradec Kralove, Czech Republic
Aihua Wang	Peking University, China
Shudong Wang	Shimane University, Japan
Suwitchaporn Witchakul	Kasetsart University, Thailand
Billy Wong	Hong Kong Metropolitan University, China
Crusher Wong	City University of Hong Kong, China
T. L. Wong	Douglas College, Canada
Papis Wongchaisuwat	Kasetsart University, Thailand
Di Wu	Central China Normal University, China
Haoran Xie	Lingnan University, China
Youru Xie	South China Normal University, China
Suwitchaporn Witchakul	Kasetsart University, Thailand
Leung Pun Wong	Tung Wah College, China
Harrison Yang	State University of New York at Oswego, USA
Ramidayu Yousuk	Kasetsart University, Thailand
Zehui Zhan	South China Normal University, China
Sha Zhu	Central China Normal University, China

Organizer



International Hybrid Learning Society

Co-organizers



Kasetsart University



Hong Kong Metropolitan University

Sponsor



Hong Kong Pei Hua Education Foundation

Contents

Keynote Papers

Reimagining Education Systems: How Research on Digital Learning Can Inform Pedagogical Practice	3
<i>Som Naidu</i>	
CiRA-Core: The Connector for Developer Teachers and User Teachers to Artificial Intelligence	13
<i>Rangsan Jomtarak, Chatree Faikhamta, Tharuesean Prasoplarb, and Kornkanok Lertdechapat</i>	
Technology-Enhanced Cybersecurity Education: A VPET Approach	20
<i>Lam-For Kwok, Chen Yi, and Sum Lam</i>	

Online and Innovative Learning

Using Network Analysis to Explore the Factors Influencing Students' Online Learning Effectiveness	35
<i>Sha Zhu, Qing Guo, Wei Qin, Shun Xu, and Harrison Hao Yang</i>	
Learning Image Processing Contents Through Game	47
<i>Suriya Natsupakpong and Suwitchaphon Suksawang</i>	
Student Satisfaction in External Online Course Provider: The Way Forward for WOU?	58
<i>Wen Huey Kan, Li Lian Yap, and Yee Hui Hoo</i>	
Construction and Innovative Application of MOOC for Large-Scale Improvement of Digital Teaching Literacy of Primary and Secondary School Teachers – Taking Guangdong Province of China as an Example	68
<i>Youru Xie, Wan Xia, Caoyin Li, and Yi Qiu</i>	
Measuring Emotion in Education Using GSR and HR Data from Wearable Devices	82
<i>Qian Dong and Rong Miao</i>	
Effect of Instruction Intervention on MOOC Forum Discussion: Student Engagement and Interaction Characteristics	94
<i>Wenyi Chen</i>	

Personalized and Individualized Learning

Emerging Techniques for Online Learning Analytics 109
Yidan Wang, Xiaoran Huang, Qijing Yu, and YuXuan Lai

Prediction of At-Risk Students Using Learning Analytics: A Literature
Review 119
Kam Cheong Li, Billy T. M. Wong, and Hon Tung Chan

Towards the Design of a Personalized Vocabulary Learning System
for University Graduates Based on Context Readability and Task Diversity 129
Gary Cheng and Di Zou

Theoretical Framework of Personal Learning Environments: SPET Model 139
Xiaoshu Xu, Yilin Sun, Jie Weng, and Yunfeng Zhang

Construction and Application of “AI + RPA + X” Intervention Field
for College Students’ Learning Adaptation 157
Yi Qiu, Zhiyang Peng, and Yi Lu

Smart Learning Environment

Utilization of Digital Human Model in Design Education: A Design
Students’ Learning Survey and Case Study 173
Hung-Hsiang Wang and Chih-Ping Chen

A Software for Learning Thai Fingerspelling Sign Language into Thai
Consonants with Machine Learning 186
Supisara Pisuchpen, Chansiri Singhtaun, and Laddawan Suwannachote

A Preliminary Study on Developing and Training the Smart Site Safety
System (SSSS) for Construction Industry in Hong Kong 199
Yan-Wai Chan, Simon K. S. Cheung, Kwan-Keung Ng, and Louise Luk

A Review of Smart Education Practices Across Disciplines 208
Billy T. M. Wong, Kam Cheong Li, and Mengjin Liu

Study on the Influencing Factors of Primary School Students’
Self-regulated Learning in the Smart Classroom: An Environmental
Perspective 218
Yinghui Shi, Xexin Jia, Caiqin Cao, Yitong Wei, and Harrison Hao Yang

Artificial Intelligence in Education

- Integration of AI Learning into Higher Education: A Case of Using
Microsoft Learn for Educators 231
Shao-Fu Li, Kwan-Keung Ng, and Lap-Kei Lee
- A Review of Data Augmentation and Data Generation Using Artificial
Intelligence in Education 242
*Kwok Tai Chui, Lap-Kei Lee, Fu Lee Wang, Simon K. S. Cheung,
and Leung Pun Wong*
- The Role of ChatGPT in Higher Education: Some Reflections from Public
Administration Students 254
*Aleksander Aristovnik, Lan Umek, Nejc Brezovar, Damijana Keržič,
and Dejan Ravšelj*
- Automated Computational Intelligence Based Course Timetabling Tool 264
Marisa Kuntasup, Pupong Pongcharoen, and Thatchai Thepphakorn
- Exploring ChatGPT-Generated Assessment Scripts of Probability
and Engineering Statistics from Bloom's Taxonomy 275
Christopher Chung Lim Kwan

Institutional Strategies and Practices

- A Survey of Evaluation Approaches in STE(A)M Education 289
Billy T. M. Wong, Kam Cheong Li, and Hon Tung Chan
- Uncertainty-Based Metamorphic Testing for Validating Plagiarism
Detection Systems 299
Pak Yuen Patrick Chan, Jacky Keung, and Zhen Yang
- Measuring the Effect of Family Structure on High-School Students' Stress
and Enthusiasm for Learning Using Tourism-Related Questionnaire 315
Pongthorn Ruksorn, Suwitchaporn Witchakul, and Naraphorn Paoprasert
- Towards Education 4.0: A Holistic Approach for Performance Evaluation
of German and Thai Universities 325
Jettarat Janmontree, Hartmut Zadek, and Woramol C. Watanabe
- Digital Technology for the Ageing Population: Perspectives
from the Younger Generation 340
Miloslava Černá, Anna Borkovcová, and Simon K. S. Cheung

Exploring Student Profile Features and Their Impact on Learning
Performance in Secondary School 349
Yicong Liang, Haoran Xie, Di Zou, Xinyi Huang, and Fu Lee Wang

Author Index 361

Keynote Papers



Reimagining Education Systems: How Research on Digital Learning Can Inform Pedagogical Practice

Som Naidu^(✉)

Technology, Education and Design Associates, Melbourne, Australia
sommnaidu@gmail.com

Abstract. The COVID-19 pandemic showed how vulnerable contemporary education systems are to change and disruption to their operational models. The fact that entire education systems globally had to shut down their learning and teaching operations during the pandemic, or move on to a minimalist form of it, requires a critical re-examination of our core business models. Educational provision globally has been under stress long before COVID-19 struck. While the demand for education and re-education grows exponentially, resources and opportunities for the provision of these opportunities are under strain globally. Preliminary assessments of the readiness and resilience of existing models, reports that contemporary education systems are unfit and unsuited to lead learning for tomorrow, and that a re-imagination and re-engineering of institutional choreographies is required. Such a rethink and re-imagination requires greater levels of openness and flexibility in its modus operandi in order to be more resilient to change and disruption. This will involve a rethink and re-imagination of everything from how students access learning opportunities, to how they are taught, assessed and supported by teachers and educational institutions, and importantly, how technology can be used to inform and support pedagogical practices.

1 Introduction

Few would have imagined that the COVID-19 pandemic could cause so much grief and havoc to our operating systems. We hadn't seen anything like it since the 1918–20 Spanish flu epidemic. Although there have been several smaller and more regional outbreaks since then, these have never had the impact of the COVID-19 pandemic. In the case of education, the impacts of the COVID-19 pandemic have been indiscriminate. From resource rich Ivy League institutions to the rest, all education systems had been gravely impacted, forcing them to shut down their operations, and resort to a minimalist form of it to remain viable (JISC 2021). The pandemic revealed how unprepared and ill-equipped the education sector globally, has been for the disruption of its educational operations and core business models.

It is imperative therefore, that we ask how it has been possible for our educational operations to be compromised so completely? Why so many of our education systems universally have had to shut down their learning and teaching activities, or resort to a minimalist form of operation? What does this say about how we have been preparing

our learners and teachers for the contemporary educational space? How prepared are educational institutions for a digitally enhanced and future-focussed education system? But more importantly, how are contemporary universities defining their places in a diverse and competitive environment, and rethinking and redesigning their institutional choreographies to meet the challenges posed by disruptive forces such as COVID-19? (Ernest & Young 2018).

2 Can Contemporary Models Lead Learning for Tomorrow?

A 2020 study of the Australian education system carried out by CISCO (<https://bit.ly/3uTW3kl>; <https://bit.ly/3Bp10UP>), asked how are Australian education institutions reimagining their core functions? The goal of this study was to understand how universities, and TAFEs (Technical and Further Education Institutions) in the Australian education context are reimagining their core functions—namely, teaching, research, administration, and campus design, in the wake of the COVID-19 pandemic. Seventy-eight percent of Australian Universities and five out of seven TAFEs responded to this survey. The survey revealed that COVID-19 pandemic served as a tipping point for open, flexible, online and distance learning (OFDL) causing campus-based operations to reconsider their business models. After decades of existing on the margins of mainstream processes, open, flexible and distance learning (OFDL) was suddenly being seen as the only viable, and possibly the most resilient form of education available (Dede & Lidwell, 2023).

The CISCO study also revealed that it is not just the physical campus spaces that are changing, but that all university functions are being reimagined in the wake of the pandemic. These include all teaching and learning functions, the student learning experience, research and industry engagement, as well as administration. The overall message of this study has been that COVID-19 will change education, and that the design of the physical campus will never be the same again.

So then, what would that rethink and recalibration of educational as well as institutional choreographies for a future focussed education look like? To allow for greatest resilience in institutional operations, such a reimagination will have to feature ideas around *openness* and *flexibility*. Openness and flexibility are value principles. As such they are unlike modes of learning such as *distance*, *blended*, *hybrid* or *face-to-face learning*. Openness and flexibility exist in all modes of learning in different degrees.

The idea of open is deeply grounded in a socio-political agenda which seeks education for all as the path to real freedom and justice (see Table 1). It comprises the practice of *open access*, *open learning* and *open scholarship*. *Open access* is about inclusive and equal access to educational opportunities without barriers such as pre-requisite entry qualifications or ability to pay. The idea of open access is based on the premise that all lives have equal value, and as such everyone has the right to education—not just the privileged few who are able to access and afford the physical campus-based experience.

The idea of open is also about *flexibility* in relation to the time, place and pace of one's learning activities. This is about freedom of choice and being able to exercise it without being locked into a mode or pattern that is based on the preferences of the educational institution. And finally, the idea of open comprises the practice of *open scholarship* which is about the adoption of a culture of openness in relation to how content and intellectual property is shared with others for use, adaptation and redistribution, and at no cost. This dimension of open educational practice is based on the premise that education, not unlike food, water, and shelter is a basic need that should be accessible to all for an equitable social structure.

Table 1. Dimensions of Open Educational Practices

Dimensions of Open education	Pedagogical affordances
Open access	This is about inclusive and equitable access to educational opportunities without barriers such as entry qualifications and ability to pay. It is based on the value that all lives have equal value
Open learning	This is about the ability to learn at anytime, anywhere and at any pace. It is based on the principle that learners ought to possess the freedom and the flexibility to choose their mode, medium, time, place and pace of study
Open scholarship	This is about releasing one's intellectual property under an open license scheme that permits no-cost access, use, adaptation and redistribution. It is based on the value that education is a basic need that should be accessible to all, if we were to achieve education for all, freedom, justice and equality

Flexible learning on the other hand, unlike modes of learning, is a state of being in which the acts of learning and teaching are increasingly freed from the limitations of its *time, place and pace*. For learners, flexibility in learning can include choices in relation to their learning activities, assessment tasks and the time, place and pace of their completion. For teachers this kind of flexibility can comprise choices in relation to the time, place and pace of their communication and engagement with learners as well as the educational organization. Flexibility is a desirable value much like how society regards fairness and equality as desirable goals. Flexibility exists in all forms of communication between teachers and learners, although to varying degrees and along several dimensions. These include flexibility in relation to learner's engagement with the subject matter, their teachers, peers, the learning environment, and the educational institution (Table 2). This also includes learners' engagement with their assessment activities, and teachers' feedback on them. It is about how learners might prefer to interact with the content and critical stakeholders in their learning environment.

Table 2. Dimensions of Flexibility

Dimensions of Flexibility	Pedagogical affordances
Learner-content engagement	This is about learners' engagement and interaction with the subject matter in ways that suit individuals, and their approaches to studying and its time, place and pace
Learner-teacher engagement	This is about choices learners have in relation to the mode and method of their engagement and interaction with their teachers and tutors
Learner-learner engagement	This is about choices learners have in relation to the mode and method of their engagement and interaction with their peers in small and large groups, and in offline and online educational settings
Learner engagement with the learning environment	This is about adaptable access, interaction and engagement with the learning environment (such as with mobile devices, Wi-Fi access and innovative use of study space)
Learner engagement with assessment activities	This is about choices learners have in relation to the fulfillment of their assessment requirements
Learner engagement with feedback	This is about choices learners have in relation to access to feedback on their learning and assessment activities
Learner engagement with the institution	This is about choices learners have in relation to their engagement with the services of the educational institution

3 Useful Lines of Inquiry—The Grand Challenge

Openness and *flexibility* in learning and teaching are relevant in any mode of study including campus-based face-to-face educational operations. As such, one size or approach to open and flexible learning does not, and will not fit all learners, teachers or disciplines. There will be a need for different approaches, with different levels of openness, flexibility, structure and guidance for different cohorts and learning contexts, while the threshold value principles of all approaches remain the same (Naidu, 2017).

The idea of openness and flexibility are not synonymous with mode or models of learning and teaching. And therefore, engagement with openness and flexibility does not mean the death of the campus-based experience. These value principles can, and do exist in *distance learning*, *online learning*, *blended learning*, *hybrid learning*, *hyflex learning*, as well as *campus-based learning* contexts. Decisions around the choice of levels of openness and flexibility in these educational contexts will vary according to their orientation and educational philosophy. Those institutions that purport to be *open education* operations, for instance, will have the balance tipped in favour of greater

openness and *flexibility*, while conventional campus-based educational settings will tip the balance in favour of less openness and flexibility in their operations.

While there are attempts at integrating openness and flexibility in all of the foregoing educational contexts, currently they exist in boutique form, on the side and the peripheries of mainstream practices. Moving these operations, to the center and into mainstream educational processes is the greatest challenge facing institutions. This requires an enterprise-wide rethink of educational and institutional choreographies. And as increasingly wider range of technologies become available with more powerful affordances for integrating openness and flexibility in learning and teaching, a useful line of inquiry ought to be to see how these can support and leverage greater levels of openness and flexibility in education systems.

While the imperative to invest in greater integration of digital technologies might seem obvious for many, how these technologies can be used to leverage openness and flexibility in education systems ought to take center stage. What would education systems that adopt technology to promote openness and flexibility look like? What would be different about them in how their education services are organized and managed? How would role and responsibility in relation to teaching be reorganized? How would academic staff be appointed to positions, and under what terms and conditions? How would academic staff have access to requisite professional development? What kinds of policy frameworks will be required for such future-focused education systems and institutions? Most importantly, what would be different about these considerations from the way they have always been (Peters et al., 2020; UNESCO, 2020)?

4 How Research on Digital Learning Can Inform Pedagogical Practice

Great teaching is about grand designs! This kind of teaching requires careful thought to, not only the subject matter that will need to be taught, but how it will be taught and what tools and technologies will be used by teachers and learners. Great teaching is when students have learned. And this *requires not only an in-depth understanding* of the subject matter content, but also the pedagogy, and the technology that will be used to mediate the transaction. Seen in this manner, teaching is a design activity not unlike the work of engineers, architects, or choreographers (Lee, 2023; Kolodner, 2023).

While great teaching is informed by a knowledge of the learning sciences, it is also an outcome of creativity (Myers, & Adams-Budde, 2016; Robinson, & Aronica, 2015). At the heart of synergies derived from this kind of expertise lies knowledge about technology, the pedagogy and the subject matter. Popularized as *technological pedagogical content knowledge* (TPACK), this framework has its origins in Shulman's (1986) idea of *pedagogical content knowledge* (PCK), which is knowledge about learning and cognition. When combined with technological, and subject matter knowledge, these three knowledge domains provide the critical ingredients for the design of productive learning experiences, and great teaching is the about the design of these learning experiences.

Seen in this way, the term *delivery* becomes an inadequate descriptor of what goes into teaching. Teaching is not about “delivering” anything to anyone (Clark, 1994). It is about the *design of productive learning experiences*, which requires getting the “*mixture right*”

between how much of it is going to be one-on-one or group-based, face-to-face, online, at a distance or blended. The one-to-one model of learning—although powerful, is not necessarily the gold standard. As such there can be no generalizable model of teaching (Dron, 2022). From independent study to group-based learning, design is critical to the development of effective, efficient and engaging learning experiences. In the case of a lecture, for instance, along with defining its focus and scope, a teacher needs to research, compile and present the subject matter content, considering issues such as sequence and timing, the audience and the ambience of the venue in order to ensure achievement of the goals of the lecture and ensure maximum impact.

The adoption of new tools and technologies to leverage key teaching and learning functions is an important part of that consideration (Fobes, & Kaufman, 2008). Digital technologies support a wide range of affordances for teaching and learning. As these tools and technologies become more accessible, teachers as well as learners will need to regularly reskill themselves in order to be effective and efficient in the contemporary educational space, where the delivery and mediation of the subject matter knowledge is but a small part of the teaching and learning transaction.

5 Useful Lines of Inquiry—What is Worth Investigating and How?

There is a lot that we already know about human learning. Therefore, putting a finger on what is worth investigating is key (Glassman 2001). This is not only about identifying a worthwhile research question but also making sure that it is a researchable question—which means that the research question is framed in such a way that it can be reliably and validly investigated. When identifying what is worth investigating, it is important to remember that you are joining a conversation, so make sure you have something “unique” to say on the topic. Most submissions for publication consideration are rejected because they make little contribution to the conversation, or issues confronting the field, or they are not contextualized, and as such not aligned with the aims and aspirations of the targeted publication outlet.

Contributions for publication consideration need to be aware of what is known, and not known about a topic, what is worth investigating, and why we need to bother with it? Consequently, the literature review is central to this task (Gredler and Sheields 2004). A good literature review is a lot more than a summary of what has been done or said on a topic. A good literature review synthesizes all that has been done and said in relation to a query or question. It is only when this is achieved, that it is possible to arrive at the research questions which point to what is not known about the topic and which is worth investigation or further research.

Clear research questions are critical to the next step in the process which is the selection of the appropriate research methodology. There are several research methods to choose from, and these vary from pure experimental designs to more ethnographic approaches, and both may comprise the use of quantitative as well as qualitative data gathering tools. A great deal of care is required in matching research questions with appropriate research methodology. Insistence on methodological rigor is crucial in this process (Naidu, 2015a, 2015b). A lack of methodological rigor or a substantial awareness of the existing literature are the most common reasons for the failure of a lot of

submissions (Mishra, Sahoo, & Pandey, 2021). The focus on typical questions is also a common problem facing submissions for publication consideration. These include over-researched questions about the difference between online and campus-based models, whether institutions should specify the mode of learning on the award or certification, the perception in industry and academia of online education, and what needs to be done so they can be treated equally.

At first glance, these lines of questioning look interesting and relevant, but on closer look they are fundamentally flawed. For starters, comparing modes of learning and teaching is not only difficult because iterations of a mode are likely to differ, but also not very useful unless the conditions of learning and teaching that are being compared are exactly the same across educational contexts. If one chooses this as a focus then a more useful line of inquiry might be to look for insights on particular attributes or affordances of learning and teaching. Table 3 provides some insights on how best to approach investigating the affordances of these attributes and particular pedagogical practices.

Table 3. Useful lines of inquiry

Pedagogical practices	Pedagogical affordances
1. Open access	<ul style="list-style-type: none"> • How does access to learning opportunities influence key educational dimensions such as participation and socio-economic development of individuals and societies?
2. Open learning	<ul style="list-style-type: none"> • How does various levels of flexibility in learning and teaching influence key educational dimensions such as participation, persistence and success with learning and teaching?
3. Open scholarship	<ul style="list-style-type: none"> • How does open access to content and products of intellectual outputs and educational resources influence learning and teaching experience design?
4. Learner-content engagement	<ul style="list-style-type: none"> • How can technology be used to support learners' engagement and interaction with the subject matter in ways that suit individuals, and their approaches to studying and their time, place and pace?
5. Learner-teacher engagement	<ul style="list-style-type: none"> • How can technology be used to provide learners with choices in relation to the mode and method of their engagement and interaction with teachers, tutors and other agents?

(continued)

Table 3. (continued)

Pedagogical practices	Pedagogical affordances
6. Learner-learner engagement	<ul style="list-style-type: none"> • How can technology be used to provide learners with choices in relation to the mode and method of their engagement and interaction with their peers in small and large groups, and in offline and online educational settings?
7. Learner engagement with the learning environment	<ul style="list-style-type: none"> • How can technology be used to support adaptable access, interaction and engagement with the learning environment?
8. Learner engagement with assessment activities	<ul style="list-style-type: none"> • How can technology be used to provide learners with choices in relation to the completion of their assessment requirements
9. Learner engagement with feedback	<ul style="list-style-type: none"> • How can technology be used to provide learners with choices in relation to their access to feedback on their learning and assessment activities
10. Learner engagement with the institution	<ul style="list-style-type: none"> • How can technology be used to provide learners with choices in relation to their engagement with the services of the educational institution

6 What Authors Ought to Know About Best Practices in Publishing

There are many things that researchers need to consider and be able to do in order to ensure that their work gets published in the targeted publications (Hartshorne, Ferdig, & Bull, 2021). Many of these considerations have to do with being familiar and conversant with the peer review, and the publications process, and how to negotiate that space. These include following the stipulated guidelines of the targeted publication, its conventions and publication style. Other factors have to do with the contribution of one's research to the wider community and society more generally.

The ultimate goal of research is to make a contribution to knowledge. In order to be able to do so, one needs to possess a very deep and complete understanding of the literature on the topic. It is only when this is achieved that one is able to identify what is not known about a topic, and therefore worth exploring and investigating. These are your research questions and only when this is known and decided upon, that it is possible to select the approach to the investigation (i.e., the research method). Problems with selecting an appropriate research method usually arise when the research question is unclear or poorly articulated. Matching the research question with a research method is crucial to the success of a project. Compromising on methodological rigor is never a good idea, and it should not be negotiable.

The importance of getting these steps right in the conduct of any research cannot be overstated, and not negotiable. These attributes of a research project, unlike its composition, argumentation and adherence to publication style etc., cannot be fixed after the research is conducted just as the foundation stones of a building cannot be replaced once the building is built. As such, the fate of the publication is sealed long before it reaches its reviewers (Naidu, 2021). Getting this right in the conceptualization and the design phases of the research project will very likely assure publication of its outputs.

References

- CISCO: The tipping point for digitization of education campuses. https://www.cisco.com/c/en_au/solutions/industries/tipping-point-report.html (2020)
- Clark, R.E.: Media will never influence learning. *Educ. Technol. Res. Dev.* **42**, 21–29 (1994). <https://doi.org/10.1007/BF02299088>
- Dede, C., Lidwell, W.: Developing a next-generation model for massive digital learning. *Educ. Sci.* **13**, 845 (2023). <https://doi.org/10.3390/educsci13080845>
- Dron, J.: Technology, teaching, and the many distances of distance learning. *J. Open, Flex. Distance Learn.* **26**(2), 7–17 (2022)
- Ernest & Young: Can the universities of today lead learning for tomorrow? The University of the Future. <https://go.ey.com/3w0HsCV> (2018)
- Fobes, C., Kaufman, P.: Critical pedagogy in the sociology classroom: challenges and concerns. *Teach. Sociol.* **36**(January), 26–33 (2008)
- Glassman, M.: Dewey and Vygotsky: society, experience, and inquiry in educational practice. *Educ. Res.* **30**(4), 3–14 (2001). <https://doi.org/10.3102/0013189X030004003>
- Gredler, M., Shields, C.: Does no one read Vygotsky’s words? commentary on glassman. *Educ. Res.* **33**(2), 21–25 (2004). <https://doi.org/10.3102/0013189X033002021>
- Hartshorne, R., Ferdig, R.E., Bull, G.: What Journal Editors Wish Authors Knew About Academic Publishing. Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/primary/p/219093/> (2021). Retrieved 3 Oct 2021
- JISC: Student digital experience insights survey 2020/21. Findings from UK higher education (pulse 1: October–December 2020). <https://repository.jisc.ac.uk/8318/1/DEI-P1-HE-student-briefing-2021-FINAL.pdf> (2021)
- Kolodner, J.L.: Learning engineering: what it is, why I’m involved, and why I think more of you should be. *J. Learn. Sci.* **32**(2), 305–323 (2023). <https://doi.org/10.1080/10508406.2023.2190717>
- Lee, V.R.: Learning sciences and learning engineering: a natural or artificial distinction? *J. Learn. Sci.* **32**(2), 288–304 (2023). <https://doi.org/10.1080/10508406.2022.2100705>
- Mishra, S., Sahoo, S., Pandey, S.: Research trends in online distance learning during the COVID-19 pandemic. *Distance Educ.* **42**(4), 494–519 (2021). <https://doi.org/10.1080/01587919.2021.1986373>
- Myers, J., Adams-Budde, M.: Creative schools: the grassroots revolution that’s transforming education. *Int. Rev. Educ.* **62**, 375–378 (2016). <https://doi.org/10.1007/s11159-016-9539-8>
- Naidu, S.: How flexible is flexible learning, who is to decide and what are its implications? *Distance Educ.* **38**(3), 269–272 (2017). <https://doi.org/10.1080/01587919.2017.1371831>
- Naidu, S.: The fate of a submission is sealed long before its consideration for publication! In: Hartshorne, R., Ferdig, R.E., Bull, G. (eds.) *What Journal Editors Wish Authors Knew About Academic Publishing*, pp. 29–36, Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/p/219093/> (2021)

- Naidu, S.: Lessons we are not learning or choosing to ignore! *Distance Educ.* **36**(3), 291–294 (2015). <https://doi.org/10.1080/01587919.2015.1083645>
- Naidu, S.: Methodological issues in educational research. *Distance Educ.* **36**(1), 1–4 (2015). <https://doi.org/10.1080/01587919.2015.1030097>
- Peters, M.A., et al.: Reimagining the new pedagogical possibilities for universities post-Covid-19. *Educ. Philos. Theory* **54**(6), 717–760 (2022). <https://doi.org/10.1080/00131857.2020.1777655>
- Robinson, K., Aronica, L.: *Creative Schools: The Grassroots Revolution That's Transforming Education*. Viking, New York (2015)
- UNESCO: *Futures of education*. <https://en.unesco.org/futuresofeducation/initiative> (2020)



CiRA-Core: The Connector for Developer Teachers and User Teachers to Artificial Intelligence

Rangsan Jomtarak¹, Chatree Faikhamta^{2(✉)}, Tharueseen Prasoplarb²,
and Kornkanok Lertdechapat³

¹ Faculty of Science and Technology, Suan Dusit University, 295 Nakhon Ratchasima Road
Dusit District, Bangkok 10300, Thailand

² Department of Education, Faculty of Education, Kasetsart University, 50 Ngamwongwan
Road, Lat Yao, Chatuchak, Bangkok 10900, Thailand
feductf@ku.th

³ Curriculum and Instruction, Chulalongkorn University, 254 Phayathai Road, Pathumwan,
Bangkok 10330, Thailand

Abstract. Technology becomes the essential connector of transdisciplinary education as known as STEM education and as the tool for enhancing learners to become a technologically literate person and use information and technology as young innovators. However, there were many studies aimed on the learning innovation developments but only a few teachers could choose or access the appropriate platform to integrated technology in their lessons. This study aimed to survey the similarities and the different views about the nature, purposes, and functions of technology in education between 17 CiRA Core agent-teachers and 69 user-teachers to using CiRA Core through an open-ended questionnaire. Data were analyzed by content analysis, and all codes were then categorized to find the patterns and become the theme. The results visualized the 3 aspects which included the nature, the purposes, and the functions present the characteristic of reflection, connection, and driving the integrated technology of CiRA Core in the instructions consecutive. The most different view between them was found at the purpose of integration of technology connects the new norm teaching (differ 47.31%), while the most similarity view between them was found at the purpose of integration of technology connects the 21st classroom. However, this study found out the most average differ view of nature of technology. That would become an important issue for the future study of professional development to emphasize this aspect for more synchronizing between the developer-teachers and the user-teachers to integrate technology in the instructional design.

Keywords: nature of technology · STEM education · technology education · CiRA-Core

1 Introduction

One of the important goals in education is to prepare students to become active citizens (Maass, Geiger, Ariza, & Goos, 2019). The idea of connecting different disciplines is not new to science education since the concept of STEM (STEM Education) is defined and

disseminated widely. To connect students to real and complex problems in everyday life (Roehrig, Dare, Ellis, & Ring-Whalen, 2021). STEM education become an alternative approach to bring learners achieve the above goals (Ra-Ngubtook, 2020).

The CiRA-Core platform is an educational innovation that promotes problem solving for students. Through the integration of technological thinking together with the Engineering Design Process (EDP) (Han & Kelley, 2022), students learn the concept of deep learning artificial intelligence regarding computing. Working like human neural networks, artificial intelligence leads to innovation in science and technology, engineering or medicine (Kittichai et al., 2021; Klinieam & Noiying, 2021). In Thailand, CiRA-Core platform has been used to train students at the secondary school level through independent study using the image detection process to innovation from Classes such as sorting coffee beans, Facial detection instead of checking student names, Food classification analysis, etc. The platform is therefore an alternative for classrooms that aim to develop students to have the characteristics of problem solvers (Sung & Kelley, 2022) based on empirical evidence to indicate the effectiveness of solutions (Roehrig et al., 2021).

Consequently, preparing teachers to be able to effectively teach STEM through CiRA-Core by integrating technology is challenge efforts (Wang, Chen, Hwang, Guan, & Wang, 2022) One of the challenges is how to connect teachers' perspectives or beliefs towards the goal of STEM, how to engage teachers use technology in STEM (Icel, 2018) from the level of using technology as a teaching tool to another level which learners as a maker or problem solver using AI. To level up teachers use artificial intelligence technology, understanding teachers' perspectives of using CiRA-Core is the main goal of this study.

1.1 Objectives

To compare teachers' views as teacher developers and teacher users regarding the CiRA-Core platform.

1.2 Literature Review

The professional development program becomes a shared learning space between a group of teacher developers (Developers) who play the role of training organizers with user teachers (Users) who play the role of training participants (Srikoom, Hanuscin, & Faikhamta, 2017). Previous studies indicated that most teachers still do not understand learning management with technology, such as the nature of technology, knowing how to use technology effectively and the notion that technology aims to respond to human needs (Bybee, 2010; Pleasants, 2020). Technology can be a process with clear goals and have the possibility of errors in technological systems (Kruse, 2013). The social context affects the creation of technology (Bybee, 2010; Kelley & Knowles, 2016). In this study, we combined the concept of the nature of technology and the integration of technology with teaching specific content knowledge (TPACK) (Mishra & Koehler, 2006), into a conceptual framework as follows;

- 1) Natures of Integration Technology (NOIT); Using knowledge along with the integration among other disciplines such as science by considering social context.

- 2) Purposes of Integration Technology (POIT); Connecting students to everyday life situations by focusing on using scientific, mathematical, or technological knowledge to solve problems. Connecting learning experiences with technology, innovation, artificial intelligence or big data develops 21st century skills.
- 3) Functions of Integration Technology (FOIT); Representing the view of technology as a subject, knowledge or information. Understanding the operation of technological processes or technological design processes.

2 Data Collection and Research Methodology

This research is qualitative research using a multiple case study in order to compare the views between two groups of research participants: Group 1, a group of 17 teacher developers (as A1, A2, A3 ... A17, respectively). These participants play a role as a leader and facilitator in using the Cira-Core platform. This group of teachers has at least 12 months of experience in using the Cira-Core platform. The second group is a group of 69 teacher users. They (as B1, B2, B3 ... B17, respectively) have the role of regular teachers who participate in training on using the Cira-Core platform. They have previous experience using the Cira-Core platform prior to training less than 3 months of experience teaching or attending training.

The researcher developed and used an open-ended questionnaire to explore the views of both groups of research participants on the CiRA-Core platform. Questions included four main areas: background in teaching with technology; Goals and limitations in organizing learning with technology and learning management related to artificial intelligence. The open-ended questionnaire was validated 3 STEM experts. The questionnaire was administered by online (Table 1).

Table 1. Examples of questions in the survey of the views of the teacher

Main points	Questions for developer group	Questions for user group
Limitations or obstacles in learning management	What problems have you encountered in organizing learning in the past 3 years ? Have you ever used the research process to solve such problems? How?	What are the obstacles in driving a curriculum that focuses on the use of artificial intelligence? What are the limitations of organizing learning with artificial intelligence in the classroom?
	Learning management related to the platform CiRA-Core	What strengths and limitations do you think the CiRA-Core platform has? Platform How does CiRA-Core change teachers' teaching or student learning?

The researcher conducted data analysis in 2 phases, according to the sequence of data analysis operations. Researchers analyzed the data obtained from the questionnaire using deductive analysis and determined codes. The researcher then read the answers that appeared in the collected questionnaires. Then code the text of the responses to the questions of the participants in the research. Using the code that was previously defined. And the code is updated periodically to cover every answer of research participants. As shown in the list of codes in Table 2.

Table 2. List of perspectives on understanding technology-integrated teaching methods

code	Code description
<i>Natural aspects of technology integration</i>	
N1 – Process on Purpose	Technology – Use the process with purpose
N2 – Involve Thinking Skills in Process	Technology – Integrate thinking skills into the process
N3 – Technology to Transdisciplinary	Technology – There is cross-disciplinary integration
N4 – Technology Depends on Social Context	Technology – consistent with the social context
N5 – Technology Could Fail	Technology – Errors can occur
<i>Target areas for technology integration</i>	
P1 – Authentic STEM	Connected technology with STEM education and real situations
P2 – Emphasized Science Program	Connected technology with classes that focus on science and math
P3 – Equity/Inequity	Connected technology – Creating equality in the classroom
P4 – Innovative Education	Connected technology – Learning towards innovation
P5 – Learning Experience	Connected technology – Learner’s learning experience
P6 – Learning Experience (with AI)	Connected technology – Artificial intelligence learning experience
P7 – New Norm Teaching	Connected technology – Organizing new ways of learning
P8 – 21 st Classroom	Connected technology – Practicing 21st Century Skills
<i>Functions of technology integration</i>	
F1 – Subj/Knowledge (Information)	Mechanisms for teaching concepts / content about technology

(continued)

Table 2. (continued)

code	Code description
F2 – Technological Thinking (CT/Coding)	Mechanisms for developing technological thinking / coding
F3 – Technological Process (Design)	Mechanisms regarding the technological design process
F4 – Technological Innovation (Robot)	Mechanisms for driving innovation / artificial intelligence
F5 – Applied Technology (Big Data)	Mechanisms regarding information management

3 Result and Discussion

Teacher Developer Group and Teacher User Group perspective on CiRA-Core Platform are different, especially in terms of nature of technology integration, goals for integrating technology and roles in integrating technology.

Part 1: Perspective on the nature of technology integration. It can be concluded that only the cross-disciplinary integration issue appeared in which teacher users viewed that the CiRA-Core platform is important and possible to use in real classroom and it can be integrated with science, mathematics, technology, and engineering (Roehrig et al., 2021). They viewed that technological work is often related to cross-disciplinary integration in order to understand problems or challenges in designing solutions (Frias, Ertas, Tate, & Back, 2015). However, the research results also found that both groups of teachers had the view that CiRA-Core Platform does not reflect technological errors. This may be because both groups of teachers viewed that CiRA-Core Platform has effectively been developed and be ready to use. They may not find any errors in the work. This findings may different from other STEM teachers how did not use CiRA-Core Platform, whow faced problems or errors in using technology in the classroom. Such problems or errors are important issues that teachers must be aware of using that technology (Ozel, Yetkiner, & Capraro, 2008).

The second aspect is the goal perspective connecting technology integration with the CiRA-Core platform. It can be concluded that the teacher developer group and the teacher user group have the most different views. The group of teacher developers has a higher perspective on using the platform in that technology play vital roles in connecting students to a new way of learning, than the group of teacher users. Since the CiRA-Core platform has important characteristics that promote student learning into the learning space in the form of artificial intelligence. Students can use the platform to design solutions that are appropriate for the situation. The group of regular teachers should understand learning management using that platform. Therefore, teachers may have less views on this aspect than the development teachers group. However, the issue that both groups of teachers have the most similar views is using the CiRA-Core platform regarding the role of technology in promoting student learning in the 21st century. It may be because both groups of teachers have common experiences in promoting student

learning in a variety of ways. Therefore, it can be linked to learning design that promotes learning skills in the 21st century (Srikoom et al., 2017).

The third aspect is driving technology integration with the CiRA-Core platform. It can be concluded that the teacher developer group and the user teacher group have different views in this area. The teacher developer group had slightly higher views in every aspect. In the role of the technological design process, the user teacher group had a slightly higher than the developer teacher group. This is because of the nature of the technological design process aimed to present obvious learning management steps that can be brought to the classroom. The groups of teacher users may find possibility to design their learning activities consistent with the important concept of technology literacy. It is also easily related to the work that responds to human needs under culture, society, economy, politics, and the environment (Kelley & Knowles, 2016). Teacher learning management then can encourage students to learn through the CiRA-Core platform meaningfully.

3.1 Recommendations

1. Preparing or supporting teachers of technology subjects or technology-based subjects to understand the nature of technology in teaching and learning. Doing so can help those teachers designing teaching that integrate technology effectively regarding to the nature of technology.
2. It is also necessary to regularly compare or check the consistency of views or understanding between the group of teacher developers and the group of users using the CiRA-Core platform. This was to ensure that teachers on both group had the same understanding of the issues raised during the workshop, such as the goals of using the CiRA-Core platform. To enhance the learning of students not just for fun but to able to solve the problems.

References

- Bybee, R.W.: Advancing STEM education: a 2020 vision. *Technol. Eng. Teach.* **70**(1), 30–35 (2010)
- Frias, K., Ertas, A., Tate, D., Back, S.: Shifting engineering education from disciplinary to transdisciplinary practice. *Int. J. Eng. Educ.* **31**(1), 94–105 (2015)
- Han, J., Kelley, T.R.: STEM Integration through shared practices: examining secondary science and engineering technology students' concurrent think-aloud protocols. *J. Eng. Des.* **33**(5), 343–365 (2022). <https://doi.org/10.1080/09544828.2022.2042498>
- Icel, M.: Implementation of STEM policy: a case study of a STEM-focused urban charter school. *J. STEM Educ.* **19**(3), 7–13 (2018)
- Kelley, T.R., Knowles, J.G.: A conceptual framework for integrated STEM education. *Int. J. STEM Educ.* **3**, 1–11 (2016). <https://doi.org/10.1186/s40594-016-0046-z>
- Kittichai, V., et al.: Deep learning approaches for challenging species and gender identification of mosquito vectors. *Sci. Rep.* **11**(1), 4838 (2021). <https://doi.org/10.1038/s41598-021-84219-4>
- Klinieam, K., Noiying, P.: Deep learning technique for a identify TE student system by face recognition. In: Proceedings of the 2021 Research, Invention, and Innovation Congress: Innovation Electricals and Electronics (RI2C), pp. 298–302 (2021). <https://doi.org/10.1109/RI2C51727.2021.9559823>

- Maass, K., Geiger, V., Ariza, M.R., Goos, M.: The role of mathematics in interdisciplinary STEM education. *ZDM* **51**(6), 869–884 (2019). <https://doi.org/10.1007/s11858-019-01100-5>
- Mishra, P., Koehler, M.J.: Technological pedagogical content knowledge: a framework for teacher knowledge. *Teach. Coll. Rec.* **108**(6), 1017–1054 (2006)
- Ozel, S., Yetkiner, Z.E., Capraro, R.M.: Technology in K-12 mathematics classrooms. *Sch. Sci. Math.* **108**(2), 80–85 (2008). <https://doi.org/10.1111/j.1949-8594.2008.tb17807.x>
- Pleasant, J.: Inquiring into the nature of STEM problems: implications for pre-college education. *Sci. Educ.* **29**(4), 831–855 (2020)
- Ra-Ngubtook, W.: Thai learners' key competencies in a VUCA world. *J. Teach. Prof. Dev.* **1**(1), 8–18 (2020). [in Thai]
- Roehrig, G.H., Dare, E.A., Ellis, J.A., Ring-Whalen, E.: Beyond the basics: a detailed conceptual framework of integrated STEM. *Disc. Interdisc. Sci. Educ. Res.* **3**(1), 1–18 (2021). <https://doi.org/10.1186/s43031-021-00041-y>
- Srikoom, W., Hanuscin, D.L., Faikhamta, C.: Perceptions of in-service teachers toward teaching STEM in Thailand. *Asia – Pacific Forum on Sci. Learn. Teach.* **18**(2), 1–23 (2017)
- Sung, E., Kelley, T.R.: Using engineering design in technology education. In: Williams, P.J., von Mengersen, B. (eds.) *Applications of Research in Technology Education: Helping teachers develop research-informed practice*, pp. 133–147. Springer, Singapore (2022)
- Wang, L.-H., Chen, B., Hwang, G.-J., Guan, J.-Q., Wang, Y.-Q.: Effects of digital game-based STEM education on students' learning achievement: a meta-analysis. *Int. J. STEM Educ.* **9**(1), 1–13 (2022). <https://doi.org/10.1186/s40594-022-00344-0>



Technology-Enhanced Cybersecurity Education: A VPET Approach

Lam-For Kwok^(✉), Chen Yi, and Sum Lam

Institute of Cyberspace Technologies, HKCT Institute of Higher Education, 2 On Shing Street,
Ma On Shan, Sha Tin, New Territories, Hong Kong SAR, China
{lflkwok, samlam}@hkct.edu.hk, alexyi@ctihe.edu.hk

Abstract. There is a shortage of skilled talents in the field of cybersecurity at all levels globally. The shortage of the skilled professionals itself becomes a real threat to the society. While education and training aim to produce graduates to fill the job market, the industry complains on their lack of practical skills. To provide more hands-on exercises in obtaining the necessary skills required to perform certain roles, a controlled virtual environment such as cyber range is adopted. A VPET approach is proposed by grouping the knowledge element with the required skill sets together supported by hands-on exercises into a learning unit. These units can be packaged into micro-credentials to provide a more flexible way of student learning or skill re-training.

Keywords: cybersecurity education · VPET · cyber range · technology-enhanced learning

1 Introduction

The use of information technology (IT) has immersed into our daily lives and provides people with a lot of convenience. Mobile devices connecting through the Internet make business operations available to people at almost anytime and anywhere. The emerging use of Internet of Things (IoT) enables numerous applications in the smart city development such as smart traffic controls and smart parking; surveillance for public safety and utilities management, etc. The widely use of IT leads to the increasing number of threats in the cyberspace.

1.1 Cybersecurity – How Serious It is

News on cyber-attacks appears every day around the world. It was reported that a 38% increase in global attacks in 2022 (Anderson, 2023) and a continued rise in attacks against all industries including educational institutions and healthcare organizations. The educational and research institutions received an average of 2,314 attacks per week per organization, an increase of more than 40% from 2021 (Check Point Software Technologies, 2023). In the latest Internet Crime Report (FBI, 2023), it was reported that FBI received more than 800,000 cybercrime-related complaints in 2022, with losses

totaling over \$10 billion. The Latest Cyber Crime Statistics (AAG, 2023) estimated that the potential losses to cybercrime for individuals is more than \$10.2 billion in the US in 2022, much higher than an estimated loss of \$6.9 billion for individuals in 2021. It is also expected eCommerce frauds to cost the retail sector \$48 billion globally in 2023.

1.2 Skill Shortage in Cybersecurity

Cybersecurity becomes an important element for business survival. While it is realized that cybersecurity is important, its success relies on the availability of cybersecurity professionals. There is a shortage of skill talents in this field at all levels globally. The situation is getting worse. The International Information System Security Certification Consortium ((ISC)²), an international nonprofit membership association for cybersecurity professionals, estimated in the Cybersecurity Workforce Study (2022) that there is a worldwide shortage of 3.4 million cybersecurity workers.

One of the reasons of skill shortage is the increase in demand for cybersecurity professionals as a result of increasing digital transformation in business (De Zan, 2019). The widely popular use of e-commerce and e-payment has extended the attack surface leading to difficulties in protecting such transactions. The technologies behind are more advanced and are changing rapidly. The skills required for security professionals are also changing which requires continuous professional development for practitioners. The high compliance requirements for data protection leads to the higher requirement in cybersecurity. The skill shortage becomes a real threat which leaves key sectors critical to our societies vulnerable to cyberattacks (World Economic Forum, 2023).

1.3 Aim and Objectives

Cybersecurity education and training seems to be an indispensable solution in providing more skill talents and close the gap of manpower shortage. In this paper, challenges in cybersecurity education were discussed; a skill-based VPET approach supported by technology enhanced laboratory in cybersecurity education was suggested with current progress and achievements reported; and implications were discussed.

2 Cybersecurity Education

2.1 Some Previous Works

Although there is a quite a number of academic programmes around the world in cybersecurity, the number of graduates still cannot fill the job market. Such shortage has lasted for some years (Mountrouidou et al., 2019; (ISC)², 2020; (ISC)², 2022), despite of efforts has been made in providing more skill talents in this area.

Caelli and Liu (2018) analyzed the cybersecurity courses offered at formal university level in Australia and found that courses available appeared quite limited in scope. The number of graduates produced in cybersecurity was far below the need in the market. To meet the demand of the manpower shortage and attract students with diverse academic backgrounds, some postgraduate courses were developed in Australia (Sitnikova, 2018).

Tsado (2019) recognized the necessity to fill the cybersecurity skill gap and suggested that a top-driven, multidisciplinary school-wide approach is the best in planning and implementing successful cybersecurity educational programs involving both technical and nontechnical disciplines.

Mishra et al. (2019) investigated the students' perceptions in cybersecurity programs about the skills, competencies, and interests needed in the cybersecurity field, and recommended cybersecurity programs trying to teach the latest tools and techniques used in industry, and to provide sufficient hands-on activities using different teaching styles.

Venter et al. (2019) argued that cyber security should be taught as essential skills at school level. Preparing children with awareness and knowledge early in primary schools can ensure that people have the skills when they need them and can remove the current gender imbalance in cyber security awareness. Triplett (2023) proposed game-based strategies to increase students' cybersecurity awareness and motivate them early in the K-12 education to pursue cybersecurity as a career.

Blazic (2021) analyzed the topics missing within the cybersecurity curricula of the higher institutions in Europe and in the courses provided by cybersecurity industry. He suggested that integration of the new topics within the existing frameworks supported by hands-on training should become a continuous practice for addressing the cybersecurity skill shortage; and collaboration between academia, governments and employers is necessary to identify the foundational knowledge and skills with appropriate actions to reduce gap in skill shortage.

2.2 Challenges

The Board Curriculum

The cybersecurity discipline consists of a broad range of specialized knowledge with a solid foundation in computer science (CS). Since the security of cyberspace covers wider areas including electromagnetic space security, physical layer to application layer security, the basic theories would then involve information and communication technology (ICT), cryptography, electronics and information engineering (EIE), computer devices, telecommunication technologies and more. It is difficult for students to fully grasp the knowledge and skills of a range of foundation courses together with some specialized courses in a traditional university curriculum.

Diverse Background of Teaching Staff.

Most of the existing teaching resources in cybersecurity also come from many different professional backgrounds, from ICT, cryptography, EIE and telecommunication, etc. Their mastery of the basic theoretical knowledge in cybersecurity is a big challenge, which requires continuous professional development for themselves.

Industrial Expectation.

Employers often find that graduates do not meet industrial expectation with too much theory and insufficient practical skills in cybersecurity. There is always a complaint that graduates of academic programmes in cybersecurity cannot go directly into the job market without further training because they do not have the necessary hands-on skills.

However, university education aims to provide theoretical knowledge in general and does not bind to using specific products; while training aims to provide learners with practical skills focusing on the current technology and methods. To this end, it is often the case that universities will have difficulties in providing suitable laboratories with up-to-date technologies for students practice. It is also difficult to recruit academic staff good in research with sufficient practical experience in academic institutions.

Education vs Training.

Teaching students to implement specific firewall rules on a specific brand of routers is training for specific skills. Education tries to explain the reason, the theory, the process behind, the potential application areas, their strength and weakness, which is referred to as knowledge (Conklin et al. 2014). However, the industry expects graduates does not only have the necessary knowledge but also acquire the practical skills on a specific discipline. Preparing students practicing with the latest technologies, the education providers need to closely work with the industry and partnering with vendors.

2.3 Technology-Enhanced Learning for Cybersecurity

The need for more practical hands-on exercises for learners in cybersecurity leads to the desire from educators to look for suitable training tools and platforms. De Leon et al. (2018) recognized that using hands-on exercises is an effective way of teaching and learning cybersecurity, which may engage learners in practicing essential skills; but preparing these exercises on an isolated and virtualized environment requires tremendous efforts. To solve these problems, they developed a system called ADLES, which enables formal declarative specifications for educational exercises, and deploy the virtual environment semi-automatically. Chou (2019) implemented a game-based learning system to promote cyber security awareness training, which attempted to provide students nine CyberSec labs to discuss real-world cyber security issues and gain knowledge to a certain cyber-attack and defense. Khan et al. (2022) proposed a game-based learning platform to enhance cybersecurity education for university students, which provides a virtual lab for students to take up cybersecurity challenges and to gain hacking skills with ethics in a safer environment. Rebecchi et al. (2022) described the design of the SPIDER cyber range replicating a customized 5G network environment and enabling interactive cyber exercises, which aims to train 5G security professionals to achieve the ability in managing and predicting security incidents, complex attacks, and propagated vulnerabilities in a collaborative manner.

In general, these attempts and efforts cannot be widely adopted. Cybersecurity laboratories are still extremely hard to setup which require a lot of skillful manpower with a long lead time. Cyber range comes into the scene under this background. The term cyber range was derived from shooting range in the military sector, which is a specialized facility designed specifically for firearm training, practice, or competitions in a safe and isolated space.

National Initiative for Cybersecurity Education (NICE) (2020) defined cyber range is a virtual environment which has the capacities of replicating the digital infrastructure (including hardware, applications and tools, or networks), in part or whole, of an organization to perform emulated cybersecurity training and testing. However, European Cyber

Security Organization (ECSO) (2020) considered that it is more precise to define cyber range as a platform with a set of core technologies for the development, delivery and use of interactive simulation environments and for the development of particular cyber range use cases. In any case, cyber range provides a virtual environment for learners experiencing and practicing cybersecurity defense against cyber-attacks in a safe manner. A cyber range has also the capacity of providing multiple scenarios as training exercises (Friedl et al., 2022).

Cyber ranges can be used for many purposes and are especially good for providing cybersecurity education and training security professionals. Some of the important use cases of a cyber range include security testing, security research, security education, development of cyber capabilities, development of cyber resilience, competence assessment, recruitment, and competitions. Depending on its intended uses, a cyber range may consist of some of these functionalities; Internet service simulation, attack simulation, user activity simulation, scenarios and content development, competency management, data collection and analysis, scoring and reporting, and instructor tools (ECSO, 2020).

3 The VPET Approach

Vocational and Professional Education and Training (VPET) tries to fill the gap of industrial expectation on graduates. In addition, VPET does not only provide an alternate path to, but broaden the learning opportunities for, both school leavers and in-service workers. To solve the problem of skilled talent shortage, it is desirable to provide cybersecurity education to young people and provide re-skilling opportunities to ICT practitioners with practical skill sets who will obtain not only academic qualification but also professional certification. To identify the essential skill sets for different roles associated with the necessary knowledge elements becomes a crucial task for course development.

The Workforce Framework for Cybersecurity (NICE Framework) (NIST, 2020) classifies cybersecurity workforce into 7 *Categories* of common cybersecurity functions; 33 distinct *Specialty Areas* of cybersecurity work; and 52 detailed grouping of cybersecurity work into *Work Roles*. A work role is divided into *Tasks* requiring specific *knowledge, skills, and abilities (KSAs)* to perform. The NICE framework suggests that cybersecurity education is multidisciplinary consisting of technical knowledge, managerial knowledge, data analytical skills, and other areas like sociology, and law. By using a defined set of knowledge and skills, learners can complete tasks described for a position (role). The set of knowledge and skills can be assessed to determine the competencies which allow education and training providers to be responsive to the needs of employers.

The European Union Agency for Cybersecurity (ENISA) (2022) similarly defined 12 role profiles in the European Cybersecurity Skills Framework. The deliverables and the main tasks are specified for each profile together with the required key skills, key knowledge and competences.

In order to understand the requirements on the skill set required by the industry, it is necessary for educational providers to work closely with industry and incorporate industrial advisory into the quality assurance mechanism. The industrial advisory team can provide advice on adopting the latest technologies, the required skill sets of various roles, the development of curriculum, etc., provide instructors and provide internship

opportunities to students. Internship and work place study is an alternate path to provide hands-on exercises to students.

In order to provide the necessary skills in cybersecurity education to fill the gap of skill shortage, it is impossible to conduct the training in traditional classrooms but in a controlled virtual environment provided by a cyber range. While employers expect graduates to have the necessary skills for the work roles, students need to learn transferable skills for future career advancement as technologies are changing rapidly. Blazic (2021) suggested that a balanced cybersecurity education should be provided with practical skills in enhancing the employability of students together with the foundations of cybersecurity science so that they are capable of updating their skills under the dynamic environment.

In addition, the curricula in cybersecurity education can be considered to incorporate contents of certain professional certifications so that students receiving VPET can attain both academic qualification and professional certification.

4 Cybersecurity Education at HKCT

4.1 Programme Design

A Higher Diploma (HD) programme in Cybersecurity (Testing and Compliance), in the Hong Kong College of Technology (HKCT), was developed in 2021 and started operation in 2022 after going through all the necessary quality assurance procedures. The HD programme, focusing on the security testing and compliance aspects in practice, aims to provide students with fundamental knowledge and skills in cybersecurity, and provide students with a near-realistic and controlled experimental environment to develop students' hands-on capacities, problem-solving skills, and self-learning ability; and thus, attempting to fill the gap of skill shortage at technician level.

Crumpler and Lewis (2019) discussed that cybersecurity skilled talents should process a broad range of specialty areas with certain essential knowledge sets and skills. The knowledge sets include an understanding of computer architecture, data, cryptography, networking, secure coding principles, operating system internals, Linux-based systems, low-level programming languages, etc. and the skills required to perform the tasks of penetration testing, secure system design, incident response, and tool development.

The design of the HD programme follows the multi-disciplinary approach and include modules covering generic skills, ICT fundamentals knowledge, specialized knowledge and skills in cybersecurity, and integration, at the corresponding technical level. The cybersecurity related courses are designed to have hands-on exercises associated with the corresponding knowledge elements.

4.2 Cyber Range Implementation

To provide a safe and controlled experimental platform for student practice, a cyber range was acquired and implemented in 2022 and named as the Heung Kong Cyber Range (HKCR). Katsantonis et al. (2023) discussed some of the weaknesses of CR implementation which include high preparation cost, high testing requirements, learning

strategy neglect, fixed workspace, ineffective assessment, and limited participants profiles.

These weaknesses were also encountered when selecting a suitable CR to support the student learning and research. Although the project was supported by a government funding, the allocated budget was insufficient to acquire a world-leading CR for sustainable operation. Instead the CR at HKCT was provided by a leading mainland supplier. The first task was to build in an English interface, in parallel to its existing Chinese interface, to the whole platform. The bilingual interface of the platform may facilitate learners with different background to have the same learning experience. Our technical team is now able to create and import new scenarios to the HKCR, in addition to the scenes coming with the CR.

The Heung Kong Cyber Range, after going through rigorous tests, is now connected to at least three cyber ranges from universities leading in the field in China. These connections with other CRs enable resource sharing in supporting teaching and learning among the universities.

4.3 Designing Hands-On Exercises

As supporting student learning is the main purpose of implementing the Heung Kong Cyber Range, how to provide hands-on exercises to support the teaching of the relevant knowledge in the curriculum becomes an urging task.

Yamin et al. (2020) considered that a cyber range is an environment in providing dedicated testbeds and infrastructures for executing the training scenarios, of which a scenario defines training requirements and the execution environment, the storyline, and the execution steps of a training exercise. Larrucea and Santamaría (2020) discussed the motivation in using cyber range as infrastructures for training purposes for both education and industry sectors. They adopted the updated taxonomy of a cyber range (Yamin et al. 2020) as a template in describing elements of an exercise on a cyber range. Since the cost on adoption and maintenance commercially available cyber ranges is too high, they proposed using free tools such as SonarQube for defining and running a cyber exercise in discovering vulnerabilities in a running application. Katsantonis et al. (2023) stated that educational CRs involve instructors, cybersecurity experts and learners. The instructor or the teacher needs to specify the scenario for the lab session and to provide instructions or hints to learners, while the security expert takes the role of scene operator to support the teacher by configuring the CR environment according to the scenario and performing other technical tasks such as injecting vulnerabilities to the scene.

In designing the hands-on exercises, it is desirable to establish a mechanism with a template for the teaching staff. In accordance with the learning outcomes, the teacher needs to identify the required skill related to each knowledge element in the curriculum, and identify what hands-on exercises are desirable for students to achieve the learning objectives. Since the teaching staff might not have the necessary skills and times working on the cyber range, and the technical support team might not have ideas on the curriculum, it is sensible to team them up to work out the practical exercises before the actual teaching starts.

Taking the module “Penetration Testing” as an example, there are four module intended learning outcomes (MILOs):

- 1) explain the purpose and describe the legal requirements of penetration testing;
- 2) describe the penetration testing process including reconnaissance, scanning, exploitation, post exploitation, and reporting;
- 3) use appropriate tools for performing penetration testing tasks; and
- 4) write penetration testing reports.

It is obvious that MILO 2 clearly states the knowledge requirements and MILO 3 specifies the corresponding skills requirements. When looking into the course content in more detail, the knowledge elements include at least setting up the testing environment, information gathering, exploitation, other attacks, post exploitation, and documentation. The teacher then identifies the knowledge elements and corresponding hands-on exercise to practice the required skill sets, as shown in Fig. 1.

<ol style="list-style-type: none"> 1. Introduction <ul style="list-style-type: none"> - Penetration testing and hacking - The penetration testing life cycle - Vulnerabilities and exploits - The history and culture of hackers 2. Setting up the testing environment <ul style="list-style-type: none"> - Introduction to Kali Linux - Getting comfortable with Kali Linux - Penetration testing tools 3. Information gathering <ul style="list-style-type: none"> - Open Source Intelligence (OSINT) - Port scanning - Enumeration 4. Exploitation <ul style="list-style-type: none"> - Finding vulnerabilities - Attack vectors and attack types - Metasploit basics - Client-side attacks - Social engineering - Penetration testing of Web application 	<ol style="list-style-type: none"> 5. Other attacks <ul style="list-style-type: none"> - Password attacks - Buffer overflows - Denial of services 6. Post exploitation <ul style="list-style-type: none"> - Maintaining access - Privilege escalation - Local information gathering - Pivoting and lateral movement 7. Documentation <ul style="list-style-type: none"> - The penetration test report - Presentation - Storage of evidence
--	--

Fig. 1. Course content with highlighted knowledge elements for hands-on exercises

<p>Hands-on exercise: skills covered</p> <p>Virtual environments</p> <ul style="list-style-type: none"> ▪ Operation environment ▪ Target environment (if any) <p>Objectives</p> <ul style="list-style-type: none"> ▪ MILOs to be achieved <p>Level</p> <p>Task description</p> <p>Steps</p> <p>Template</p>	<p>Hands-on exercise: Port scanning</p> <p>Virtual environments</p> <p>Operation environment</p> <ul style="list-style-type: none"> ▪ Acunetix - a web application security scanner running on Microsoft Windows platform and automatically looks for vulnerabilities such as SQL Injection and cross site scripting. This tool can scan websites or web applications using HTTP/HTTPS protocol, and others such as ASP.net, PHP, Java, Ruby. <p>Target environment</p> <ul style="list-style-type: none"> ▪ a website with vulnerabilities <p>Objectives</p> <ul style="list-style-type: none"> ▪ MILO 2: describe the penetration testing process including reconnaissance, scanning, exploitation, post exploitation, and reporting. ▪ MILO 3: use appropriate tools for performing penetration testing tasks. <p>Level: introductory</p> <p>Task description: find vulnerabilities of a website in the Reconnaissance phase</p> <p>Steps:</p> <ul style="list-style-type: none"> ▪ Installation of Web scanning tool WVS ▪ Tasks for web scanning with WVS ▪ WVS scan result analysis <p>An example</p>
--	---

Fig. 2. A template for specifying hands-on exercises with an example

A series of in-class exercises are to be designed on the cyber range by the support team. By specifying the items in the template, as shown in Fig. 2, the support team will

have a standardized format in describing the hands-on exercises, which may facilitate the communication between the teaching team and the support team.

In the actual laboratory session, the support team will need to set up the virtual environments and other necessary tools on the cyber range. Students are required to perform the task as specified on the worksheet in this controlled environment and to complete a report with data analysis and results discussion. The assessment of a student's ability to perform the technical task is based on the records obtained from the cyber range and quality of the lab report.

4.4 Packaging Skill-Based Units into Micro-credentials

Requirements on cybersecurity skills are changing rapidly with the technologies. To address the gap of skill shortage in this area, it is important to provide both school leavers receiving formal education and practitioners receiving skills re-training flexible ways to develop their knowledge, skills and competencies. According to the European Commission (2021), "*micro-credentials certify the learning outcomes of short-term learning experiences, for example a short course or training. A micro-credential is a short, competency-based recognition that allows an educator to demonstrate mastery in a particular area.*" In this sense, the module "Penetration Testing" can well be packaged into a micro-credential unit as it fills the skill gap in the specific area by defining the knowledge and skill to fulfil the specific learning outcomes with assessment to ascertain the competency of certain abilities to perform the task of penetration testing. When a learner obtaining a number of related micro-credential units in the same area, the learner may then be considered to have the abilities to work for certain position (role) as discussed in Sect. 3.

4.5 CR-Based Talent Cultivation

Collaboration with other universities was established in order to gain access to and learn from more experts in cybersecurity. To fully utilized the Heung Kong Cyber Range, a joint training program on gaining the practical ability of network security for Computer Science major undergraduate students with a university in Hong Kong.

Namin et al. (2016) suggested that competition-based training workshops are effective tools in simulating students' motivation in gaining cybersecurity knowledge where participants take part in competition in solving the challenging problems.

A "Cybersecurity Vocational Skills Competition" was jointly organized with the Hong Kong Computer Society in April 2023 with 77 participants, from both practitioners in related industry and college students, qualified into the final contest. The competition was composed of both knowledge examination and practical assessment, and was set according to the National Vocational Skills Standards for Network and Information Security Administrator issued by the Ministry of Human Resources and Social Security. In order to prepare participants for the competition, 60 h of initial trainings on relevant knowledge elements and practical skills using the HKCR were provided and opened to participants online. The competition was well recognized by the authority. The first five winners were issued a certificate obtaining the standard equivalent to a Network and Information Security Administrator (Advanced) Certificate from the Ministry of Industry

and Information Technology. Those participants achieving certain level on both technical knowledge and practical skills in the competition were issued a certificate at the middle level.

Based on the achievement from the “Cybersecurity Vocational Skills Competition”, a framework of “Network Security Talent Certification Standards” and the corresponding courses containing certain micro-credential units were proposed for accreditation under the qualification framework locally. The proposed certification standards, with reference to the National Vocational Skills Standards, enable students to obtain professional and academic qualifications from both China and Hong Kong.

5 Conclusion

There is a shortage of skilled talents in the field of cybersecurity at all levels globally. The shortage of the skilled professionals itself becomes a real threat to the society. While education and training aim to produce graduates to fill the job market, the industry complains continuously on their lack of practical skills.

In order to provide more hands-on exercises in obtaining the necessary skills required to perform certain roles, a controlled and virtual environment such as cyber range is adopted, in which students can play around safely. How to enhance traditional classroom education in cybersecurity supported by the use of a cyber range is a problem needed to be solved.

A VPET approach is adopted by grouping the knowledge element with the required skill sets together, designing hands-on exercises with a template specifying the learning objectives, the virtual environment, the task and the steps etc., and assessing the competency of a student by evaluating the performance records from the cyber range and the lab report. These units can be packaged into micro-credential units to provide a more flexible way of student learning or skill re-training. This VPET approach aims to make some contributions in closing the gap of skill shortage in cybersecurity.

Acknowledgement. The authors would like to thank the support from the cyber range vendor and its support team. This work is partially supported by the Enhancement and Start-up Grant Scheme for Self-financing Post-secondary Education (ESGS/2021/APP-19).

References

- AAG: The Latest Cyber Crime Statistics (updated Aug 2023). <https://aag-it.com/the-latest-cyber-crime-statistics/> (2023)
- Anderson, J.L.: Global cyberattacks increased 38% in 2022. <https://www.securitymagazine.com/articles/98810-global-cyberattacks-increased-38-in-2022> (2023)
- Blazic, B.J.: The cybersecurity labour shortage in Europe: moving to a new concept for education and training. *Technol. Soc.* **67**(2021), 10176 (2021)
- Caelli, W.J., Liu, V.: Cybersecurity education at formal university level: an Australian perspective. *J. Colloquium Inform. Syst. Secur. Educ. Editon* **5**(2), (2018)
- Check Point Software Technologies: Cyber Security Report 2023. <https://pages.checkpoint.com/cyber-security-report-2023.html> (2023)

- Chou, T.S.: Multi-Learning Techniques for Enhancing Student Engagement in Cybersecurity Education (2019). <https://doi.org/10.18260/1-2--33127>
- Conklin, W.A., Cline Jr. R.E., Roosa T.: Re-engineering cybersecurity education in the US: an analysis of the critical factors. In: 47th Hawaii International Conference on System Science, IEEE (2014). <https://doi.org/10.1109/HICSS.2014.254>
- Crumpler, W., Lewis, J.A.: The Cybersecurity Workforce Gap, Center for Strategic and International Studies. <https://www.csis.org/analysis/cybersecurity-workforce-gap> (2019). Retrieved on 1 September 2023
- De Leon, D.C., Goes, C.E., Haney, M.A., Krings, A.W.: ADLES: specifying, deploying, and sharing hands-on cyber-exercises. *Comput. Secur.* **74**, 12–40 (2018). <https://doi.org/10.1016/j.cose.2017.12.007>
- De Zan, T.: Mind the Gap: The Cyber Security Skills Shortage and Public Policy Interventions. University of Oxford, GCSEC – Global Cyber Security Center (2019)
- European Commission: Proposal for a COUNCIL RECOMMENDATION on a European approach to micro-credentials for lifelong learning and employability. <https://data.consilium.europa.eu/doc/document/ST-9237-2022-INIT/en/pdf> (2021)
- European Cyber Security Organization: Understanding cyber ranges: from hype to reality. Technical report (2020)
- European Union Agency for Cybersecurity (ENISA): European Cybersecurity Skills Framework Role Profiles (2022)
- Federal Bureau of Investigation (FBI): FBI Internet Crime Report 2022. https://www.ic3.gov/Media/PDF/AnnualReport/2022_IC3Report.pdf (2023)
- Friedl, S., Glas, M., Englbrecht, L., Böhm, F., Pernu, G.: ForCyRange: an educational IoT cyber range for live digital forensic. In: Drevin, L., et al. (eds.) WISE 2022, IFIP AICT 650, pp. 77–91 (2022)
- International Information System Security Certification Consortium (ISC)²: Cybersecurity Professionals Stand Up to a Pandemic. Cybersecurity Workforce Study 2020. <https://www.isc2.org/research/> (2020)
- International Information System Security Certification Consortium (ISC)²: Cybersecurity Workforce Study 2022. <https://www.isc2.org/research/> (2022)
- Katsantonis, M.N., et al.: Cyber range design framework for cyber security education and training. *Int. J. Inf. Secur.* **22**, 1005–1027 (2023). <https://doi.org/10.1007/s10207-023-00680-4>
- Khan, M.A., Merabet, A., Alkaabi, S., et al.: Game-based learning platform to enhance cybersecurity education. *Educ. Inf. Technol.* **27**, 5153–5177 (2022). <https://doi.org/10.1007/s10639-021-10807-6>
- Larrucea, X., Santamaría, I.: Designing a cyber range exercise for educational purposes. In: Yilmaz, M., Niemann, J., Clarke, P., Messnarz, R. (eds.) Systems, Software and Services Process Improvement: 27th European Conference, EuroSPI 2020, Düsseldorf, Germany, September 9–11, 2020, Proceedings, pp. 302–312. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-56441-4_22
- Mishra, S., Pinchot, J., Cellante, D., Pullet, K.: Student perceptions of skills and competencies needed for cybersecurity programs and careers. *Issues Inform. Syst.* **20**(2), 9–17 (2019)
- Mountroudidou, X., et al.: Securing the human. In: Proceedings of the Working Group Reports on Innovation and Technology in Computer Science Education, pp. 157–176 (2019). <https://doi.org/10.1145/3344429.3372507>
- Namin, A.S., Aguirre-Muñoz, Z., Jones, K.S.: Teaching cyber security through competition: an experience report about a participatory training workshop. In: Annual International Conference on Computer Science Education: Innovation & Technology, pp. 98–104 (2016)
- National Initiative for Cybersecurity Education (NICE): The cyber range: A guide. Technical report (2020)

- NIST: Workforce Framework for Cybersecurity (NICE Framework), NIST Special Publication 800-181 Revision 1 (2020). <https://doi.org/10.6028/NIST.SP.800-181r1>
- Rebecchi, F., et al.: A Digital Twin for the 5G Era: the SPIDER Cyber Range. In: IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), Belfast, United Kingdom, pp. 567–572 (2022)
- Sitnikova, E.: Developing postgraduate cyber security programs at the Australian Centre for Cyber Security at the University of New South Wales Canberra at ADFA. *J. Colloquium Inform. Syst. Secur. Educ. (CISSE)* **5**(2), 1–23 (2018)
- Triplett, W.J.: Addressing cybersecurity challenges in education. *Int. J. STEM Educ. Sustain.* **3**(1), 47–67 (2023)
- Tsado, L.: Cybersecurity Education: the need for a top-driven, multidisciplinary, school-wide approach. *J. Cybersecurity Educ. Res. Pract.* **2019**(1), 4 (2019)
- Venter, I.M., Blignaut, R.J., Renaud, K., Venter, M.A.: Cyber security education is as essential as “the three R’s.” *Heliyon* **5**(12), e02855 (2019)
- World Economic Forum: The cybersecurity skills gap is a real threat – here’s how to address it. <https://www.weforum.org/agenda/2023/05/the-cybersecurity-skills-gap-is-a-real-threat-heres-how-to-address-it/> (2023)
- Yamin, M.M., Katt, B., Gkioulos, V.: Cyber ranges and security testbeds: scenarios, functions, tools and architecture. *Comput. Secur.* **88**, 101636 (2020). <https://doi.org/10.1016/j.cose.2019.101636>

Online and Innovative Learning



Using Network Analysis to Explore the Factors Influencing Students' Online Learning Effectiveness

Sha Zhu¹, Qing Guo¹, Wei Qin¹, Shun Xu², and Harrison Hao Yang³(✉)

¹ Central China Normal University, Wuhan 430079, China

² Hubei University of Technology, Wuhan 430068, China

³ State University of New York at Oswego, Oswego 13126, USA

harrison.yang@oswego.edu

Abstract. Despite the common pursuit of students' online learning effectiveness, there is limited research on the factors influencing the effectiveness, particularly when considering the interplay of multiple factors. This study investigated the online learning experiences of 121,391 students in Wuhan during the COVID-19 period. A network analysis was used to reveal the impact of information literacy, online learning presence, and online learning engagement on online learning effectiveness. The study also identified the variations in the influence of these factors between student groups with high and low-effectiveness levels. The findings indicate the following: 1) Emotional engagement, information perception awareness, information application awareness, and information behavior positively contribute to online learning effectiveness; 2) Emotional engagement is associated with higher levels of strength, betweenness, closeness and expected influence in both groups. Information laws and regulations have stronger betweenness in the high-effectiveness group, while emotional engagement is stronger in the low-effectiveness group. The results inform recommendations for enhancing future online learning practices.

Keywords: online learning; information literacy · effectiveness · network analysis

1 Introduction

The global COVID-19 pandemic was widespread over the past few years, leading to billions of students transitioning to online learning (UNESCO, 2020). Despite the ongoing return to face-to-face learning, the educational reforms catalyzed by COVID-19 are likely to remain, with an ongoing evolution towards a blended approach of online and face-to-face learning. Despite the importance of online learning, studies have found that its effectiveness is less than ideal for most students (Chen et al., 2010). In an effort to enhance students' online learning effectiveness, numerous studies have investigated factors that could potentially affect the effectiveness of online learning.

Some studies have found that the maturity of online learning technologies (Lackmann et al., 2021; Özhan & Kocadere, 2020) is an objective factor influencing online learning effectiveness. Additionally, subjective factors related to students themselves, such as their abilities, attitudes, and cognition, including information literacy (Prior et al., 2016) and online learning engagement (Yang et al., 2018), have been found to be significant influencing factors on students' online learning effectiveness. It is widely believed that teachers cannot control the technology; however, factors at the student level can help teachers understand how to improve online learning. Therefore, researchers recommend focusing on student-level factors, rather than other external factors, to provide tangible guidance for effective online teaching (Marks et al., 2005).

However, previous studies have some limitations with respect to student-level factors. First, studies often selectively examine the impact of a few or even a single key factor on online learning effectiveness, neglecting the interplay of multiple variables (Hu & Li, 2017). Second, previous empirical studies on online learning often have had small sample sizes; these small samples may result in selection bias from non-representative samples, making it difficult to generalize the research findings (Yang, 2018). Finally, previous research has been unevenly distributed across educational levels, with a primary focus on undergraduate and graduate students. Few studies have been conducted with middle school students (Yang, 2018).

To address the research limitations described above, this study conducted a large-scale survey of middle school students' online learning experiences during the COVID-19 period, with a sample of 121,391 students. This study applied network analysis methods, treating multiple factors as a systemic whole and constructing them into an interconnected network. The goal was to explore how multiple factors interactively influence their online learning effectiveness. Specifically, this study focused on the following research questions. (1) For all students, what are the key influential factors in the overall network of online learning effectiveness? (2) For students with different levels of online learning effectiveness, what are the differences in the factors influencing the network? This study will provide a new perspective of systemic view for the research of influential factors of online learning effectiveness. Meanwhile, the results of the study will provide instructional guidelines for students' online learning.

2 Review of Factors Influencing Online Learning Effectiveness

Few studies have investigated the systematic effects of multiple factors involved in online learning effectiveness, but they have validated student-level factors that may influence such effectiveness. These factors are described as follows.

2.1 The Dependent Variable

Online learning effectiveness is the dependent variable of this study. Online learning effectiveness refers to the knowledge and skills acquired by students during a period of online learning. Previous studies have mostly used students' academic performance (PC) or satisfaction (SA) as the main indicators to measure online learning effectiveness (Wei & Chou, 2020; Guo et al., 2023). PC primarily reflects student achievement in terms

of knowledge and skills, with less emphasis on assessing emotional goals. SA reflects students' emotional experiences in online learning, but does not focus on the mastery of knowledge and skills. Therefore, this study applied both SC and SA as indicators to measure learners' online learning outcomes.

2.2 Information Literacy and Online Learning Effectiveness

The concept of information literacy has attracted widespread attention since it was first proposed by Zurkowski (1974). The meaning of the term has also continuously expanded as time has passed, and as technology has developed and advanced. Today, information literacy is considered an essential and comprehensive form of literacy for individuals to survive in an increasingly intelligent world. Our previous research proposed a framework for student information literacy, with four dimensions: information awareness and attitude, information knowledge and skills, information thinking and behavior, and information social responsibility. These dimensions have nine specific indicators: information perception awareness (A1), information application awareness (A2), information security awareness (A3), information scientific knowledge (B1), information application skills (B2), information thinking (C1), information behavior (C2), information ethics (D1), and information laws and regulations (D2) (Author, 2020). Information literacy is an important factor for engaging students in successful online learning. Students with a high level of information literacy efficiently access, organize, and critically use information. This stimulates active online learning behaviors and generates positive online learning experiences (Buchanan et al., 2016; Prior et al., 2016).

2.3 Online Learning Presence and Online Learning Effectiveness

The concept of online learning presence originates from the Community of Inquiry (CoI) model. The model assumes that meaningful learning experiences are generated through interaction between students and teachers in online learning communities (Wertz, 2022). The online learning presence is created by developing three interdependent elements: cognitive presence (CP), teaching presence (TP), and social presence (SP) (Shea & Bidjerano, 2009). CP refers to students' ability to explore, integrate, and solve problems in the online environment. TP refers to students' experience with respect to the teachers' selection, organization, presentation, and design of course content and learning activities. SP refers to the state whereby students map their personal characteristics onto the online learning environment, presenting themselves as authentic individuals to other learners (Garrison et al., 2010). Online learning presence significantly impacts students' online learning effectiveness. A higher level of online learning presence indicates that students are likely to engage in more constructive online learning behaviors and achieve higher online learning effectiveness (Shea & Bidjerano, 2009, Zhang et al., 2016).

2.4 Online Learning Engagement and Online Learning Effectiveness

“Online learning engagement” refers to the sustained positive state shown by students as they actively participate in learning activities and interact with others in an online learning

environment (Yang et al., 2018). Online learning engagement encompasses behavioral engagement (BE), cognitive engagement (CE), affective engagement (AE), and social engagement (SE) (Fredricks et al., 2004; Rimm-Kaufman et al., 2015; Bae et al., 2020). BE refers to specific behavioral manifestations during the online learning process; CE primarily refers to the use of learning strategies in the online learning environment; AE mainly represents students' emotional experiences during online learning; and SE refers to students' interpersonal interactions with teachers during online learning (Hu & Li, 2017; Hoi & Hang, 2021). Researchers have observed that positive online learning engagement effectively promotes online learning practices and enhances online learning effectiveness (Lin et al., 2019).

3 Method

Detailed descriptions of the participants and their screening process, the utilized instruments, and the data analysis approach are provided below.

3.1 Participants and Procedure

Due to COVID-19, Wuhan was one of the first cities to implement large-scale online learning, resulting in almost every student having extensive experience with the approach. This makes those students an appropriate research sample. With the support of the Wuhan Education Bureau, this study explored the topic of online learning among middle school students across the city, using a cluster sampling method. The survey distribution approach was organized from top to bottom, with the Education Bureau first distributing the survey notice and online survey link to each district's education bureau. Seventh and eighth-grade students from 234 middle schools in 14 districts of Wuhan were informed about the purpose of the study and were requested to participate. Subsequently, each district's education bureau issued specific survey notices to the selected schools within their jurisdiction. Finally, each selected school organized seventh and eighth-grade students to complete the survey at the school computer room. All participants were requested to complete the entire questionnaire within 40 min, and the data was stored in CSV format. After removing duplicate and missing values, a total of 121,391 questionnaires were received by the study team, including responses from 62,919 males (51.83%) and 58,472 females (48.17%). The sample included 65,346 students in the seventh-grade (53.83%) and 56,045 students in the eighth-grade (46.17%).

3.2 Instruments

To assess online learning effectiveness, this study measured students' SA and PC with respect to online learning. The SA scale was adapted from Kuo et al. (2014) and consisted of three items. The reliability analysis indicated a Cronbach's alpha coefficient of 0.909, a composite reliability (CR) value of 0.910, and an average variance extracted (AVE) value of 0.771, meeting the fit criteria (Fornell and Larcker, 1981). PC was assessed based on the average scores for three main subjects during the students' online learning period: Chinese, Mathematics, and English. The scores were grouped into five levels:

“Above 110” was assigned a score of 5, “100–109” was assigned a score of 4, “90–99” was assigned a score of 3, “80–89” was assigned a score of 2, and “Below 80” was assigned a score of 1. This is because the majority of the students in this study scored above 80 and many teachers regard 80 as a passing line.

To assess information literacy, this study utilized a standardized test for middle school students, designed in a previous study (Zhu et al., 2021). The content validity of the test was verified using the Delphi method, in which multiple rounds of expert consultations were conducted to seek feedback about the consistency of the test items and their respective information literacy indicators. The final test consisted of 50 multiple-choice questions, covering the 9 indicators of information literacy. Each indicator was scored based on the number of correctly answered questions.

To measure online learning presence, this study adapted a scale developed by Law et al. (2019), with three sub-scales: social presence, cognitive presence, and teaching presence. The Cronbach’s alpha coefficients for these three sub-scales were 0.895, 0.937, and 0.927 respectively; the CR values were 0.895, 0.937, and 0.927, respectively; and the AVE values were 0.631, 0.680, and 0.717, respectively. The results of the analysis indicated good questionnaire quality.

To measure online learning engagement, the study applied a scale developed by Hoi et al. (2021), with four sub-scales: behavioral engagement, cognitive engagement, affective engagement, and social engagement. The Cronbach’s alpha coefficients for these four sub-scales were 0.907, 0.923, 0.930, and 0.904 respectively; the CR values were 0.908, 0.924, 0.930, and 0.904, respectively; and the AVE values were 0.711, 0.752, 0.769, and 0.703 respectively. The results of the analysis showed good questionnaire quality.

All the scales listed above used a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The scales were translated from English to Chinese. Two graduate students performed parallel translations of the scales under committee coordination. After the translation was completed, an educational technology expert with more than 20 years of teaching experience in both the United States and China conducted a bilingual evaluation of the translated items. Feedback was used to adjust the wording of several items, enhancing the overall readability of the scales.

3.3 Data Analysis

To explore the interactive mechanism of multiple factors, this study adopted a network analysis approach to construct a complex network of factors related to online learning effectiveness. Network analysis refers to the construction of a network that includes variables and their interconnections to present system characteristics, examine the relationships between elements, identify key elements, and guide interventions.

For research question 1, the MGM package in R was first used to construct an overall network of factors influencing online learning effectiveness for all participants. Non-parametric bootstrap calculations (1,000 iterations) were performed to validate the robustness and accuracy of the network structure. The characteristics of the links in the overall network were qualitatively analyzed to identify the influencing factors. Centrality indices, including strength, betweenness, closeness, and expected influence, were calculated to quantitatively measure the importance of each factor. The strength

represented the influence of a node in the network; betweenness indicated the number of shortest paths that pass through a node, reflecting the node's role as an intermediary; closeness represented the degree to which a node is in a central position in the network; and expected influence was the influence of a node after excluding negative effects.

For research question 2, K-means clustering was first conducted using SPSS 24.0 to divide students into two groups (high-effectiveness and low-effectiveness) based on their satisfaction with online learning and academic scores. Then, the MGM package in R was used to construct separate networks of influencing factors for the two groups. The NCT package was applied to analyze the overall differences between the networks of the two groups. The centrality indices of the nodes in each group's network were calculated, and further comparisons were made to determine the key differences in influencing factors between the two groups.

4 Results

The results of the network analysis of the factors influencing online learning effectiveness, as well as the results of the network comparisons between the high- and low-effectiveness groups, are reported below.

4.1 Overall Analysis of Networks of Influential Factors

Figure 1 shows the overall network of factors influencing students' online learning effectiveness, generated using a network analysis of data from all participants. The network consists of 18 interconnected nodes, with each node representing an influencing factor. The surrounding circles around the nodes indicate the degree to which the node can be explained by other network nodes. Thicker lines indicate stronger correlations between nodes. Green lines represent positive correlations between nodes. Figure 2 shows the results of the network stability analysis. A smaller degree of overlap between the red and black lines in Fig. 2 (left), and a higher number of black squares in Fig. 2 (right), indicate greater differences between the connections and stronger stability. Figure 2 indicates that the network constructed in this study is highly stable, indicating the network analysis results are highly reliable.

An analysis of the overall network shows that information literacy, online learning presence, and online learning engagement are all positively correlated with students' online learning effectiveness. Specifically, SA has strong positive correlations with AE and PC; and SC has strong positive correlations with A1, A2, and C2. Figure 3 shows the centrality measures of each node. AE, SA, and SC have high levels of strength, betweenness, closeness, and expected influence in the network. This indicates that AE has a strong influence in the network, occupying a central position and playing an important mediating role in the mutual influence of other nodes. Additionally, CP and CE have a high level of strength in the network, but are not directly connected to online learning effectiveness (SA and PC).

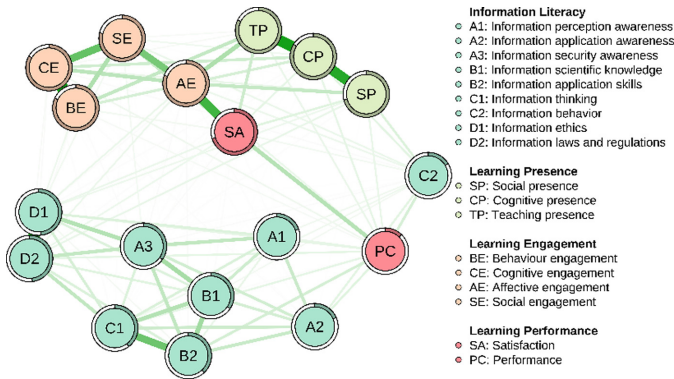


Fig. 1. Overall network of factors influencing online learning effectiveness

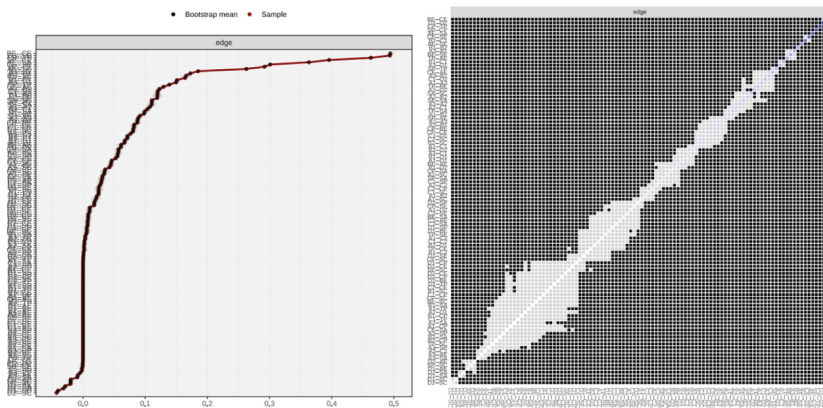


Fig. 2. Stability assessment of the network

4.2 Network Comparison of High- and Low-Effectiveness Groups

Figure 4 shows the networks of factors influencing online learning effectiveness for high- and low-effectiveness groups. The overall network difference test results indicate significant differences in the network structure ($M = 0.097$, $p < 0.001$) and overall network strength ($S = 0.1180$, $p < 0.001$; 7.829 for the high-effectiveness group and 7.947 for the low-effectiveness group). This indicates that the network of factors influencing online learning effectiveness is comparable when assessing the high- and low-effectiveness groups.

Figure 5 shows the centrality indices of the nodes in the networks for the high- and low-effectiveness groups. It shows that the strength, closeness, and expected influence of each factor in the two networks are similar. Of these, AE has higher levels of strength, closeness, and expected influence in both the high- and low-effectiveness groups. In terms of betweenness, C1, D1, and D2 show stronger intermediacy in the high-effectiveness group. In contrast, SE and AE show stronger intermediacy in the low-effectiveness group.

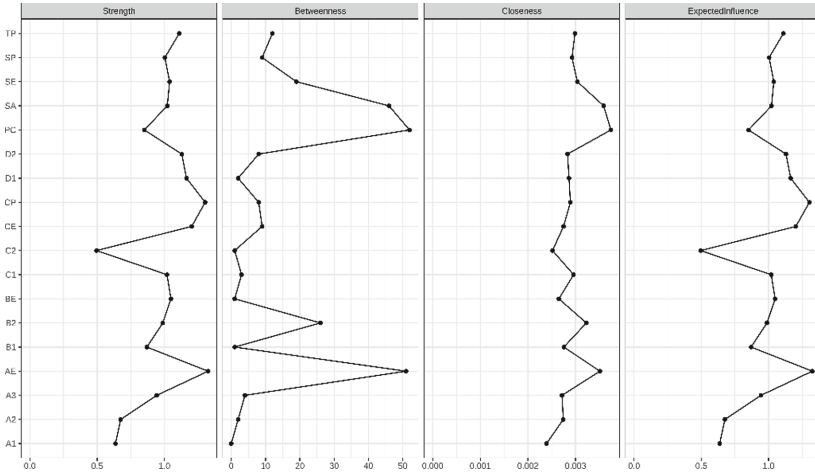


Fig. 3. Centrality indices for each node in the network. The abbreviations of the factors in the figure are the same as in Fig. 1.

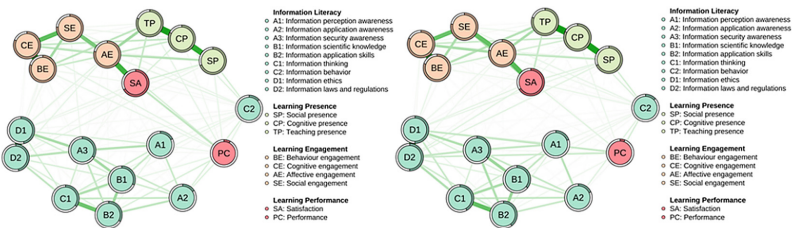


Fig. 4. Networks of influencing factors for high- (right) and low- (left) effectiveness groups

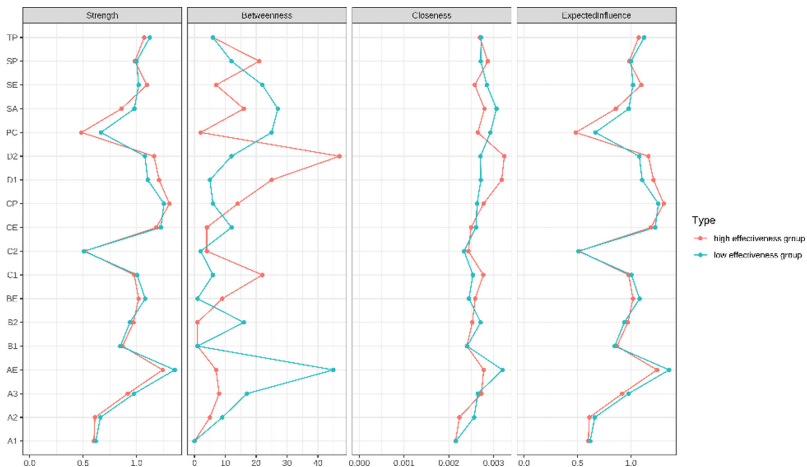


Fig. 5. Centrality indices for each node in the high- and low-effectiveness group networks. The abbreviations of the factors in the figure are the same as in Fig. 1.

5 Discussion and Conclusion

This study comprehensively investigated the factors influencing online learning effectiveness for middle school students. It systematically explained the interactive effects between these factors and the effectiveness of online learning using network analysis methods. The results inform meaningful conclusions that provide guidance for enhancing online learning effectiveness. The following discussion focuses on the most important findings and recommendations.

5.1 Overall Analysis of Networks of Influential Factors

In the overall network, AE is a key node that positively enhances students' SA and PC; it also plays an important mediating role. Through AE, factors related to online learning presence and engagement both positively impact SA and SC. This finding confirms the results of previous studies (Zhan & Mei, 2013; Doo & Bonk, 2020). When students engage in online learning, they often have limited opportunities for meaningful interaction with peers and teachers, which may result in learner isolation and passive learning (Banna et al., 2015). However, a higher level of AE may help students effectively reduce feelings of loneliness, enhance their satisfaction, and consequently improve the effectiveness of online learning (Martin and Bolliger, 2018).

Additionally, CP does not directly influence online learning effectiveness; however, it has a significant influence within the network and can indirectly affect online learning effectiveness through important mediating factors such as AE. Previous studies found that students with higher CP are more likely to focus on online learning behaviors, increase their engagement in learning activities, and actively participate in online learning (Doo & Bonk, 2020). This ultimately promotes online learning effectiveness. Information literacy also exerts a strong positive influence, especially with respect to A1, A2, and C2. Previous studies found that an abundance of information can easily lead to information overload and confusion for students (Avcı & Ergün, 2022). Having a high level of information literacy enables students to effectively cope with the challenges of online learning, adapt to the online learning environment, and enhance online learning effectiveness (Avcı & Ergün, 2022).

5.2 Overall Analysis of Networks of Influential Factors

Comparing the network structure of the high- and low-effectiveness groups revealed that AE had a strong influence in both groups, which echoes the results derived from the overall network. This result reaffirms that the emotional acceptance of online learning, and being able to actively participate, are key to achieving positive online learning outcomes (Martin and Bolliger, 2018). Information literacy factors, including C1, D1, and D2, showed strong betweenness in the high-effectiveness group; these factors were less mediated in the low-effectiveness group. This result may be because students in the high-effectiveness group were able to effectively and rationally acquire and apply information when participating in the online learning process. This may generate a stronger sense of self-efficacy and a willingness to take on self-directed learning tasks in the online learning environment.

In addition, students in the high-effectiveness group may be able to transform problems into informational activities to solve them, thereby achieving excellent online learning effectiveness (Prior et al., 2016). Online learning engagement, including SE and AE, are important mediators for students in the low-effectiveness group. This may be because students in the low-effectiveness group are less capable of independent learning and lack the ability to apply information to solve problems (Zhu et al., 2020). This highlights the importance of them being able to actively participate in online learning, which should affect the rest of their online learning experience and influence their online learning effectiveness.

5.3 Implications

This study highlights both theoretical and practical implications. From a theoretical perspective, the study advances a method influencing factor inquiry, by analyzing the interactions among multiple factors using network analysis. This idea informs new conclusions about the mechanisms that influence students' online learning effectiveness, and provides methodological inspiration for research in other fields.

From a practice perspective, study findings point to recommendations for improving students' online learning practices. First, online learning should stimulate students' affective engagement, to enhance satisfaction with online learning and promote learning effectiveness. Two effective approaches include creating authentic learning situations that are close to real life, and designing challenging learning tasks to stimulate student interest and motivation. Second, it is important to cultivate students' information literacy, given the important mediating role of information literacy in enhancing online learning effectiveness. Specifically, students should be provided with convenient online learning tools and clear learning guidelines and supports, thus helping them to develop the awareness, ability and habits related to finding, processing and using information. In particular, focus should be placed on helping low-effectiveness students develop awareness related to solving information problems, to enhance their online learning efficacy and promote learning effectiveness.

5.4 Limitations

As with any research, this study has some limitations, which highlight opportunities for future research. First, this study primarily utilized a cross-sectional research design to collect data, which may not lead to an exact causal relationship. Second, this study focused on three sets of factors based on previous research experience: information literacy, online learning presence, and online learning engagement. However, the factors affecting students' online learning effectiveness are complex and diverse, and there may be factors not considered in this study. In the future, a longitudinal research design could further examine the relationship of factors influencing online learning effectiveness. Meanwhile, future research should select additional factors to explore the mechanism of their influence on students' online learning effectiveness.

Acknowledgments. The work was supported by a grant from the National Natural Science Foundation of China (No. 62107019); Key Projects in Provincial Philosophy and Social Sciences (No. 22D043); Teaching and Research Project of Hubei University of Technology (No. 2022057).

References

- Avcı, Ü., Ergün, E.: Online students' LMS activities and their effect on engagement, information literacy and academic performance. *Interact. Learn. Environ.* **30**(1), 71–84 (2022)
- Bae, C.L., DeBusk-Lane, M.L., Lester, A.M.: Engagement profiles of elementary students in urban schools. *Contemp. Educ. Psychol.* **62**, 1–13 (2020)
- Banna, J., Lin, M.F.G., Stewart, M., Fialkowski, M.K.: Interaction matters: strategies to promote engaged learning in an online introductory nutrition course. *J. Online Learn. Teach./Merlot.* **11**(2), 249 (2015)
- Buchanan, S.M.C., Harlan, M.A., Bruce, C., Edwards, S.: Inquiry based learning models, information literacy, and student engagement: a literature review. *Sch. Libr. Worldw.* **22**(2), 23–39 (2016)
- Chen, P.D., Lambert, A.D., Guidry, K.R.: Engaging online learners: the impact of web-based learning technology on student engagement. *Comput. Educ.* **54**(4), 1222–1232 (2010)
- Doo, M.Y., Bonk, C.J.: The effects of self-efficacy, self-regulation and social presence on learning engagement in a large university class using flipped Learning. *J. Comput. Assist. Learn.* **36**(6), 997–1010 (2020)
- Fornell, C., Larcker, D.F.: Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* **18**(1), 39–50 (1981)
- Fredricks, J.A., Blumenfeld, P.C., Paris, A.H.: School engagement: potential of the concept, state of the evidence. *Rev. Educ. Res.* **74**(1), 59–109 (2004)
- Guo, Q., Zeng, Q., Zhang, L.: What social factors influence learners' continuous intention in online learning? A social presence perspective. *Inf. Technol. People* **36**(3), 1076–1094 (2023)
- Hoi, V.N., Le Hang, H.: The structure of student engagement in online learning: a bi-factor exploratory structural equation modelling approach. *J. Comput. Assist. Learn.* **37**(4), 1141–1153 (2021)
- Hu, M., Li, H.: Student engagement in online learning: a review. In: 2017 International Symposium on Educational Technology (ISET), pp. 39–43. IEEE (2017)
- Kuo, Y.C., Walker, A.E., Schroder, K.E., Belland, B.R.: Interaction, Internet self-efficacy, and self-regulated learning as predictors of student satisfaction in online education courses. *Internet Higher Educ.* **20**, 35–50 (2014)
- Lackmann, S., Léger, P.M., Charland, P., Aubé, C., Talbot, J.: The influence of video format on engagement and performance in online learning. *Brain Sci.* **11**(2), 128 (2021)
- Law, K.M., Geng, S., Li, T.: Student enrollment, motivation and learning performance in a blended learning environment: the mediating effects of social, teaching, and cognitive presence. *Comput. Educ.* **136**, 1–12 (2019)
- Lin, L. C., Hung, I. C., Kinshuk, Chen, N.S.: The impact of student engagement on learning outcomes in a cyber-flipped course. *Educ. Technol. Res. Dev.* **67**, 1573–1591 (2019)
- Marks, R.B., Sibley, S.D., Arbaugh, J.B.: A structural equation model of predictors for effective online learning. *J. Manag. Educ.* **29**(4), 531–563 (2005)
- Martin, F., Bolliger, D.U.: Engagement matters: student perceptions on the importance of engagement strategies in the online learning environment. *Online Learn.* **22**(1), 205–222 (2018)
- Özhan, ŞÇ., Kocadere, S.A.: The effects of flow, emotional engagement, and motivation on success in a gamified online learning environment. *J. Educ. Comput. Res.* **57**(8), 2006–2031 (2020)
- Prior, D.D., Mazanov, J., Meacheam, D., Heaslip, G., Hanson, J.: Attitude, digital literacy and self efficacy: flow-on effects for online learning behavior. *Internet High. Educ.* **29**, 91–97 (2016)
- Rimm-Kaufman, S.E., Baroody, A.E., Larsen, R.A., Curby, T.W., Abry, T.: To what extent do teacher-student interaction quality and student gender contribute to fifth graders' engagement in mathematics learning? *J. Educ. Psychol.* **107**(1), 1–17 (2015)

- Shea, P., Bidjerano, T.: Cognitive presence and online learner engagement: a cluster analysis of the community of inquiry framework. *J. Comput. High. Educ.* **21**(3), 199–217 (2009)
- UNESCO.: Education: From disruption to recovery UNESCO. <https://en.unesco.org/covid19/education%20response> (2020). Retrieved 24 May 2020
- Wei, H.C., Chou, C.: Online learning performance and satisfaction: do perceptions and readiness matter? *Distance Educ.* **41**(1), 48–69 (2020)
- Wertz, R.E.: Learning presence within the Community of Inquiry framework: an alternative measurement survey for a four-factor model. *Internet Higher Educ.* **52**, 100832 (2022)
- Yang, D., Lavonen, J.M., Niemi, H.: Online learning engagement: factors and results-evidence from literature. *Themes eLearn.* **11**(1), 1–22 (2018)
- Zhan, Z., Mei, H.: Academic self-concept and social presence in face-to-face and online learning: perceptions and effects on students' learning achievement and satisfaction across environments. *Comput. Educ.* **69**, 131–138 (2013)
- Zhang, H., Lin, L., Zhan, Y., Ren, Y.: The impact of teaching presence on online engagement behaviors. *J. Educ. Comput. Res.* **54**(7), 887–900 (2016)
- Zhu, S., Wu, D., Yang, H.H., Li, Y., MacLeod, J.: Understanding K-12 students' information literacy in informal learning environments: a literature review. *Int. J. Innov. Learn.* **27**(4), 432–449 (2020)
- Zhu, S., Yang, H.H., Wu, D., Chen, F.: Investigating the relationship between information literacy and social media competence among university students. *J. Educ. Comput. Res.* **59**(7), 1425–1449 (2021)
- Zurkowski, P.G.: The information service environment: Relationships and priorities (Related Paper No. 5). National Commission on Libraries and Information Science (1974)



Learning Image Processing Contents Through Game

Suriya Natsupakpong^(✉) and Suwitchaphon Suksawang

Institute of Field Robotics, King Mongkut's University of Technology Thonburi,
126 Pracha Uthit Rd, Bang Mot, Thung Khru, Bangkok, Thailand
suriya.nat@kmutt.ac.th

Abstract. This research is design and development an educational game for learners to gain knowledge in image processing topics while playing the game. The game is created to combine educational content with fun playing elements and enhance strategic thinking skills. This makes it an effective tool for both formal and informal learning. The user assumes the role of a robot named Enigma in the secret town and has five missions to complete within that place and return safely. The missions consist of the sampling and quantization quest, the light and color adjustment quest, the depth of field quest, the hue color space quest, and the lens distortion and camera calibration quest. Each mission contains information and knowledge related to the mission stored in the user's journey book, which the user can open and review lessons at any time while playing the game. The game is designed in first- and third-person perspective in a 3D environment and developed using Unity3D on Windows operating system. The research trial tested 10 users playing the game for an average of 60 min. Each user takes the quiz before and after playing the game. Out of a full score of 20, the average scores before and after learning from this game were 11.30 points and 16.70 points, respectively. The experimental results show that users gained knowledge of image processing from this game. The experiment also evaluates the usability in terms of usefulness, ease of use, ease of learning, and satisfaction with the average scores of 4.60, 4.15, 4.73, and 4.43, respectively.

Keywords: learning · educational game · student engagement · image processing

1 Introduction

There are various learning styles that help learners access the content and skills they need. The popular framework is VARK model, which developed by Niel Fleming in the late 1980s. The term "VARK" stands for Visual, Auditory, Reading/Writing, and Kinesthetic, which represent the four main sensory modalities through which people tend to learn. An educational game is a type of game designed with the primary purpose of promoting learning, skill development, and educational goals. These games are intentionally created to combine entertainment and educational content, providing players with an engaging and interactive way to gain new knowledge, skills, and perspectives. Incorporating the VARK learning styles into educational games can improve the learning experience by

accommodating various preferences and engaging students in various ways. An image processing course typically focuses on teaching the fundamental concepts, techniques, and tools used in the analysis, manipulation, enhancement, and interpretation of digital images. Learners learn through video media as a lecture-style teaching in the knowledge section and practice the programming skills by doing assignments. Learners are unable to remember and understand all theories and contents with lecture-based learning. Integrating games into context solves this problem. This research focuses on the impact of learning in the image processing contents using educational games.

In this research, we suggest an image processing teaching game to improve the user learning experience. The remainder of the paper is structured as follows. Section 2 examines the relevant literature. The research methodology is detailed thoroughly in Sect. 3. Section 4 presents the experimental results. This is followed by concluding observations in Sect. 5.

2 Literature Review

Digital games are increasingly being used as teaching and learning tools because they are fun and engaging (Plass et al., 2019). The purpose of educational games or game-based learning is to give learners an interactive and interesting way to study and acquire specific skills or a specific learning outcome (Noemí & Máximo, 2014). Gamification, on the other hand, refers to learning processes or motivation-boosting activities to enhance the user experience of interactive systems through game design (Seaborn & Fels, 2015). Learners who completed the gamified experience or game-based learning got better scores in practical assignments and in overall score (Domínguez et al., 2013; Liu et al., 2020). A comprehensive research initiative examining how students respond to game-based learning within programming courses has been conducted (Mathew et al., 2019; Zhao et al., 2022). The findings indicate that the integration of these games into the teaching and learning process has yielded positive outcomes. Moreover, a game-based learning strategy may be beneficial in helping students' acquisition of 21st century skills (Qian & Clark, 2016). In engineering education, the game-based learning using virtual and augmented reality is also studied to promote the experimental learning theory, which suggests utilizing experiences as the foundational elements for observing, reflecting, engaging in abstract conceptualization, and constructing knowledge (Vaz De Carvalho, 2019). In industrial training, game-based learning is utilized to practice assembling work pieces with hand gestures in a virtual environment without the assistance of an expert (Natsupakpong & Jailungka, 2020).

3 Research Methodology

3.1 System Overview Design and Development

The overall workflow of the system will begin with the main login page. After logging in with a personal code, the system will display scores and progress through different levels. When the player clicks to start the game, they will enter a storyline where they can choose the content they want to learn. The player will control the character's movement in a 3D format using the keyboard and mouse. The game will involve solving puzzles designed to be related to the content of image processing.

3.2 Story and Quest Design

The contents and puzzles of the game are related to basic image processing knowledge, lens fundamentals, and image enhancement aligned with the learning outcomes of the basic image processing course. Through gameplay, players gained knowledge and developed strategic thinking skills. The player takes on the role of robot named Enigmo in the secret town and has five missions to complete within that place and return safely. In the first section, it begins with puzzles related to basic image processing knowledge. The initial scene involves player finding the way into the secret town, passing through two puzzle rooms (Fig. 1) as follows:

Image Puzzle: This is played like a question-and-answer game where players must press the correct image button indicated and pass through a concealed door located behind the image. The content used here involves Sampling and Quantization. Questions involve selecting images with different pixel sizes and with specified intensity values.

Maze Puzzle: This section demonstrates the impact of image processing through a player's perspective. Camera views incorporate Unity's post-processing effects, including Grain, Vignette, and Color Grading, to manipulate the player's visual perception. Items throughout the maze improve visibility when interacted with them. By adjusting post-processing effects, players experience altered perspectives and enhanced visibility. The content used here involves the light and color adjustment.

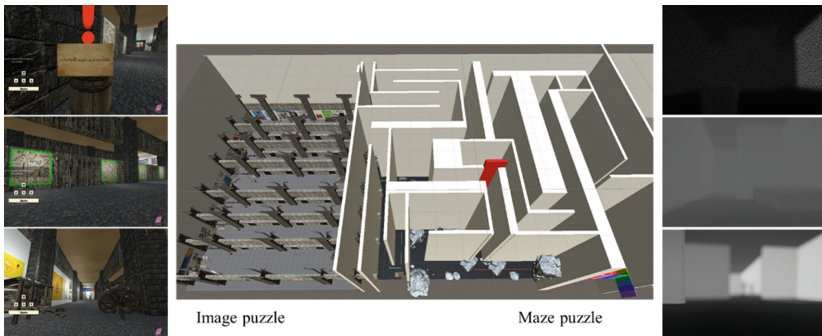


Fig. 1. Image puzzle and maze puzzle scenes.

There will be three more puzzles when the player reach in secret town (Fig. 2) as follows:

Blurred Image Puzzle: Players must focus on stone columns at various distances to decipher passwords in sharp clarity. This puzzle educates players about camera focusing and image blurring. Using Unity's Depth of Field post-processing effect, players can modify focal length values using mouse scrolling. A number on each stone column indicates the focal point. Incorrect focus results in a blurry image and obscured number, emphasizing the need for accurate focusing.

Color Match Puzzle: Six signboards with changing colors appear in a sequence. Players match the sequence by pressing corresponding color buttons. The game adjusts the

player’s vision, initially restricting color perception. Collecting items enhances color vision, making this puzzle solvable. Players need to memorize the color sequence and press the buttons in the correct order to open the exit door.

Lens Puzzle: This puzzle focuses on adjusting lenses for optimal photography by teaching lens distortion and Camera Calibration concepts.

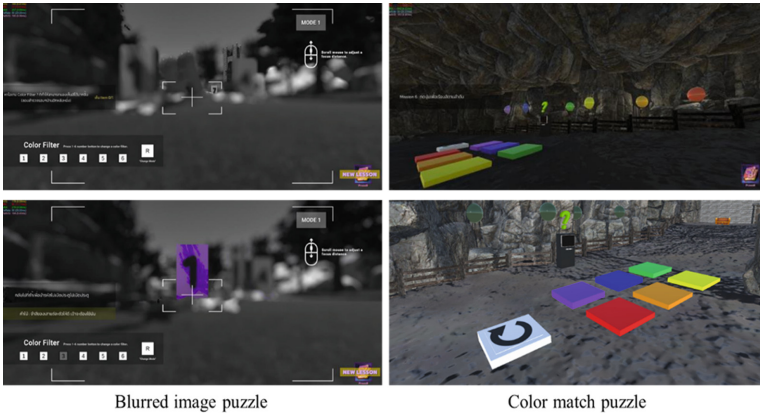


Fig. 2. Blurred image puzzle and color match puzzle scenes.

3.3 Scene Design and Character Controller

The created scenes are divided into two parts: the user interface and the 3D game environment. The user interface system was designed in scene to significantly improve gaming and assist players to learn and grasp the game ideas depicted in Fig. 3. For example, the “New Lesson” notice appears when a new lesson connected to a quest is available, and the user can examine that lesson by pressing the “B” button. By pressing the “E” button, the user can interact with the game item, which displays a message when the user gets close to it. Missions and suggestions are also displayed to help players progress through the game.

The 3D game environment includes the decorated terrain, vegetation, structures, and scene components. There are four scenes in totals shown in Fig. 4 as follows:

Starting Scene: This functions as a lobby for selecting the next scene. Players can either enter the tutorial mode to learn the game mechanics or access the main learning content mode.

Forest Village Scene: This scene prepares players for the main learning content. Players begin their journey through the forest to find their way to the house that leads to the secret town.

Small House Interior Scene: This scene immerses players into the main storyline with puzzles intertwined with image processing content, ranging from easy to challenging. There are two puzzle contents within the room.

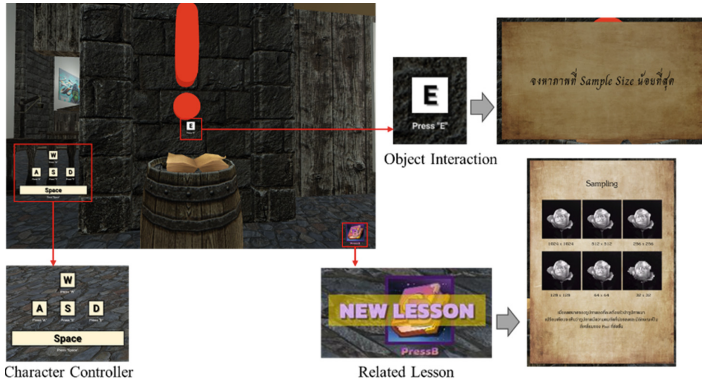


Fig. 3. Example of user interface used in the game.

Secret Town Scene: This scene adds more image processing content. Puzzles are divided into three parts, with increased exploration based on the storyline.



Fig. 4. Example of scene design.

For character control system, players can control the game using the keyboard and mouse. Basic buttons “W,” “A,” “S,” and “D” are used for moving forward, turning left, moving backward, and turning right, respectively. The “Spacebar” is used for jumping. Additionally, there are extra buttons like “E” used for collecting items based on the game’s status. The game also features various camera angles. The default view is a third-person perspective, but in some scenes, it switches to a first-person view. At times, the camera angle adjusts to different points in the scene to add variety to the game’s dimensions.

3.4 Game Development and Testing

The game was designed in first- and third-person perspective in a 3D environment and was created with Unity3D on Windows. The game was tested with 10 users ranging from

high school to bachelor's degree students. The testing process comprised three steps as follows:

Pre-Learning Phase: Players received an introduction to the academic content of the game for learning about image processing. Basic game information was explained before entering the game. Each game consisted of four image puzzle challenges and one maze challenge where players had to collect a total of five items. In the blurry image puzzle, there were two sub-challenges, while the color capture puzzle had one challenge, resulting in a total of eight challenges. Players needed to pass each challenge from the first to the last. Subsequently, players took a pre-test before delving into the in-game learning content about image processing.

Learning Phase: Upon entering the game, players encountered the main menu where they could either start playing the game or exit. After starting the game, players entered the tutorial level with content guiding them to find the correct house. Hints for the task were shown on the left side of the user interface, and explanations for using keyboard buttons were given at the bottom of the hint. In the first tutorial level, players became familiar with the environment, mechanics, and the first-person perspective before engaging in the actual content. The tutorial section wasn't overly challenging but depended on how quickly players could interpret the task and find the correct solution. Once players found the correct house, they entered the planned content. In this content, players engaged in strategic thinking by understanding the goal from the task and hints given through the user interface. Then they analyzed the task to find the correct answer before proceeding to the next challenge. Strategic thinking and planning were crucial for players to complete each challenge as swiftly as possible. After executing the challenges, players evaluated their results for the most accurate adjustments. Throughout the game, players were given opportunities to try both incorrect and correct approaches, but they couldn't progress solely through guessing. Each challenge had specific content, which led to new information being added through the user interface for players. After completing the game, players and developers discussed the game's content according to each level's objectives. Additional content, feedback, errors, and suggestions for improvement were summarized for future adjustments.

Post-Learning Phase: After completing gameplay, players took a post-test to assess the effectiveness of their learning from the in-game content related to image processing.

4 Results and Discussion

The pre- and post-test performance assessment average scores were 11.30 and 16.70, respectively, out of a total of 20. It illustrates that players who engaged in learning through the educational game on image processing exhibited statistically significant improvement in their learning outcomes. The experimental results show that the average time for the user to complete the game was 60 min, and the evaluation of usability was collected from the satisfaction assessment in 5 levels from score 1 to 5 with the meanings of strongly disagree, disagree, neutral, agree, and strongly agree, respectively (Likert, 1932). Table 1 shows how the evaluation criteria are interpreted based on the average level. The score

interval is computed by subtracting the highest and lowest scores and dividing by the number of levels. The usability evaluation is separated into four categories: usefulness, ease of use, ease of learning, and satisfaction, and is based on the USE questionnaire (Lund, 2001). Table 2 displays the question and evaluation outcomes. Figure 5 shows a bar graph of average scores with standard distribution values for each question.

The interesting points show that all users agree that the educational game can help user learn the image processing content. Half of the users can play the game without a manual. However, when users need to know how to play something, gaming instructions should be provided. During the testing of the game, several issues were identified. The display was not consistent on screens with varying resolutions, leading to distorted visuals. The lack of brightness adjustment on the screen affected visibility of details in darker areas. Additionally, there were some errors within the game in specific areas. There are user suggestions regarding enhancing the learning experience within the game. These suggestions include the ability to review previously learned content, teleporting to previously visited points, saving game progress for later play, and enhancing the overall experience with additional sound effects. Thus, the game requires further development and testing to ensure accuracy and completeness.

Table 1. Evaluation Criteria.

Score Interval	Meaning
4.21 – 5.00	Strongly agree
3.41 – 4.20	Agree
2.61 – 3.40	Neutral
1.81 – 2.60	Disagree
1.00 – 1.80	Strongly disagree

The evaluation results for usability in Fig. 6 show that the average scores for ease of learning, usefulness, and satisfaction were at a strongly agree level with a score of 4.73, 4.60, and 4.43, respectively. The average score for ease of use was at an agree level, with scores of 4.15. The study had significant drawbacks, including a small sample size and a short amount of time between pre- and post-test. It can be enhanced in future research to make the experimental results better.

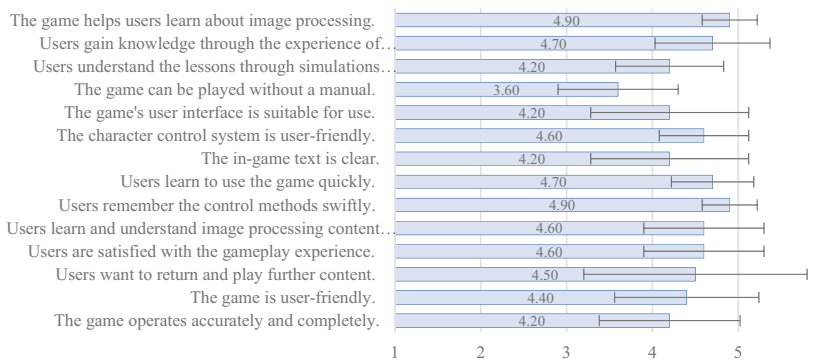
Table 2. Questions and Evaluation Results.

n = 10		Strongly agree n (%)	Agree n (%)	Neutral n (%)	Disagree n (%)	Strongly disagree n (%)	Mean	SD
Usefulness	The game helps users learn about image processing	9 (90)	1 (10)	0 (0)	0 (0)	0 (0)	4.90	0.32
	Users gain knowledge through the experience of playing the game	8 (80)	1 (10)	1 (1)	0 (0)	0 (0)	4.70	0.67
	Users understand the lessons through simulations within the game	3 (30)	6 (60)	1 (1)	0 (0)	0 (0)	4.20	0.63
Ease of use	The game can be played without a manual	1 (10)	4 (40)	5 (5)	0 (0)	0 (0)	3.60	0.70
	The game's user interface is suitable for use	4 (40)	5 (50)	0 (0)	1 (10)	0 (0)	4.20	0.92
	The character control system is user-friendly	6 (60)	4 (40)	0 (0)	0 (0)	0 (0)	4.60	0.52
	The in-game text is clear	4 (40)	5 (50)	0 (0)	1 (10)	0 (0)	4.20	0.92
Ease of learning	Users learn to use the game quickly	7 (70)	3 (30)	0 (0)	0 (0)	0 (0)	4.70	0.48

(continued)

Table 2. (continued)

n = 10		Strongly agree n (%)	Agree n (%)	Neutral n (%)	Disagree n (%)	Strongly disagree n (%)	Mean	SD
	Users remember the control methods swiftly	9 (90)	1 (10)	0 (0)	0 (0)	0 (0)	4.90	0.32
	Users learn and understand image processing content rapidly	7 (70)	2 (20)	1 (10)	0 (0)	0 (0)	4.60	0.70
Satisfaction	Users are satisfied with the gameplay experience	7 (70)	2 (20)	1 (10)	0 (0)	0 (0)	4.60	0.70
	Users want to return and play further content	8 (80)	1 (10)	0 (0)	0 (0)	1 (10)	4.50	1.30
	The game is user-friendly	6 (60)	2 (20)	2 (20)	0 (0)	0 (0)	4.40	0.84
	The game operates accurately and completely	5 (50)	2 (20)	3 (30)	0 (0)	0 (0)	4.20	0.82

**Fig. 5.** The Evaluation Results of Usability: Usefulness, Ease of Use, Ease of Learning, and Satisfaction.

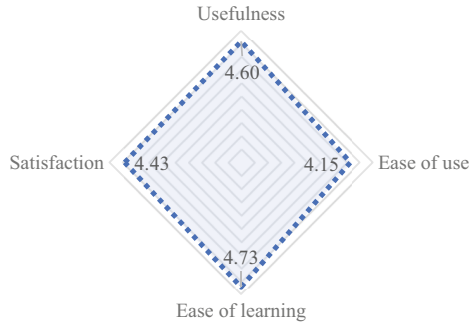


Fig. 6. The Average Score of Evaluation Results in 4 Areas of USE Questionnaire.

5 Conclusion

This research involves the design and development of an educational game focused on teaching the concepts of image processing. The primary objective is to facilitate knowledge acquisition through gameplay, enhancing the players' understanding of lesson content. The game encompasses various topics, including Sampling and Quantization Levels, Basic Image Enhancement involving light and color, Hue Color space, Depth of Field, as well as Lens Distortion and Camera Calibration. These topics are integrated into every level of the game. The game was tested with a group of 10 users, each spending an average of around 60 min playing. The performance assessments were conducted using quiz before and after playing game, and the average scores were 11.30 and 16.70, respectively. This indicates that the learning game contributes to improved retention and comprehension of the content. Moreover, the usability with the gameplay experience was evaluated. The results of the assessment revealed that the average scores for ease of learning, usefulness, and satisfaction were all at a highly agree level, with scores of 4.73, 4.60, and 4.43, respectively. The average score for ease of use was 4.15.

For future research directions, there are several enhancements that can be considered. First, increasing the number of questions in each level could provide a more comprehensive learning experience. This could involve randomizing and shuffling the order of questions and answer choices, allowing players to encounter new challenges when revisiting the game. Furthermore, improving the clarity of scoring mechanisms could enhance the players' understanding of their performance. Clear and detailed scoring systems can help players track their progress and motivate them to strive for better results. Adding additional audio elements to the game, such as sound effects and background music, can contribute to a more immersive and engaging experience for players. This auditory dimension could enhance the overall gameplay and make it more captivating.

Acknowledgements. This research project is supported by King Mongkut's University of Technology Thonburi (KMUTT), Thailand Science Research and Innovation (TSRI) and National Science, Research and Innovation Fund (NSRF) in the Fundamental Fund, Fiscal Year 2023.

References

- Domínguez, A., Saenz-De-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., Martínez-Herráiz, J.J.: Gamifying learning experiences: practical implications and outcomes. *Comput. Educ.* **63**, 380–392 (2013)
- Likert, R.: A technique for the measurement of attitudes. *Arch. Psychol.* **140**, 5–55 (1932)
- Liu, Z.-Y., Liu, Z.-Y., Shaikh, Z., Gazizova, F.: Using the concept of game-based learning in education. *Int. J. Emerg. Technol. Learn.* **15**(14), 53–64 (2020)
- Lund, A.: Measuring usability with the USE questionnaire. *Usability User Experience Newslett. STC Usability SIG.* **8**(2), 3–6 (2001)
- Mathew, R., Malik, S.I., Tawafak, R.M.: Teaching problem solving skills using an educational game in a computer programming course. *Inform. Educ.* **18**(2), 359–373 (2019)
- Natsupakpong, S., Jailungka, P.: Part assembly learning by using hand gestures with leap motion. *Int. J. Technol. Learn.* **27**(1), 35–45 (2020)
- Noemí, P.-M., Máximo, S.H.: Educational games for learning. *Univ. J. Educ. Res.* **2**(3), 230–238 (2014)
- Plass, J.L., Mayer, R.E., Homer, B.D.: *Handbook of Game-based Learning*. The MIT Press (2019)
- Qian, M., Clark, K.R.: Game-based Learning and 21st century skills: a review of recent research. *Comput. Hum. Behav.* **63**, 50–58 (2016)
- Seaborn, K., Fels, D.I.: Gamification in theory and action: a survey. *Int. J. Hum Comput Stud.* **74**, 14–31 (2015)
- Vaz De Carvalho, C.: Virtual Experiential Learning in Engineering Education. In: *Proceedings – Frontiers in Education Conference, FIE, 2019-October* (2019)
- Zhao, D., Muntean, C.H., Chis, A.E., Rozinaj, G., Muntean, G.M.: Game-based learning: enhancing student experience, knowledge gain, and usability in higher education programming courses. *IEEE Trans. Educ.* **65**(4), 502–513 (2022)



Student Satisfaction in External Online Course Provider: The Way Forward for WOU?

Wen Huey Kan^(✉), Li Lian Yap, and Yee Hui Hoo

School of Business Administration, Wawasan Open University, 54, Jalan Sultan Ahmad Shah,
10050 Geogetwon, Penang, Malaysia
whkan@wou.edu.my

Abstract. Education profoundly impacts life quality, highlighting the importance of a strong education system. This study explores student satisfaction with external online course providers in distance learning. It focuses on cognitive beliefs, technology, and social context, rooted in Bandura's Social Cognitive Theory.

During the COVID-19 pandemic, institutions like Wawasan Open University (WOU) rapidly shifted to e-learning, ensuring continuity and well-being. The rising demand for online education requires providers to offer engaging courses.

This study quantitatively analyzes factors influencing student satisfaction and contentment with online courses. Data collection utilized a cross-sectional online survey integrated into WOU's Learning Management System. Structural Equation Modeling (SEM) provided deep insights.

Findings emphasize the significance of student contentment with course material, impacting career development and coaching services. Skill-building and degree integration in online learning align with labor market needs. WOU can enhance programs through iterative collaboration with online course providers.

In essence, understanding and improving student satisfaction with online course providers elevates education quality and empowers students. Innovations and student-centered strategies are crucial for distance learning institutions like WOU in the dynamic online education landscape. This research holds profound implications for distance learning's future and the role of online course providers.

Keywords: Interactive Learning · Distance Education · External Online Course Provide

1 Introduction

In response to the global COVID-19 pandemic, educational institutions have undergone unprecedented shifts in instructional approaches. Wawasan Open University (WOU), like many others, has swiftly transitioned to online learning, ensuring education's continuity while prioritizing safety. However, as the demand for online education surges, institutions such as WOU encounter new challenges in sustaining student satisfaction and delivering quality distance learning experiences through external online course providers.

Student satisfaction holds a pivotal role in the efficacy of any educational system, impacting academic engagement, motivation, and learning outcomes. In the dynamic

landscape of online education, understanding the determinants of student satisfaction becomes imperative for improving educational policies and teaching methods.

This study delves into distance learning to explore student satisfaction with external online course providers.

Moreover, the study examines Interactive Learning's link to Student Satisfaction, focusing on students' active engagement, participation, and collaboration in online courses. Understanding this relationship is pivotal for effective online learning strategies that prioritize engagement and positive experiences.

By investigating these dimensions of student satisfaction with external online course providers, this research contributes insights to online education. Findings will guide WOU's student-centric online learning experiences and benefit other institutions navigating the competitive landscape of international online education.

In conclusion, unraveling elements impacting student satisfaction is crucial as the global education system embraces distance learning. The subsequent sections will detail the methodology used to explore the relationships between key factors and Student Satisfaction. The study aims to illuminate Wawasan Open University's path and enable institutions to thrive in the evolving online education era, nurturing students' academic and professional growth through exceptional distance learning experiences.

1.1 Problem Statement

The global COVID-19 pandemic has necessitated a rapid shift towards online learning in educational institutions, including the esteemed Wawasan Open University (WOU). As distance learning becomes the new norm, ensuring student satisfaction with external online course providers has emerged as a crucial challenge for WOU. To enhance the overall quality of education and meet the demands of the dynamic online education market, it is imperative to investigate the factors influencing student satisfaction in this context.

This study aims to explore the impact of Computer Self-efficacy, System Functionality, and Content Feature on students' perception of Interactive Learning, and subsequently, how Interactive Learning influences Student Satisfaction. Specifically, the research will examine the role of Computer Self-efficacy in students' confidence and competence in using digital technology for educational purposes. It will assess students' satisfaction with the functionality of the online learning system (System Functionality) and their perception of the quality and relevance of the online course content (Content Feature).

Furthermore, the study will investigate the extent to which Interactive Learning, which encompasses active participation, engagement, and collaboration in the online courses, contributes to students' overall satisfaction with the external online course provider. Understanding the interplay of these variables can provide valuable insights into designing effective and student-centric online learning experiences.

By addressing these research objectives, the study seeks to inform Wawasan Open University's strategies for improving the online learning environment, promoting student engagement, and enhancing satisfaction with external online course providers. The findings will be instrumental in guiding educational institutions' efforts to thrive in the competitive landscape of international online education, fostering students' academic

and professional advancement, and elevating the overall quality of distance learning experiences.

1.2 Research Objectives

1. **Identify the Factors Influencing Student Satisfaction:** The primary objective of this study is to identify and examine the key factors that influence student satisfaction with external online course providers in a distance learning environment. By comprehensively analyzing factors such as learners' cognitive beliefs, technological environment, and social setting, the study aims to shed light on the multifaceted elements that contribute to students' contentment with online education.
2. **Assess Student Satisfaction with Online Course Content:** The second objective of this research is to assess student satisfaction specifically in relation to the content of online courses offered by external providers. This assessment encompasses an exploration of the quality, relevance, and engagement value of the course materials. By evaluating students' perspectives on the content feature, the study seeks to gain insights into how the design and delivery of course content impact their overall satisfaction and learning experiences.

2 Literature Review

In the rapidly evolving landscape of modern education, the question of student satisfaction with online learning experiences has assumed paramount importance due to its far-reaching implications for educational quality and effectiveness. The rapid transition to distance learning, catalyzed by the global COVID-19 pandemic, has thrust into the spotlight the need to comprehensively understand the multifaceted factors that contribute to students' contentment with external online course providers. This literature review embarks on a comprehensive exploration of recent research endeavors, delving into the intricate determinants of student satisfaction within the complex context of online education. By delving into the nuances of this pivotal area, this synthesis aims to offer a comprehensive understanding that aligns seamlessly with the research objectives of the present study.

2.1 Factors Influencing Student Satisfaction

Within the intricate tapestry of factors that shape student satisfaction, certain determinants have emerged as critical pillars in the endeavor to create effective online learning environments. Amidst these, cognitive beliefs stand out as central determinants, encompassing the crucial notions of self-efficacy and performance expectations (Bandura, 1986). Self-efficacy, representing an individual's perception of their own competence to accomplish tasks, holds a formidable influence over students' motivation, engagement levels, and overall satisfaction with online courses (Richardson et al., 2012). As the investigation delves deeper, the technological milieu, encapsulating the functionality of systems and attributes of content, has ascended to prominence as a vital determinant of students' experiential dimensions (Wang & Newlin, 2002). Platforms boasting

user-friendly and intuitive interfaces possess the potential to elevate satisfaction levels by mitigating technological-related frustrations, thereby fostering an environment conducive to effective learning (Wu & Hiltz, 2004). Furthermore, the intricate interplay of the social fabric, defined by interactions and a nurturing climate, significantly amplifies the sense of belonging and community, exerting a profound influence on student satisfaction (Rovai, 2002).

2.2 Assessment of Course Content and Satisfaction

At the core of the discussion lies the quality and pertinence of online course content, playing a pivotal role in shaping students' perceptions of their educational journey. Importantly, students' assessments of content structure, clarity, interactivity, and alignment with learning objectives wield direct implications for their engagement and overall satisfaction (Ali & Leeds, 2009). Moreover, the potency of content catering to diverse learning modalities and fostering experiential richness is profound, substantially contributing to an augmented sense of satisfaction (Hsu & Wang, 2015). Notably, a discernible thread links student satisfaction with course content to their overarching contentment with the trajectory of learning (Chiu et al., 2017).

2.3 Synthesis and Research Gap

While the academic arena has witnessed an abundance of studies scrutinizing the multifaceted factors influencing student satisfaction within the realm of online education, a conspicuous gap persists in the nuanced comprehension of how these very factors manifest within the distinctive context of external online course providers. The precise interplay and manifestation of cognitive beliefs, technological environment, and social milieu in shaping student satisfaction within this specific context remain a realm ripe for exploration. Addressing this conspicuous gap assumes paramount importance, not solely in refining the landscape of distance learning but also in substantially enhancing the holistic experiential dimensions of students.

2.4 The Fundamental Theory

The study is grounded in Bandura's Social Cognitive Theory (Bandura, 1986), which guides the exploration of student satisfaction with external online courses in distance learning. This well-established theory's ability to predict human behavior and facilitate behavior change is recognized. The study extends its applicability to the unique context of distance learning, identifying three key factors that shape student satisfaction:

1. Cognitive Beliefs (Self-efficacy and Performance Expectations): Students' self-efficacy influences satisfaction, impacting confidence in using tools, engaging with materials, and managing online learning.
2. Technological Environment (System Functionality and Content Features): The study highlights how online platforms affect satisfaction, from usability to content features.
3. Social Environment (Interaction and Learning Climate): Virtual interactions and community creation influence satisfaction.

Incorporating these factors—cognitive beliefs, technological environment, and social context—into the study’s framework provides a comprehensive understanding of student satisfaction in distance learning. By leveraging Bandura’s Social Cognitive Theory and applying it to the digital realm, the study offers insights to enhance online education quality and align it with student needs and expectations.

3 Methodology

3.1 Research Design

This study adopts a quantitative research design to explore student satisfaction with external online course providers in the context of distance learning. The quantitative approach facilitates systematic data collection and analysis, offering valuable insights into the factors influencing student satisfaction and content-related contentment with online courses. By employing structured surveys and standardized measurements, this method allows for objective assessment of variables such as cognitive beliefs, technological environments, and content quality. The quantitative framework enhances replicability and generalizability, contributing to a growing body of research in online education. Overall, this approach effectively investigates the complex dynamics of student satisfaction within the digital learning landscape.

3.2 Sample Size and Participants

Central to the success of any empirical study is the judicious selection of participants and an adequately sized sample that mirrors the diversity and complexity of the population under scrutiny. This study’s participants constitute students enrolled at the esteemed Wawasan Open University (WOU), representing those who have traversed the contours of external online courses on their academic journey. Employing a cross-sectional survey methodology, the study casts its net widely to encompass students across varied academic programs and disciplines. The determination of an appropriate sample size is guided by the principles of representativeness and inclusivity, ensuring that the findings gleaned resonate with the broader tapestry of student experiences.

3.3 Data Collection

In the wake of the unprecedented global upheaval catalyzed by the COVID-19 pandemic, the study adroitly navigates the challenges of data collection through innovative digital means. At the heart of this endeavor is the meticulously designed structured questionnaire, meticulously calibrated to capture the intricate nuances of student satisfaction. This instrument is seamlessly woven into WOU’s Learning Management System (LMS), harnessing the power of technology to administer the survey via a user-friendly online Google Form. This strategic integration not only addresses the practical limitations posed by the pandemic but also embraces the accessibility and convenience that technology affords in the realm of data collection.

3.4 Research Instrument

The structured questionnaire stands as the conduit through which the voices of the participants find resonance and representation in the study's empirical narrative. A labor of methodological precision, this instrument is thoughtfully nurtured by weaving together threads from a tapestry of sources. It draws insights from the rich corpus of existing literature on student satisfaction, finds theoretical resonance in Bandura's Social Cognitive Theory, and synchronizes with the contemporary pulse of online education research. The instrument, characterized by a judicious amalgamation of closed-ended and Likert-scale questions, navigates the multifaceted spectrum of variables. It delves into students' perceptions of course content quality, the dynamic tapestry of interactions with instructors and peers, the scaffolding of self-efficacy beliefs, the terrain of performance expectations, the terrain of system functionality, and the kaleidoscope of content features.

3.5 Data Analysis Technique

The collected data is analyzed using Structural Equation Modeling (SEM), a robust statistical technique. SEM allows for the examination of latent variables and their relationships, providing a comprehensive understanding of the complex interplay among the identified factors influencing student satisfaction. The technique facilitates the assessment of direct and indirect effects, enabling the identification of the most significant determinants of student satisfaction in the online learning context.

4 Findings

The research endeavors focused on unraveling the intricate tapestry of student satisfaction with external online course providers in the context of distance learning. Through meticulous investigation, several key findings have come to light, shedding valuable insights on the factors, beliefs, and environments that shape students' contentment and engagement within this educational landscape.

The study's first objective, centered around identifying the factors influencing student satisfaction, unearthed a trio of critical dimensions. Cognitive beliefs emerged as influential cornerstones, with self-efficacy and performance expectations significantly impacting students' overall satisfaction. Those who held higher self-efficacy beliefs exhibited greater motivation and engagement, leading to enhanced contentment with their online courses. The technological environment showcased its prominence by exerting a substantial impact on students' satisfaction. The usability and functionality of online learning platforms emerged as key determinants, with user-friendly interfaces contributing positively to students' learning experiences. The social environment's role was equally profound, with interaction and a positive learning climate fostering a sense of community and belonging, thereby influencing students' satisfaction.

Delving into the second research objective, the assessment of student satisfaction with online course content underscored the centrality of content quality. Students' perceptions of content relevance, organization, and alignment with learning objectives played a decisive role in shaping their overall satisfaction. Courses featuring content that

catered to diverse learning styles and facilitated meaningful interactions garnered higher satisfaction ratings. Moreover, the study revealed a notable correlation between student satisfaction with course content and their overall satisfaction with the entire learning experience, underscoring the pivotal role of content in shaping students' perceptions.

In addition to these dimensions, the Structural Equation Modeling (SEM) analysis results reinforced and enriched the understanding of the relationships among the identified factors. The SEM analysis provided a deeper insight into how cognitive beliefs, technological functionality, social engagement, and course content quality interrelate to collectively influence student satisfaction. The analysis revealed direct and indirect effects, highlighting the intricate pathways through which these factors impact students' contentment and engagement within the realm of online education.

In synthesis, the research findings, coupled with the SEM analysis results, illuminate a comprehensive picture of student satisfaction within the realm of external online course providers. The interplay of cognitive beliefs, technological functionality, social engagement, and course content quality emerges as a multifaceted tapestry that defines the student experience. These insights offer educational institutions, such as Wawasan Open University (WOU), a roadmap for enhancing their online education offerings. By optimizing technology, fostering social interaction, and curating high-quality content, institutions can cultivate an environment that not only meets but surpasses students' expectations. The research findings and SEM analysis results contribute to the ongoing discourse on distance learning practices, offering actionable insights that can shape the future of online education and elevate the overall quality of the student experience.

5 Research Outcomes

The culmination of this comprehensive study on student satisfaction with external online course providers in the context of distance learning has yielded valuable insights with far-reaching implications for the field of online education. Through an exploration of the multifaceted factors that influence student contentment, the research has provided nuanced findings that underscore the significance of various dimensions in shaping the online learning experience.

First and foremost, the investigation into the factors influencing student satisfaction has revealed that cognitive beliefs, including self-efficacy and performance expectations, play a pivotal role in determining students' overall contentment. Students who possess higher levels of self-efficacy tend to exhibit greater motivation, engagement, and satisfaction with their online courses. Additionally, the technological environment emerged as a critical determinant of satisfaction. Students' perceptions of the usability and functionality of the online learning platform significantly impact their learning experiences and satisfaction levels. Furthermore, the social environment, characterized by interaction and a supportive learning climate, has been identified as a crucial component influencing students' sense of belonging and overall satisfaction within the online learning community.

In line with the second research objective, the assessment of student satisfaction with online course content has illuminated the importance of high-quality and engaging course materials. Students who perceive the content as relevant, organized, and aligned with

their learning objectives exhibit higher levels of satisfaction with the courses. Notably, satisfaction with course content has been found to correlate positively with overall satisfaction with the learning experience, highlighting the central role of content in shaping students' perceptions of the course.

Overall, the research outcomes suggest that a holistic approach to student satisfaction involves a careful balance between cognitive beliefs, technological functionality, social engagement, and course content quality. These findings have significant implications for educational institutions, including Wawasan Open University (WOU), as they navigate the dynamic landscape of online education. Leveraging these insights, institutions can refine their course design, technological platforms, and social interactions to enhance student satisfaction, promote effective learning, and better align educational offerings with the demands of the contemporary online education market. The research outcomes thus contribute to the ongoing evolution of distance learning practices and the continuous improvement of the student experience within the realm of online education.

5.1 Limitations and Future Directions

There are limitations to consider in interpreting the study's findings. The use of an online survey within Wawasan Open University's Learning Management System for data collection might have introduced sample bias, potentially excluding students with limited technology access and limiting the representation of diverse perspectives. The cross-sectional design used offers a snapshot rather than establishing causal relationships between variables. Future longitudinal research could provide insights into changing student satisfaction over time and the effects of interventions. Additionally, while efforts were made to ensure self-report measures' validity for assessing student perceptions, inherent response biases and subjectivity could impact data accuracy.

Looking ahead, several research avenues could enhance our understanding of online student satisfaction. Longitudinal studies could show how factors like cognitive beliefs, technological environment, social interactions, and content quality evolve over time, shaping overall satisfaction. Employing mixed-methods, combining quantitative data with qualitative insights, could provide a comprehensive understanding of satisfaction drivers. Exploring experimental designs would assess intervention effectiveness and establish causal links between factors and satisfaction. Comparative studies comparing satisfaction in online courses to traditional settings could offer broader insights. Investigating contextual influences like cultural variations could inform tailored strategies. Acknowledging these potential research directions can amplify this study's impact, refining online education and enhancing the student experience.

5.2 Conclusion

In the rapidly evolving landscape of modern education, where online learning has taken center stage, this study delves into the intricate realm of student satisfaction with external online course providers in distance learning. By exploring dimensions like cognitive beliefs, technological environments, social interactions, and course content quality, this research offers valuable insights for both academic institutions and the broader landscape of online education.

The findings highlight the crucial role of cognitive beliefs, especially self-efficacy and performance expectations, in shaping overall student satisfaction. Higher self-efficacy correlates with increased engagement, motivation, and contentment with online courses. The technological environment's importance is evident, as user-friendly platforms positively impact the learning experience. Similarly, meaningful interactions and a supportive learning environment positively influence students' sense of belonging.

These insights have significant implications for institutions like Wawasan Open University (WOU) navigating the online education landscape. By strategically leveraging cognitive beliefs, technological functionality, social engagement, and content quality, institutions can create a tailored learning environment that meets diverse student needs. This research provides guidance for refining online education, fostering interactions, and optimizing course content delivery.

Amidst global educational transformations, this study's insights guide educators, administrators, and policymakers toward strategies that enhance the student experience. By embracing innovation, prioritizing student-centered design, and promoting continuous improvement, institutions position themselves at the forefront of online education, addressing the evolving needs of learners. Ultimately, this study contributes to distance learning discourse and lays the groundwork for an educational future that transcends boundaries, enriches lives, and nurtures holistic growth.

References

- Al-Atabi, M., Warsi, A.: Investigating the critical factors influencing students' satisfaction in online courses during the COVID-19 pandemic. *Educ. Inform. Technol.* 1–18 (2020)
- Al-Quzwini, O.H., Dennis, C., Hegazy, A.I.: Factors influencing university students' behavioural intention to adopt online learning: a study from a developing country. *J. Mark. High. Educ.* **29**(2), 225–250 (2019)
- Ali, A., Leeds, E.M.: A framework for assessing satisfaction and learning in a web-based information system. *J. Comput. Inform. Syst.* **50**(1), 121–128 (2009)
- Bandura, A.: *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall (1986)
- Bates, A.W.: *Teaching in a Digital Age: Guidelines for Designing Teaching and Learning*. Tony Bates Associates (2015)
- Brown, T.A.: *Confirmatory Factor Analysis for Applied Research*, 2nd edn. The Guilford Press (2015)
- Byrne, B.M.: *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*, 3rd edn. Routledge (2016)
- Cavanaugh, C.S.: Student achievement in elementary and high school. In: *Online learning: Best Practices for K-12 Online Programs*, pp. 63–86. Lawrence Erlbaum Associates (2005)
- Chiu, C.M., Hsu, M.H., Sun, S.Y., Lin, T.C., Sun, P.C.: Factors affecting online nursing course effectiveness: a qualitative research. *Nurse Educ. Today* **50**, 64–70 (2017)
- Chen, P.S.D., Wu, H.C.: Learner satisfaction and learning effectiveness: a comparison of web-based and 2D simulation-based immersive environments. *Br. J. Edu. Technol.* **46**(1), 122–130 (2015)
- El Alaoui, R.N., Miftah, S.: Exploring factors affecting students' satisfaction in online learning environments during the COVID-19 pandemic: a comparative study. *Educ. Inform. Technol.* 1–21 (2020)

- Garrison, D.R., Anderson, T., Archer, W.: Critical inquiry in a text-based environment: Computer conferencing in higher education. *Internet High. Educ.* **2**(2–3), 87–105 (2000)
- Grace, J.B., Bollen, K.A.: *Interpreting Structural Equation Models: From the Simple to the Complex*. Guilford Press (2019)
- Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E.: *Multivariate Data Analysis*, 8th edn. Cengage Learning (2019)
- Hew, K.F., Cheung, W.S.: Attracting student participation in asynchronous online discussions: a case study of peer facilitation. *Comput. Educ.* **51**(3), 1111–1124 (2008)
- Hsu, Y.C., Wang, S.Y.: Effects of personalization in multimedia-enhanced problem-based learning environments. *Br. J. Edu. Technol.* **46**(1), 117–129 (2015)
- Hoyle, R.H.: *Handbook of Structural Equation Modeling*. Guilford Press (2012)
- Kebritchi, M., Lipschuetz, A., Santiago, L.: Issues and challenges for teaching successful online courses in higher education: a literature review. *J. Educ. Technol. Syst.* **46**(1), 4–29 (2017)
- Kline, R.B.: *Principles and Practice of Structural Equation Modeling*, 4th edn. The Guilford Press (2015)
- Islam, A.K.M.N., Beer, M., Slack, F.: Factors influencing students' usage behaviour and performance in E-Learning: examining the moderating role of gender. *Comput. Hum. Behav.* **49**, 121–133 (2015)
- Jung, I., Lei, J.: Examining the factors affecting online learners' satisfaction in higher education: a structural equation modeling approach. *Int. Rev. Res. Open Distrib. Learn.* **20**(1), 57–78 (2019)
- Lam, S.S., Nguyen, T.T.: Understanding the impact of perceived enjoyment, perceived social interaction, and perceived instructor responses on students' satisfaction and continued intention to use MOOCs. *Comput. Educ.* **141**, 103612 (2019)
- Liaw, S.S., Huang, H.M.: Exploring the relationship between e-learning system, learners' self-regulatory efficacy and e-learning usage. *Comput. Hum. Behav.* **45**, 168–175 (2015)
- Means, B., Toyama, Y., Murphy, R., Bakia, M., Jones, K.: *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. US Department of Education (2009)
- Picciano, A.G.: Beyond student perceptions: issues of interaction, presence, and performance in an online course. *J. Asynchronous Learn. Netw.* **6**(1), 21–40 (2002)
- Rahimi, E., Yadollahi, S.: Investigating the effect of interaction types on learners' satisfaction, motivation, and learning in an online learning environment. *Interact. Learn. Environ.* **28**(3), 310–324 (2020)
- Richardson, J.T., Sheldon, K.M., King, J.: Students' achievement goals, learning-related emotions and academic achievement. *Educ. Psychol.* **32**(2), 247–262 (2012)
- Rovai, A.P.: Building classroom community at a distance: a case study. *Education Tech. Research Dev.* **50**(1), 5–16 (2002)
- Sun, P.C., Tsai, R.J., Finger, G., Chen, Y.Y., Yeh, D.: What drives a successful e-learning? An empirical investigation of the critical factors influencing learner satisfaction. *Comput. Educ.* **144**, 103701 (2019)
- Tallent-Runnels, M.K., et al.: Teaching courses online: a review of the research. *Rev. Educ. Res.* **76**(1), 93–135 (2006)
- Tinto, V.: *Leaving college: Rethinking the causes and cures of student attrition*, 2nd edn. University of Chicago Press (1993)
- Wang, Q., Newlin, M.H.: Predictors of web-student performance: the role of self-efficacy and reasons for taking an online course. *Comput. Hum. Behav.* **18**(2), 151–163 (2002)
- Wu, D.D., Hiltz, S.R.: Predicting learning from asynchronous online discussions. *J. Asynchronous Learn. Netw.* **8**(2), 139–152 (2004)
- Zainal, N., Zaid, N.M., Razak, R.C.A.: Student satisfaction with e-learning: examining the impact of learning experience on academic performance. *Procedia Soc. Behav. Sci.* **234**, 355–363 (2017)



Construction and Innovative Application of MOOC for Large-Scale Improvement of Digital Teaching Literacy of Primary and Secondary School Teachers – Taking Guangdong Province of China as an Example

Youru Xie^(✉), Wan Xia, Caoyin Li, and Yi Qiu

School of Educational Information Technology, South China Normal University, No. 55,
West of Zhongshan Avenue, Tianhe District, Guangzhou, China
xieyouru@aliyun.com

Abstract. Primary and secondary school teachers are vital of high-quality basic education. Improving their digital teaching literacy is crucial to fostering a professional teaching force in the digital era. Massive open online course (MOOC) has characteristics like large-scale, openness, interactivity, and flexibility. Thus, MOOC holds promise for improving teachers' digital teaching literacy on a large scale. This study clarified the connotation elements of primary and secondary school teachers' digital teaching literacy. Guided by this literacy's requirement, the study established a CCD development model of MOOC from construction concept, teaching content, design and development, and developed “*Innovative Application of Digital Teaching Textbooks*” MOOC. For Guangdong Province's primary and secondary school teachers in China, two rounds of design-based research were conducted, forming a path for MOOC application to improving their digital teaching literacy. Effect analysis showed high participation among these teachers, affirming the MOOC's efficacy improving their digital teaching literacy and application level.

Keywords: primary and secondary school teachers · digital teaching literacy · MOOC construction · innovative application

1 Introduction

Primary and secondary school teachers are the backbone of the high-quality development of basic education. Teaching is their fundamental duty, and digital teaching is an inevitable choice for classroom reform. Improving the digital teaching literacy of these teachers is crucial to fostering a high-quality and professional teaching force in the digital era. As a key way to improve the support service capacity of online education, Massive open online course (MOOC) has the characteristics of large-scale, openness, interactivity, flexibility, which is an effective way for the training of primary and secondary school teachers to improve their digital teaching literacy. Addressing the current problems of

unspecified standards and obvious lack of digital teaching literacy of these teachers, this study clarified the connotation elements of primary and secondary school teachers' digital teaching literacy. Guided by this literacy's requirement, the study established a CCD development model of MOOC from construction concept, teaching content, design and development, and developed "*Innovative Application of Digital Teaching Textbooks*" MOOC. For Guangdong Province's primary and secondary school teachers in China, two rounds of design-based research were conducted, forming a path for MOOC application to improving their digital teaching literacy. The research will address the following questions:

1. What are the connotation elements of digital teaching literacy for primary and secondary school teachers?
2. How to develop, construct and innovatively apply MOOC that improve the digital teaching literacy of primary and secondary school teachers?
3. What is the effect of the application of MOOC to improve the digital teaching literacy of primary and secondary school teachers?

2 Literature Review

2.1 A Research Review on Digital Teaching Literacy of Primary and Secondary School Teachers

Teaching literacy is one of the important criteria for evaluating teachers' educational abilities. Currently, there is limited research on the digital teaching literacy of primary and secondary school teachers. For instance, Záhorec et al. (2019) assessed the digital teaching literacy of Slovak primary and secondary school teachers through a questionnaire survey and identified the literacy's cultivation needs. Marín and Castañeda (2022) reviewed the relevant concepts and evolution of digital teaching literacy, proposing strategies for cultivating teachers' digital teaching literacy. Yan and Liu (2022) analyzed seven representative frameworks for teachers' digital literacy, highlighting the European Union's focus on "digitalized teaching." Nguyen and Habók (2023) reviewed 33 papers, revealing that 78% of assessment frameworks relate to teachers' teaching skills. In summary, there is currently no unified understanding of teachers' digital teaching literacy in the academic community. Therefore, further exploration is needed to clarify the connotation and elements of teachers' digital teaching literacy.

2.2 A Research Review on the Development and Innovative Application of MOOC

MOOC's global application and continuous development have positioned it as a "digital foundation" for advancing high-quality education. Over the past decade, MOOC has seen advancements across several domains, including development frameworks, resource structures, activity organization, course interactions, support services, learning assessment, and innovative application. Examples include Feng's (2018) "FLTC (Facilitator-Learner-Text-Context)" MOOC development model, Kim et al.'s (2019) integration of virtual reality into MOOC resources, Sun's (2021) interactive analysis framework of

online course video, Wang et al.’s (2017) learning activity design using *Instructional Interaction Hierarchical Model*, Wang’s (2021) proposal for Coursera to automatically identify at-risk learners and intervene through information push, Duan and Wu (2023) “student-centered, teacher, and peer auxiliary” student self-assessment paradigm, King et al.’s (2018) use of professional development MOOCs to improve language literacy among English teachers. Despite extensive research on MOOC development and application, addressing the needs of substantially improving teachers’ digital teaching literacy on a large scale through efficient MOOC design, development, and flexible application remains a worthy avenue of exploration.

3 Research Process and Methods

This study applies various research methods, following the steps of “Status Investigation—Literacy Clarification—MOOC Development—MOOC Application—Effect Analysis”, as shown in Fig. 1.

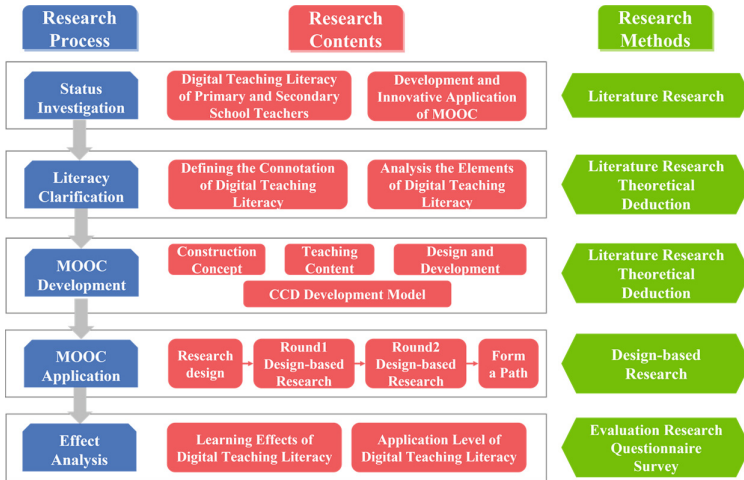


Fig. 1. Research Process and Methods.

4 Research Content

4.1 Connotation Elements of Digital Teaching Literacy

Definition of Digital Teaching Literacy. Literacy refers to the ability of people to use the knowledge and skills they have learned to effectively analyze, reason, communicate, and to pose, solve, and interpret problems in major subject areas as well as in a variety of contexts (OECD, 2009, p. 13). Teaching literacy is a fundamental requirement for teachers, and many scholars have discussed it, such as Zhu (2023) believed that teaching

literacy includes teaching knowledge, teaching ability and teaching affection. Xu (2007) pointed out that teaching literacy emphasizes the standard and correct attitude of teachers in terms of theory, knowledge, ability, and so on. Drawing on PISA's definition of literacy, this study argues that teaching literacy refers to the comprehensive ability of teachers to use subject knowledge and skills to create contexts, pose, analyze and solve problems in the teaching process. Based on the above views, this study further proposes that digital teaching literacy of primary and secondary school teachers refers to the comprehensive ability of primary and secondary school teachers to make appropriate use of digital technologies and resources, apply subject knowledge and skills, create contexts in the teaching process, and pose, analyze and solve problems, optimize, with the aim of innovating and transforming teaching activities.

Analysis of Elements of Digital Teaching Literacy. This study combines the connotation of digital teaching literacy and summarizes the components of digital teaching literacy into four dimensions, the specific contents are shown in Table 1.

4.2 Construction of MOOC to Improve the Digital Teaching Literacy

Based on the *curriculum design theory*, this study establishes a CCD development model from MOOC construction concept (Concept), MOOC teaching content (Content), and MOOC design and development (Develop) concerning the MOOC development model proposed by Zou et al. (2020), as shown in Fig. 2.

Clarify the Concept of MOOC Construction. Guangdong Province is located in the south of China, with a developed economy but unbalanced development, and a large population, which has 1,094,600 primary and secondary school teachers (Guangdong Provincial Department of Education, 2023). Basic education presents the characteristics of “large volume, long front, high investment, and imbalance”, and there is still a large gap between the overall quality of the teaching force and the requirements of the development of the cause of basic education for teachers, which is urgently needed to make up the short board and shrink the gap (Guangdong Provincial Department of Education, 2022), the use of MOOC to massively improve the digital teaching literacy of primary and secondary school teachers is an important way to promote the construction of the teacher team. Therefore, addressing the current problems of unspecified standards and the obvious lack of digital teaching literacy of primary and secondary school teachers, this study internalizes the demand for high-quality development of the classroom as a higher-order pursuit of the learner. The concept of MOOC construction is “to enable primary and secondary school teachers to make appropriate use of digital technology and resources, apply subject knowledge and skills, improve digital teaching awareness, innovate digital teaching design, carry out digital teaching implementation, and change digital teaching evaluation”, and this study further determines the higher-order objective of MOOC, such as cultivating people with political education, imparting knowledge, and cultivating competence. Through the design for integration of professionalism and political education “high-level objective positioning—innovative content reconstruction—flexible adjustment of activities—high-quality and appropriate resources”, the standard requirements of digital teaching literacy are deepened into the

Table 1. Dimensions and descriptions of digital teaching literacy.

The first dimension	The second dimension	Descriptive
Digital Teaching Awareness	Digital Teaching Perception	I understand the value of digital technology in economic, social, and educational development, and recognize the opportunities and challenges that the development of digital technology brings to changes in teaching and learning.
	Digital Teaching Willingness	I have willingness to learn and use digital technologies and have the motivation to utilize digital technologies for teaching practices and teaching innovation.
	Digital Teaching Determination	I can overcome the difficulties and challenges faced in digital teaching practice, and I strongly believe in and continue to explore digital teaching practice.
Digital Teaching Design	Digital Teaching Objective Design	I can focus on the core literacy of the discipline, scientifically determine the teaching objectives based on student learning data, and focus on the organic fusion of discipline teaching and political education.
	Digital Teaching Content Design	I can reconstruct teaching content and realize interdisciplinary knowledge integration by using digital technologies such as knowledge mapping and data mining under the requirements of the “ <i>Double New</i> ”.
	Digital Teaching Activity Design	I can reasonably use digital teaching resources and digital tools such as generative AI, flexibly adjust teaching activities, and realize rich and active teacher-student, student-student, and human-machine interaction according to the requirements of “ <i>Double Reduction</i> ”.

(continued)

Table 1. (continued)

The first dimension	The second dimension	Descriptive
	Digital Teaching Evaluation Design	I can benchmark the concept of “teaching-learning-evaluation” integration, synthesize multi-source, multi-dimensional, and multi-modal teaching data to carry out concomitant, evidence-based, and value-added evaluations, so as to implement the fundamental task of cultivating moral character.
Digital Teaching Implementation	Digital Teaching Process Reorganization	I can use digital technology to dynamically organize a reasonable and complete teaching process with subject-specific characteristics, achieve iterative optimization of the process based on teaching process data, and change the way teachers teach and students learn.
	Digital Teaching Structure Reengineering	I can utilize digital technology to flexibly create a teaching structure that is bidirectionally empowered by human and machines and continues to evolve, establish a closed loop of mutual evidence of teaching elements and interactive data, and change the structure of classroom teaching.
Digital Teaching Evaluation	Digital Value Judgment	I can benchmark teaching objectives and academic quality standards, utilize multi-source, multi-dimensional, and multi-modal teaching data to make value judgments about student learning outcomes.

(continued)

Table 1. (continued)

The first dimension	The second dimension	Descriptive
	Digital Teaching Diagnosis	I can utilize multi-source, multi-dimensional, multi-modal teaching data to gain insight into student problems and make evidence-based diagnoses and rational attributions.
	Digital Teaching Improvement	I can utilize multi-source, multi-dimensional, and multi-modal teaching data to make adjustments and improvements to teaching design and implementation

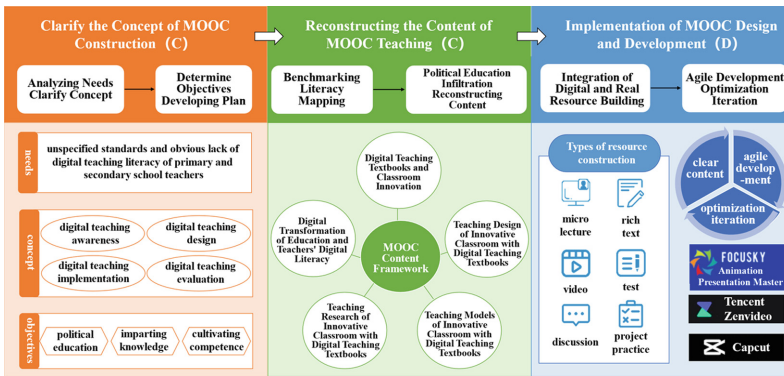


Fig. 2. CCD development model.

professional qualities of learners, and primary and secondary school teachers are given the innovative ability to cope with the educational problems and challenges, which fully reflects the high-level, innovative and challenging nature of the MOOC.

Reconstructing the Content of MOOC Teaching. This study hits the cutting-edge dynamics and development trend of the innovative application of digital teaching textbooks, adopts the results-oriented reverse design concept, and under the guidance of the higher-order objective, benchmarks the digital teaching literacy standard requirements, dynamically reconstructs the teaching content of MOOC by means of knowledge map, and gives the teaching content the momentum of continuous growth, as shown in Fig. 3. As shown in Fig. 4, teaching content are carried by a network of knowledge points related to classroom teaching innovation, teaching resources, design, models and research. Each module is precisely directed towards the elements of digital teaching literacy of primary and secondary school teachers, allowing for the precise penetration of multidimensional and multiform cases of political education. The knowledge map records the learning data of each node, providing the basis to carry out blended learning and practice.



Fig. 3. Knowledge map.

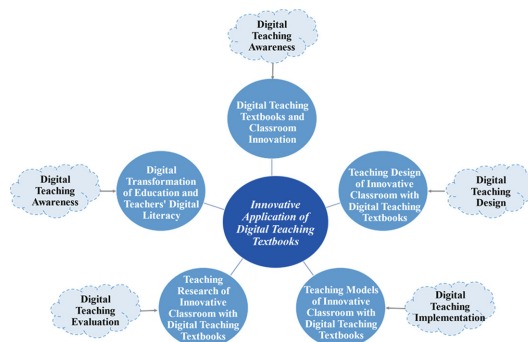


Fig.4. Match between the content of the “*Innovative Application of Digital Teaching Textbooks*” MOOC and the digital teaching literacy of primary and secondary school teachers.

Implementation of MOOC Design and Development. This study employed tools like *Focusky Animation Presentation Master*, *Tencent Zenvideo*, and *Capcut* to execute the design and development of the MOOC following the process of “clear content-agile development-optimization iteration”. Emphasizing the integration of digital and physical resources, it encompassed various resources like videos, texts, and cases, including micro-lecture, videos, discussion, rich texts, tests, project practice, etc.

The MOOC featured 37 high-quality videos totaling over 420 minutes and 62 rich text resources comprising teaching design plans, examples, and research. MOOC also set up project practices with a certain degree of challenge, specifically to improve the innovative application ability of primary and secondary school teachers with digital teaching textbooks.

4.3 MOOC Application for Improving Digital Teaching Literacy

Research Design of MOOC Application. This study relied on the *Xueyin Online platform “Innovative Application of Digital Teaching Textbooks”* MOOC to carry out two rounds of application respectively. The research object was primary and secondary school teachers in Guangdong Province of China, who independently select courses for learning. The specific research design is shown in Table 2.

Table 2. Design-based research schedule.

Design-based Research	Research Target	Research Objective	Post-testing
First round (second semester 2022)	Primary and secondary school teachers enrolled in MOOC studies (1786)	A preliminary exploration of the MOOC application process for improving digital teaching literacy of primary and secondary school teachers	<ol style="list-style-type: none"> 1. Learning Effects on Digital Teaching Literacy for Primary and Secondary School Teachers 2. Application level of digital teaching literacy for primary and secondary school teachers
Second round (first semester 2023)	Primary and secondary school teachers enrolled in MOOC studies (1248)	Based on the previous round, observe and analyze whether the process of MOOC application is reasonable, whether the function of human-machine collaboration is highlighted, and summarize the improvement	

First Round of Design-based Research: Initial Exploration. *Design and Implementation.* During this round of research, relying on the *Xueyin Online platform, Teaching Design Studio Support Service Center, Tencent Meeting, QQ*, etc., the course team provided corresponding learning support services to guide teachers to learn the MOOC following the process of “Watch Video – Online Discussion – Project Practice – Learning Evaluation”, as shown in Fig. 5. In the “Watch Video” phase, the course team regularly released learning suggestions and progress reminders based on the 5-week content. In the “Online Discussion” phase, the team carried out interactive communication based on students’ questions, guided teachers to participate in asynchronous discussions of each course module, fostering ongoing generation and prompting deep reflection through likes and replies. The “Project Practice” phase focused on a results-oriented project: the “innovative application of national curriculum digital teaching textbooks for teaching design or teaching research paper writing,” which drove teachers to address deficiencies, assess their ability to use digital teaching textbooks to innovatively solve real educational problems, and evaluate the learning effectiveness. In addition, primary and secondary school teachers participated in the “*Double Integration and Double Creation*” *Teacher Information Literacy Improvement Practices of Guangdong Province* to further apply and develop their knowledge and skills. In the “Learning Evaluation” phase, the course team judged teachers’ learning performance, considering video learning, chapter quizzes, and project practice. Targeted responses were provided for self-constructed work order

inquiries, and artificial intelligence (AI) and other tools were used to offer personalized Q&A, suggest future learning directions, and encourage teachers' self-reflection and growth.

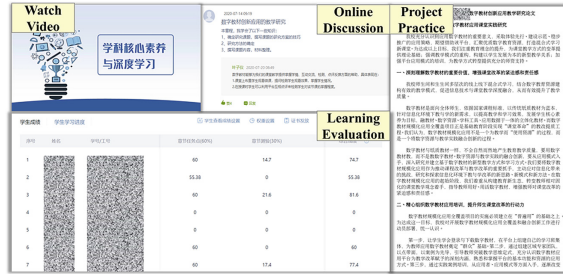


Fig.5. First round of MOOC application.

Evaluation and Reflection. In the first round, teachers met deadlines, but the ability to meet teachers' individual needs and the flexibility to adjust the arrangement of teaching activities still need to be further improved. Therefore, in the second round of research, it is necessary to focus on refining MOOC processes and enhancing human-machine collaboration.

Second Round of Design-based Research: Optimization and Improvement. *Design and Implementation.* The second round of practice further summarized the MOOC application phases, focusing on the three main lines of “Synchronous—Live Interaction + Q&A Feedback”, “Asynchronous—Case Study + Recorded Analysis”, “Hybrid—Project Practice + Human-machine Interaction”, as shown in Fig.6. Moreover, the course team emphasized the functional role of generative AI, introducing it into learning communities, course discussion forums, live Q&A activities, etc., to achieve inspirational interactions, improve teachers' learning experience, and meet their individual needs in an agile manner. In the “Synchronous—Live Interaction + Q&A Feedback” phase, we provided learners with personalized learning paths and guidelines, and generative AI offered real-time problem-solving support to teachers. In the “Asynchronous—Case Study + Recorded Analysis” phase, combined with the direction of the development of national education digitalization and the cutting-edge trend of classroom change, focused on the “Double New” “Double Reduction”, digital transformation of education, teachers' digital teaching literacy and other topics, we flexibly adjusted teaching activities, and generative AI provided personalized problem diagnosis, Q&A session, situational dialogue, and enlightening interactions. In the “Hybrid—Project Practice + Human-machine Interaction” phase, teachers focused on the theme of project practice, used generative AI to carry out in-depth dialogues and output results. Throughout teachers' project practice, we pushed excellent cases on demand to help teachers migrate and apply, regulate teachers' learning process, and promote teachers' effective learning.

Evaluation and Reflection. In the second round, teachers completed the course and actively utilized the knowledge and skills to participate in the “Double Integration and

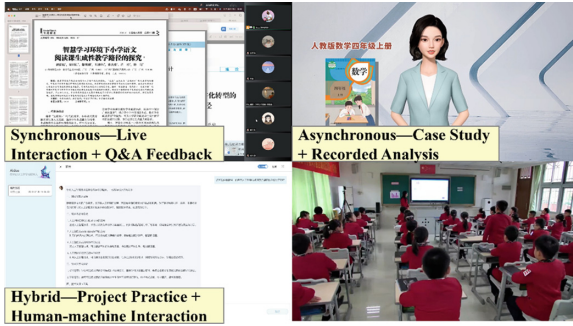


Fig. 6. Second round of MOOC application.

“Double Creation” Teacher Information Literacy Improvement Practices of Guangdong Province, promoting their own innovation and development through the integration of teaching and learning in practice. The use of generative AI also strengthens teachers’ interest in learning. However, it is still necessary to improve teachers’ data awareness in subsequent practice and guide them to actively apply intelligent technologies to further optimize, change and innovate the teaching process structure.

The Path of Innovative Application of MOOC. After two rounds of design-based research, this study formed a path for MOOC innovative application to improve the digital teaching literacy of primary and secondary school teachers according to the steps of the “design for integration of professionalism and political education—development of the integration of digital and physical—collaborative implementation of human-machine—multi-dimensional evidence-based evaluation”, as shown in Fig. 7.

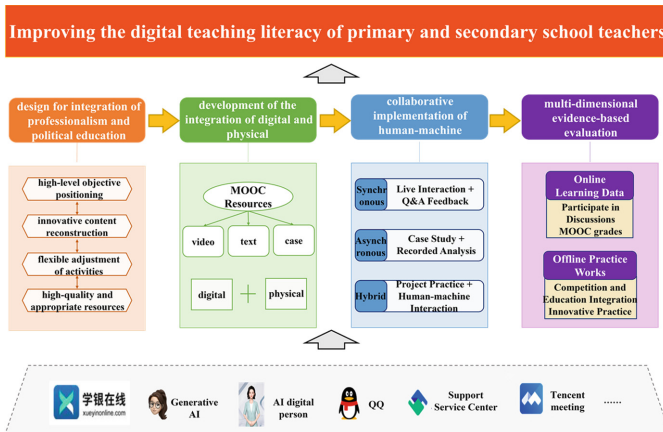


Fig. 7. Innovative application paths of MOOC.

5 Research Effect

5.1 Analysis of Learning Effects on Digital Teaching Literacy

This study conducted statistical analysis on the relevant learning data from two rounds of the “*Innovative Application of Digital Teaching Textbooks*” MOOC, as shown in Table 3. These findings illustrate that the learning effects of digital teaching literacy of primary and secondary school teachers were favorable during both rounds of MOOC application, with a high level of online participation. Notably, the second round demonstrated more significant advantages in pass rates, suggesting that after two rounds of design-based research adjustments and improvements, the MOOC was better able to support the knowledge acquisition and interactive engagement of primary and secondary school teachers, effectively enhancing their digital teaching literacy.

Table 3. Analysis of data from two rounds of MOOC application.

	First Round	Second Round
Course Enrollments	1786	1248
Number of Teachers Who Passed	1683	1240
Pass Rate	94.23%	99.36%
Number of Participants in Interaction	833	640
Number of Posts	2892	3092

5.2 Analysis of Application Level of Digital Teaching Literacy

Based on the connotation elements of digital teaching literacy of primary and secondary school teachers, and referring to relevant teaching literacy scales, this study adapted and formed the evaluation indicators of digital teaching literacy of primary and secondary school teachers. The study created a corresponding online scale using the Likert five-point scoring method. The Cronbach’s α reliability coefficient for this scale is 0.952, indicating high reliability and suitability for use.

Based on the results of digital teaching literacy measurement of primary and secondary school teachers before course commencement, and during two rounds of MOOC application, this study categorized teachers’ MOOC scores into four levels. According to the ratio “Level A: Level B: Level C: Level D = 2:3:3:2”, 30 teachers from each round with similar scale scores were randomly selected for analysis. The research team analyzed teachers’ integrated works in teaching and competition, which involved submitting an innovative digital technology application teaching case, including teaching design plans, teaching video recordings, and reflective texts. Two scholars independently scored the integrated works based on the digital teaching literacy evaluation framework. The Kappa consistency test was conducted on the scoring results, yielding a Kappa value of 0.843, indicating good consistency in the scoring results. The study performed

an independent samples t-test on the integrated works scores of 60 primary and secondary school teachers, with results shown in Table 4. It is evident that the average score of teachers' works increased from 30.97 to 32.17. The score for each dimension in the second round were slightly higher than those in the first round, with the dimension of digital teaching design showing the most significant improvement, which demonstrates that the "*Innovative Application of Digital Teaching Textbooks*" MOOC effectively promotes the improvement of the application level of digital teaching literacy.

Table 4. Analysis of integrated works scores.

	Round 1			Round 2			F	Sig.
	N	Mean	SD	N	Mean	SD		
Integrated works of competition and education	30	30.97	7.636	30	32.17	8.538	.066	.568
Digital Teaching Awareness	30	8.30	2.744	30	8.60	2.990	.126	.687
Digital Teaching Design	30	12.30	2.718	30	12.87	3.003	.357	.447
Digital Teaching Implementation	30	5.33	1.561	30	5.47	1.717	.063	.754
Digital Teaching Evaluation	30	5.03	1.629	30	5.23	2.176	.952	.688

6 Conclusion

This study clarified the connotation elements of digital teaching literacy of primary and secondary school teachers; guided by this literacy's requirement, it established a CCD development model of MOOC from construction concept, teaching content, design and development, and developed the "*Innovative Application of Digital Teaching Textbooks*" MOOC; for Guangdong Province's primary and secondary school teachers in China, two rounds of design-based research were conducted, and formed a path for MOOC application to improve their digital teaching literacy. Effect analysis showed high participation among these teachers, affirming the MOOC's efficacy improving their digital teaching literacy and application level.

We hope that the digital teaching literacy scale for primary and secondary school teachers, the CCD development model, the "*Innovative Application of Digital Teaching Textbooks*" MOOC, and the path of MOOC innovative application can provide both theoretical and practical references to cultivate primary and secondary school teacher force with digital teaching literacy.

References

- Duan, T., Wu, B.: The student self-assessment paradigm in MOOC: an example in Chinese higher education. *Comunicar* **31**(75), 115–128 (2023)
- Feng, Y.: Construction of “FLTC” Model of MOOC development from the perspective of “Paradigm” fusion. *e-Educ. Res.* **39**(08), 64–69 (2018)
- Guangdong Provincial Department of Education: 2022 Guangdong Provincial Educational Development Statistical Bulletin. http://edu.gd.gov.cn/zwgknew/sjfb/content/post_4196860.html (2023)
- Guangdong Provincial Department of Education: Extraordinary Ten Years· Guangdong Education | Five major systems promote the high-quality development of basic education in Guangdong. <http://edu.gd.gov.cn/hdjlpt/live/index/index/records/pc?pid=2955&jump=0&siteId=168> (2022)
- Kim, H., Nah, S., Oh, J., Ryu, H.: VR-MOOCs: a learning management system for VR education. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 1325–1326 (2019)
- King, M., Luan, B., Lopes, E.: Experiences of timorese language teachers in a blended massive open online course (MOOC) for continuing professional development (CPD). *Open Praxis* **10**(3), 279–287 (2018)
- Marín, V.I., Castañeda, L.: Developing Digital Literacy for Teaching and Learning. In: *Handbook of Open, Distance and Digital Education*, pp. 1–20. Springer Nature Singapore, Singapore (2022). https://doi.org/10.1007/978-981-19-0351-9_64-1
- Nguyen, L.A.T., Habók, A.: Tools for assessing teacher digital literacy: a review. *J. Comput. Educ.* (2023). <https://doi.org/10.1007/s40692-022-00257-5>
- OECD: The PISA 2009 Assessment Framework: Key Competencies in Reading, Mathematics, and Science. OECD, Paris (2009)
- Sun, T.: Modes of interaction in online course videos and optimization strategy. *Chin. J. Distance Educ.* **01**, 57–65 (2021)
- Wang, Y.: A review of the global MOOC movement in 2019. *Chin. J. Distance Educ.* **05**, 68–75 (2021)
- Wang, Z., Zhao, H., Chen, L.: Using instructional interaction hierarchical model to design learning activities. *Chin. J. Distance Educ.* **06**, 39–47+80 (2017)
- Xu, D.: *The New Construction of Chinese Teacher’s Accomplishment under the Background of New Courses*, Dissertation. Zhejiang Normal University (2007)
- Yan, G., Liu, L.: Research on teachers’ digital literacy and its cultivation path: a comparative analysis based on the seven digital literacy frameworks for teachers in the European union. *Int. Comp. Educ.* **44**(03), 10–18 (2022)
- Záhorec, J., Hašková, A., Munk, M.: Teachers’ professional digital literacy skills and their upgrade. *Eur. J. Contemp. Educ.* **8**(02), 378–393 (2019)
- Zhu, D., Peng, H.: An empirical study on the evaluation model of teacher’s interdisciplinary instruction competency. *J. East China Normal Univ. (Educ. Sci.)* **41**(02), 1–13 (2023)
- Zou, Y., Xie, Y., Qiu, Y., Li, J., Bai, J.: Development and application of a massive open online course (MOOC) “online teaching in the epidemic period” based on the object-content-develop (OCD) model. In: 2020 Ninth International Conference of Educational Innovation through Technology (EITT), pp. 58–63 (2020)



Measuring Emotion in Education Using GSR and HR Data from Wearable Devices

Qian Dong¹ and Rong Miao²(✉)

¹ School of Physical Education, Shanghai University of Sport, Shanghai, China

² Graduate School of Education, Peking University, Beijing, China

st16625m@gse.pku.edu.cn

Abstract. The essential role of emotions in education is widely valued, which has attracted enormous interest in emotion recognition. Wearable devices that collect real-time and fine-grained subjective physiological data make it possible to complex emotional dynamics in education. However, less studies are known about its potential role in examining the emotional changes of different levels of students under specific pedagogical actions in face-to-face chess classroom. Throughout objective GSR and HR data from wearable devices and subjective data from students and teachers in chess classrooms, it is found students at different levels had experienced different emotions for the same pedagogical actions. Moreover, teachers should pay more attention to grasping the teaching content of students at different levels, setting appropriate challenges of difficulty, and using brackets and prompts at appropriate times for challenging assignments. These findings evidence the feasibility of data analysis model of wearable devices but also serve as a reference point and future investigations on the explanation of how and why students' emotions varied of more specific pedagogical actions in education.

Keywords: emotion recognition · wearable devices · chess education · galvanic skin response · heart rate

1 Introduction

The significance of emotion recognition during the learning process is critical in education since it is linked to learners' motivation, interest, attitude, aspiration, and attention [1]. It has become a trend to use wearable devices to collect various real-time data to indicate students' emotional states. Researchers have been interested in examining emotion since 2012, but due to the complexity of emotion measurement, experimental environments have largely been restricted to laboratories. With the development of technology over the past few years, experimental settings have gradually expanded to real-life teaching scenarios, such as face-to-face lectures. As for measuring signals, most studies concentrated on indicators obtained from human bodies, such as Apicella's research on electroencephalogram (EEG) [2] and Antoniou's work utilizing electrodermal activity (EDA) to access emotion [3]. In comparison to EEG, skin temperature (ST), blood volume pressure (BVP), galvanic skin response (GSR) and heart rate (HR) data were

the most widely adopted methods for measuring emotion in education [4]. This paper aims to study the emotion in a learning process to better design the learning materials and educational procedures. Among all the studies, the effects of specific pedagogical actions on emotion in education research have been rather rare, not to mention paying attention to students at different levels. Besides, unlike many studies in physical education, it is noteworthy that no research has been conducted on the use of wearable devices for emotional measurement in face-to-face chess classrooms.

This paper aims to examine the emotional changes of different levels of students via objective GSR and HR data from a wearable device and subjective data from students and teachers under specific pedagogical actions in face-to-face chess education classrooms. In doing so, it not only sheds light on students' real-time emotional feedback to specific activities for teachers during the learning process but also provides insight and inspiration to aid in the ongoing quest for better educational design in the future. Such knowledge might differ from those of the original authors in chess education.

2 Literature Review

2.1 Emotion Recognition in Education

Emotions have a strong connection to cognitive functions and are crucial for cognitive processes such as information processing, memory performance, attribution decision-making, and intellectual manipulation in education. The research on emotion recognition in education has gradually shifted from what to how. From 2008 to 2012, early research focused on delicately classifying/predicting emotions. When objective things align with one's own needs, individuals generate positive emotions such as satisfaction, liking, pride, happiness, etc. Conversely, negative emotions such as anger, fear, and pain are triggered.

Ba & Hu [4] found that there were 30 types of emotions measured from the previous work and among them, "engagement" appeared the most in education, such as Shen, Callaghan, and Shen examined learners' engagement from physiological signals in ICT environments to explore how to optimize the digital learning environment based on the dynamic evolution of learners' emotions [5]. Studies examining the connections between emotions and other learning-related factors and assessing the effects of instructional designs on students' emotions have been carried out since 2018. Researchers try to use cognitive load theory, self-regulated learning, and embodied cognition theory to explain the relationship between students' learning performance and emotions [6–8].

Researchers have been interested in studying emotion, but due to the difficulty of measuring emotion, experimental settings have mostly been limited to laboratories. Over the past few years, as technology has advanced, experimental settings have steadily expanded to include real classrooms. In a laboratory environment, it is difficult to induce certain emotional reactions in humans, and data under natural conditions makes it difficult to ensure the quantification of research needs. Hence, experimental research on emotions requires a high level of rationality in research methods and effectiveness in research design.

The Japanese education community has taken the lead in introducing physiological indicators measurement and analysis into classroom teaching research, such as the GSR

responses of students measured in different classrooms by Nobuo Hori of Fukui University and Masashi Honma of Miyagi Education University [9]. Chinese researchers have deployed wearable neurophysiological collection devices for the first time to record students' physiological signals, such as skin electricity and heart rate, for 90 min continuously in real math exam scenarios [10]. Among all the studies, the effects of specific pedagogical actions on emotion in education research have been rather rare, not to mention paying attention to students at different levels. There is a study gap on the impacts of specific pedagogical actions on students' emotions in all studies, and there are surprisingly few studies that have concentrated on the effects of specific pedagogical activities on the emotions of students at different levels in actual teaching contexts.

2.2 Emotional Measurement Methods

Emotion refers to an attitudinal experience associated with a person's material or physiological needs, typically lasting for a short period, ranging from a few seconds to a few minutes. Emotions usually begin with the evaluation of the personal meaning of an event and are then manifested in three parts: subjective experience, external expression, and physiological response. These three aspects correspond to the three indicators of emotional measurement, namely the subjective experience within an individual, external behavioral performance, and physiological changes within the body [11].

The subjective experience method usually uses standardized scales to have participants report their subjective feelings. Such as adjective checklist, dimension scale, rank scale, SAM emotion scale, emotion mood scale, differentiation emotion scale, etc. This measurement method is simple and convenient, without the need for specialized analytical techniques or the introduction of complex instruments and equipment. However, it is difficult to avoid situations such as the disguise or concealment of the subject, and there is a lack of objectivity.

The external manifestations include actions, expressions, postures, gestures, etc., among which facial expressions are the most distinctive. The measurement tools include early measurement tools such as circular scales and 3D pattern maps, as well as modern measurement systems based on facial expressions and facial action encoding recognition technology.

Human emotional information can be perceived through visual signals (such as facial expressions, head movements, body posture, gestures, etc.), speech signals (such as tone, intensity, frequency, duration, etc.), and physiological signals (such as heart rate, electromyography, electroencephalography, electrodermal response, thermal infrared imaging, etc.). Physiological signals are controlled by the autonomic nervous system and endocrine system, making it difficult to be controlled by subjective will. Other indirect methods for identifying results are more objective, realistic, and persuasive, while physiological indicators are more robust and objective [12, 13].

Emotional measurement can be measured through various methods. The most extensively used wearable methods for assessing emotion in education were heart rate (HR) and galvanic skin response (GSR) data. However, in actual measurement, emotions and a certain physiological indicator are not a one-to-one correspondence, so when measuring emotions, it is not necessary to rely solely on individual indicators for measurement

but rather to use multiple methods comprehensively. Therefore, this article adopts a combination of subjective and objective methods for emotional measurement.

3 Methods

3.1 Participants

We selected S institution's classroom, a municipal off-campus education institution of the Beijing Municipal Education Commission as the research site. The school had strict requirements and required the submission of a health commitment letter to ensure physical health before attending classes. Due to the susceptibility of emotions to health conditions, this regulation ensured the suitability of the study subjects. The chess teacher D had many years of teaching experience in primary and secondary schools.

A total of 42 students (Chinese, 37 males and 6 females) were recruited for the experiment. They ranged in age from 8–14 years old. All the subjects were in good health, without any psychiatric illnesses. These students had some chess experience. Because this course was a new one, the teaching teacher adjusted the class based on the performance of each class's chess skills in the previous semester, dividing it into advanced class, intermediate class, and elementary class. However, according to interviews with teacher C from the previous semester, the differences in students' chess skills among the three classes were relatively small. At the online parent meeting held before class, the teacher introduced the experiment and wearable devices to the parents, and the parents expressed their support for this. During the course, students were required to wear wearable bracelets voluntarily. The specific participants' information is shown in Table 1.

Table 1. Participants information

Class	Chess level	Male	Female	Age	Participants wearing wearable devices
A	Advanced	17	1	9–14	5 males
B	Intermediate	12	1	8–11	4 males
C	elementary	7	4	8–10	6(2 males, 4 females)

3.2 Emotion Recognition Tools

3.2.1 Wearable Neurophysiological Measuring Instrument-wrist Type: Collecting Objective Physiological Data

The wearable device has obtained CE and RoHS certification, weighing less than 30 g. The right side of the dial includes four indicator lights: power, function, Bluetooth, and charging. The battery life of the watch is greater than 48 h, and it can collect raw data such as pulse waves and skin electricity. It can be transmitted and exported locally through USB to ensure data security. The data format is universal and can be imported into

third-party analysis tools such as SPSS and MATLAB with just one click. The sampling rate of the pulse wave is 20 Hz, the accuracy of heart rate is ± 2 BPM, and the sampling rate of skin electrography is 40 Hz, with an accuracy of 0.01 μ S. Due to the younger age of the selected participants, this wristwatch was chosen to prevent students from touching it. The wristwatch will automatically start collecting data without charging, and there will be no interruption in data collection caused by students' misoperation. In class, the teacher can collect the heart rate and electrodermal data of each student using a wristwatch.

3.2.2 SAM Self Emotional Assessment Scale: Collecting Subjective Perception Data

Due to the younger age of students and their slower reading speed, the Self-Assessment Manikin (SAM) was selected to make judgments about their current emotional state in a relatively short time. The SAM scale used in this study is a graphical self-assessment report based on Russell's arousal valence dimension. The scale was designed by Lang et al. in 1980 based on Osgood's semantic differentiation scale. SAM scores emotions from two dimensions: valence and arousal. Valence distinguishes the types of emotions, ranging from negative to positive emotions. Arousal characterizes the level of emotional activity and is generally considered the allocation of attention resources. These two dimensions characterize emotions together. Each dimension adopts a set of cartoon portrait images. Students can choose the valence dimension from sadness to pleasure based on their current emotional state, as well as the arousal dimension from relaxed to tense.

3.3 Experimental Procedure

In this study, each class had the same experimental procedure. It included three stages: (1) preparation, (2) data collection, (3) data analysis. Some of these stages can be divided into sub-phases. The detailed experimental procedure (see Fig. 1) and their corresponding inquiry phases are described as follows.

In the preparation stage, three parts were conducted in the experiment: collecting students' information, determining course information, and wearing wearable devices. In the first two parts, the teacher collected basic information about students and the courses. Then, according to voluntary principle, teachers helped students wear wearable devices for students.

In the data collection stage, event sampling in experience sampling methods were used for these four parts. In the first part, the teacher gave a lecture on chess puzzles as examples and showed instructions on how to deal with solving puzzles. After that, each student was provided with a SAM scale to fill before doing puzzles (T1). During the process of students' completing the puzzles independently, teachers observed the students' performance. At the 7th minute, students needed to fill out the SAM scale again (T2). When the students finished solving puzzles, the SAM scale had to be filled out by each student (T3). Then, the teacher started to explain the puzzles while asking the students to correct the answers. After the teacher had finished speaking, the students completed the

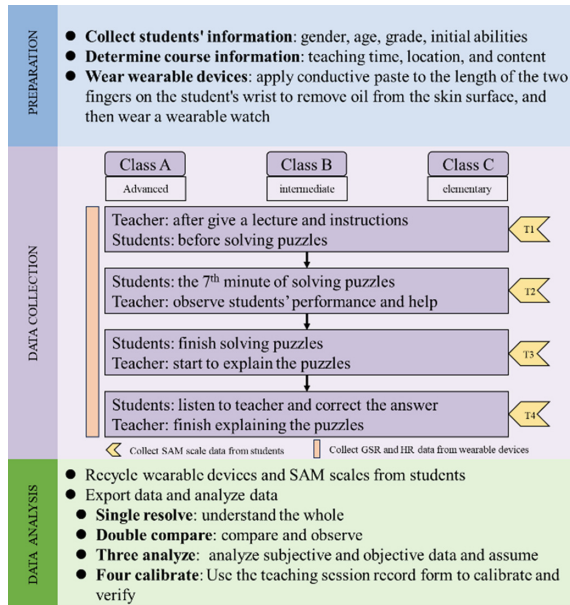


Fig. 1. Diagram of experiment design

SAM scale a final time (T4). Wearable devices assisted teachers in gathering students' GSR and HR data throughout the whole process to understand their emotions.

The data analysis stage includes two sub-phases: recycle wearable devices and SAM scales from students, and export and analyze data. Based on the data analysis model of wearable devices by Dong [14], it forms a data model including four stages: (1) single resolve, (2) double compare, (3) three analyze, (4) four calibrate, which can help teachers analyze data by highlighting key steps. Figure 1 shows a figure consisting of different types of lines. Elements of the figure described in the caption should be set in italics, in parentheses, as shown in the sample caption.

4 Results

4.1 Single Remove

The study conducted repeated measurement analysis of variance in emotions (including valence and arousal dimensions) of the participants in Class A, Class B, and Class C. The dependent variables are potency and arousal, with student level (advanced, intermediate, elementary) and course teaching stage (before starting to solve the puzzles, the 7th minute to solve the puzzles, finishing solving the puzzles, and starting to explain the puzzles, and finishing the explaining the puzzles) as the internal factors of the subjects.

4.1.1 Emotions in Different Teaching Stages

Valence Dimension

Table 2 presents the mean and standard deviations of the valence dimension of emotions acquired by students in three classes at T1, T2, T3, and T4.

Table 2. Descriptive analysis of the valence dimension of emotions in different teaching stages

Time	Class	Mean	SD	Students' number
T1	A	5.06	1.392	18
	B	6.08	1.038	13
	C	5.45	1.036	11
	Total	5.48	1.254	42
T2	A	3.83	1.339	18
	B	4.23	1.878	13
	C	4.73	0.786	11
	Total	4.19	1.435	42
T3	A	4.17	1.295	18
	B	4.15	1.676	13
	C	5.91	1.221	11
	Total	4.62	1.577	42
T4	A	5.11	1.183	18
	B	5.54	1.506	13
	C	6.18	1.079	11
	Total	5.52	1.311	42

In the four teaching stages of the three classes, the valence range of students was 3–7, and the overall trend of valence was first decreasing and then increasing, indicating that students had transitioned from a happy state to a more neutral state and then to a happier state.

After the Mauchly spherical test, the null hypothesis of potency for different teaching stages was established ($p = 0.918 > 0.05$), indicating that the results of each repeated measurement were unrelated. The results of the within-subject effect test could be seen, indicating significant differences in potency among different teaching stages ($p = 0.000 < 0.05$). There was a significant difference in the potency of students between T1 and T2 ($p = 0.000 < 0.05$), and T1 and T3 ($p = 0.009 < 0.05$) stages; There was a significant

difference in the potency of T2 and T4 ($p = 0.000 < 0.05$); There was a significant difference in the potency of T3 and T4 ($p = 0.011 < 0.05$).

Arousal Dimension.

Table 3 presents the mean and standard deviations of the arousal dimension of emotions acquired by students in three classes at T1, T2, T3, and T4.

Table 3. Descriptive analysis of the arousal dimension of emotions in different teaching stages

Time	Class	Mean	SD	Students' number
T1	A	2.56	1.723	18
	B	1.77	1.363	13
	C	2.27	1.421	11
	Total	2.24	1.543	42
T2	A	3.33	1.879	18
	B	4.08	2.100	13
	C	2.91	1.973	11
	Total	3.45	1.978	42
T3	A	2.89	1.906	18
	B	4.00	2.121	13
	C	2.00	1.414	11
	Total	3.00	1.975	42
T4	A	2.39	1.754	18
	B	1.92	1.188	13
	C	1.73	1.849	11
	Total	2.07	1.614	42

According to the Mauchly spherical test, the spherical hypothesis of arousal in different teaching stages was valid ($p = 0.839 > 0.05$), indicating that the results of each repeated measurement were unrelated. The results of the within-subject effect test could be seen, indicating that there was a significant difference in arousal in different teaching stages ($p = 0.000 < 0.05$). There was a significant difference in arousal between students in T2 and T4 ($p = 0.000 < 0.05$), and there was a significant difference in arousal between T3 and T4 ($p = 0.024 < 0.05$).

4.1.2 Emotions in Different Classes

The paired comparison showed that there was a significant difference in valence between the elementary and advanced classes ($p = 0.015 < 0.05$), and there was no significant difference in arousal dimension among the three classes.

4.1.3 Emotions of Different Chess Classes in Different Teaching Stages

The paired comparison shows that the interaction between the teaching process and the national elephant class has an impact on student valence and arousal.

From the perspective of valence, there was a significant difference in the valence of T1 and T2 among students ($p = 0.008 < 0.05$), while there was a significant difference in the valence of T2 and T4 ($p = 0.015 < 0.05$) in the advanced class. In the intermediate class, there was a significant difference in the valence of T1 and T2 ($p = 0.000 < 0.05$), T1 and T3 ($p = 0.000 < 0.05$), and T3 and T4 ($p = 0.045 < 0.05$) among students. In the elementary class, there was a significant difference in the valence of T2 and T4 ($p = 0.038 < 0.05$) among students.

From the perspective of arousal, there was no significant difference in the arousal of students in the four stages between the advanced and elementary classes. In the intermediate class, there were significant differences in arousal among students in T1 and T2 ($p = 0.000 < 0.05$), T1 and T3 ($p = 0.001 < 0.05$), and T2 and T4 ($p = 0.003 < 0.05$).

4.2 Double Compare

Based on understanding the overall subjective feelings of students, a comparison was made by combining subjective observations from teachers.

According to the progress of the teaching process and the teacher's observation, the students in the three classes were focused and did not show any resistance when doing the puzzles, which was consistent with the positive emotions filled in by the students. During the teacher's lecture, they carefully watched the demonstration and raised their hands to answer questions. After the lecture, the students returned to a calm and relaxed state, consistent with the arousal trend filled out by the students. Among them, students in the intermediate class had the most serious attitude during class were outgoing, and showed a high level of interaction. Students in advanced class exhibited a phenomenon of polarization, some students remained silent, while others interacted very actively. Students in the elementary class were more introverted, and the class was quieter when doing puzzles, with less discussion.

The difference between the three classes was reflected in the fact that the valence of students in the advanced class slowly increased, while the valence of students in the intermediate class decreased relatively steadily, while the valence of students in the elementary class increased significantly at T2. By comparing the teaching activities, it could be seen that the three classes had completed the first four relatively difficult and basic puzzles of one-step killing, two-step killing, and tactical exercises at T2. They began to tackle the second two computationally intensive and skillful puzzles. In the advanced and intermediate classes, the teacher did not give students any prompts. Gradually, some students in the advanced class were able to complete and answer correctly, while students in the intermediate class still failed to answer. The findings in the advanced and intermediate classes validate Wood et al.'s finding that as difficulty increases, the level of operation increases, but after reaching a certain level, the increase in difficulty leads to a decrease in the level of operation[15]. In the elementary class, on the one hand, the teacher guided the direction of solving puzzles and individual tutoring, which improved

the effectiveness of the students. On the other hand, some students had misjudged the answer to the puzzles, and many students mistakenly believed that the answer was correct and felt enjoyable.

4.3 Three Analyze

At this stage, physiological data such as HR and GSR collected by wearable devices were compared and cross-validated with the teaching session record form. Firstly, we checked the HR data. If the HR data was continuous and within the normal heart rate range, it indicated that the data was valid. Next, we checked the GSR data. If there were no breakpoints and there were no abnormal mutation values in the GSR data, it indicated that the data was valid. Subsequently, based on the four stages recorded in the teaching process record form, the corresponding curves of HR and GSR were segmented and viewed. Finally, we connected the four points of the segment into an approximately continuous curve and observed its trend.

Due to significant individual differences in GSR, taking S5 students in the advanced class as an example, we compared and tested the objective physiological data collected from wearable devices and objective teaching session records. As shown in Fig. 2, the heart rate curve was continuous, and there was no mutation in GSR data. Subsequently, according to the teaching process record, students started to solve puzzles at 8:37(T1). 8:42 was the time that students worked on the puzzles for 7 min(T2). Students finished the puzzles and the teacher prepared to teach at 9:00(T3). The teacher finished teaching at 9:37(T4). It was speculated that the arousal level of the four nodes of students showed a trend of first increasing and then slowly decreasing.

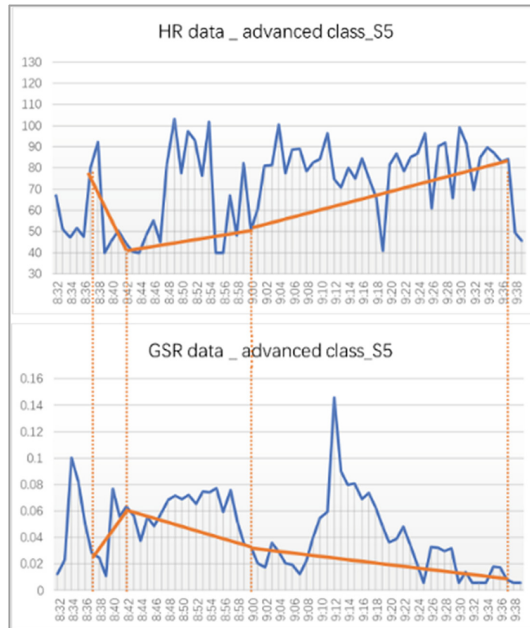


Fig. 2. HR and GSR data of S5 in the advanced class

4.4 Four Calibrate

At this stage, the trend in the subjective analysis hypothesis of students collecting data on wearable devices was compared with the trend changes in the SAM scales filled out by students in the four teaching stages for calibration testing.

Through comparison, it was found that the objective GSR data at the four nodes showed a consistent trend with the arousal data filled out by S5 as well as the overall trend of the class. This indicated that the objective GSR data could reflect students' subjective feelings, and confirmed the findings of Khalfa that changes in emotional arousal can trigger skin electrical changes [16]. Shi Zhongzhi et al. found that GSR increased with increasing arousal levels, but GSR could not distinguish between positive and negative emotions [17]. This study used subjective SAM scale validity dimension and arousal dimension data, combined with objective GSR data, to distinguish positive and negative emotions.

5 Conclusions and Discussion

In our research, we experimented to examine the emotional changes of three different levels of students via objective GSR and HR data from a wearable device and subjective data from students and teachers under specific pedagogical actions in face-to-face chess education classrooms.

The study using face-to-face teaching of chess as an example validated the feasibility of using subjective and objective data to understand the real-time emotions of students with different levels during the learning process. Overall, two new findings have been identified: 1) The data analysis model of wearable devices is clear to use, but the analysis software developed with the wearable wrist needs to be strengthened. For example, the software could support batch export and calculation, but could not choose to view individual and overall data of the class. Besides, it could choose the time interval for drawing, but not support choosing the start point for drawing; 2) Students at different levels had experienced different emotions for the same pedagogical actions. In terms of course integration, teachers should pay more attention to grasping the teaching content of students at different levels, setting appropriate challenges of difficulty, and for difficult exercises, using brackets and prompts at appropriate times, so that students can have a better experience of the teaching content. In terms of strategic innovation, teachers should choose appropriate teaching strategies based on teaching objectives, such as paying attention to the allocation of practice time and explanation time, using teaching tools such as chessboards, hanging boards, and multimedia videos to provide students with more availability and stimulate their interest in learning.

The limitations of this study lie in two aspects. Not all students fully wore the wearable devices in this study because of the small number of wearable devices. Besides, this experiment was a short-term experiment focused on puzzle solving as the specific pedagogical actions in chess. This leaves room to think and explore the process to provide more evidence on the explanation of how and why students' emotions varied of more specific pedagogical actions in other subjects.

References

1. Antoniou, P.E., et al.: Biosensor real-time affective analytics in virtual and mixed reality medical education serious games: cohort study. *JMIR Serious Games* **8**(3), e17823 (2020)
2. Apicella, A., Arpaia, P., Frosolone, M., Improta, G., Moccaldi, N., Pollastro, A.: EEG-based measurement system for monitoring student engagement in learning 4.0. *Sci. Rep.* **12**(1), 5857 (2022)
3. Antoniou, P.E., Arfaras, G., Pandria, N., Athanasiou, A., Ntakakis, G., Babatsikos, E., et al.: Biosensor real-time affective analytics in virtual and mixed reality medical education serious games: Cohort study. *JMIR Serious Games*. **8**(3), e17823 (2020)
4. Ba, S., Xiao, H.: Measuring emotions in education using wearable devices: a systematic review. *Comput. Educ.* **200**, 104797 (2023)
5. Shen, L., Callaghan, V., Shen, R.: Affective e-Learning in residential and pervasive computing environments. *Inf. Syst. Front.* **10**(4), 461–472 (2008)
6. Zhao, G., Zhang, L., Chu, J., Zhu, W., Hu, B., He, H., et al.: An augmented reality based mobile photography application to improve learning gain, decrease cognitive load, and achieve better emotional state. *Int. J. Hum.-Comput. Inter.* **39**(3), 643–658 (2022)
7. Malmberg, J., Järvelä, S., Holappa, J., Haataja, E., Huang, X., Siipo, A.: Going beyond what is visible: What multichannel data can reveal about interaction in the context of collaborative learning? *Comput. Hum. Behav.* **96**, 235–245 (2019)
8. Geršak, V., Vitulić, H.S., Prosen, S., Starc, G., Humar, I., Geršak, G.: Use of wearable devices to study activity of children in classroom; Case study-Learning geometry using movement. *Comput. Commun.* **150**, 581–588 (2010)
9. Li, J.: Teaching physiology: a science exploring human physiological information in teaching. *Educ. Res.* (010), 52–58 (1992)
10. Qu, Z., Chen, J., Li, B., et al.: Measurement of high-school students' trait math anxiety using neurophysiological recordings during math exam. *IEEE Access* **99**, 1 (2020)
11. Wang, H.: Research on Emotion Recognition Technology Based on Multi-channel Physiological Signals, pp. 3–4. Hunan University Press, Changsha (2016)
12. Han, Y., Dong, Y., Bi, J.: Physiological data representation of emotions in learning analysis: a prospective application of electrodermal response. *Modern Educ. Technol.* **28**(10), 12–19 (2018)
13. Wang, H.: Research on Emotion Recognition Technology Based on Multi-channel Physiological Signals, pp. 7–8. Hunan University Press, Changsha (2016)
14. Dong, Qian, Qu, Ximei, Miao, Rong: Data Analysis Model of Wearable Devices in Physical Education. In: Cheung, Simon K S., Lee, Lap-Kei., Simonova, Ivana, Kozel, Tomas, Kwok, Lam-For. (eds.) *ICBL 2019. LNCS*, vol. 11546, pp. 225–238. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-21562-0_19
15. Wood, C.G., Hokanson, J.E.: Effects of induced muscular tension on performance and the inverted U function. *J. Pers. Soc. Psychol.* **95**(5), 506–510 (1965)
16. Khalfa, S., Isabelle, P., Jean-Pierre, B., et al.: Event-related skin conductance responses to musical emotions in humans. *Neurosci. Lett.* **328**(2), 145–149 (2002)
17. Shi, Z.: *Cognitive Science*, p. 419. China University of Science and Technology Press, Anhui (2008)



Effect of Instruction Intervention on MOOC Forum Discussion: Student Engagement and Interaction Characteristics

Wenyi Chen^(✉)

Zhongkai University of Agriculture and Engineering, No. 501, Zhongkai Road, Guangzhou, Guangdong Province, China
winniecwy@126.com

Abstract. This study delves into the effect of instruction intervention on a MOOC forum discussion to ascertain its impact on students' participation and the characteristics of their interactions. Using a mixed-method approach combining social network analysis (SNA) and inductive qualitative analysis, the relationships and underlying interactions within an EFL MOOC discussion were examined. The findings revealed that instruction intervention by teachers effectively augmented students' engagement within the discussion area, fostering both instructor-learner and learner-learner interactions. Additionally, the conventional hub-and-spoke structure, where the instructor serves as the central node in the MOOC forum, was found to be subject to change through instruction intervention. Network analysis of the original post and a comparative assessment of two instances highlighted the direct reply tie definition as the most informative for revealing interaction relationships. It is recommended that teachers and teaching assistants actively and strategically participate in MOOC discussions to facilitate independent learning in the era of intelligent education.

Keywords: instruction intervention · learner interaction · network construction · MOOC forum discussion · social network analysis

1 Introduction

Educational institutions worldwide have increasingly embraced online instruction to cater to students across various educational levels, aligning with the progress of intelligent education. This shift has been facilitated by the widespread use of the Internet and the emergence of new online platforms, where Massive Open Online Courses (MOOCs) stand out as a consequential phenomenon. While interaction with peers and instructors is acknowledged as a potent means of learning support (Smith, 2018), the effectiveness of traditional learner-instructor communication in MOOCs is being questioned due to the large number of students involved and the diversity of their backgrounds, needs, and motivations (Johnson et al., 2020). Research highlights that students often experience feelings of isolation in online education, resulting in high dropout rates, increased boredom, and diminished achievement (Jones & Smith, 2017; Lee et al., 2019). Notably,

Gillani and Eynon (2019) assert that the discussion forums, fostering semi-synchronous exchanges among thousands of participants worldwide, distinguish MOOCs from earlier forms of online learning. Therefore, examining the dynamics of interactions within these forums promises fresh insights into the pedagogical value and potential of MOOCs.

1.1 Study on MOOC Forum Engagement

Prior research indicates that engagement in MOOC forums involves active participation in course activities (Smith, 2018), and lack of engagement increases the likelihood of discontinuation (Wang et al., 2020). MOOC discussion forums often suffer from low student involvement, with a few dominating the discourse (Jones & Smith, 2017; Xie & Zhang, 2020). To improve MOOC effectiveness, alternative approaches to student engagement are necessary. Suggestions include implementing discussion prompts that foster meaningful interactions and emphasizing self-regulation, teaching presence, and social presence as predictors of engagement and persistence (Reeve, 2012). Maintaining a positive forum atmosphere and encouraging high-quality posts are crucial for MOOC organizers (Wang et al., 2020).

1.2 Study on MOOC Forum Interaction

Previous studies indicate that in MOOCs, learner interactions are often instructor-centric, with limited learner-learner interactions (Xie & Zhang, 2020; Wu et al., 2021). The instructor's role and status can inadvertently create a hub-and-spoke structure where they dominate communication (Gillani & Eynon, 2019). Additionally, learners tend to engage more with peers and form stronger connections within content-related networks in MOOC forums (Gillani & Eynon, 2019; Xie & Zhang, 2020). Therefore, this study focuses on analyzing ties within content-related discussions to understand instructor-learner interactions within the learning community.

1.3 Instruction Intervention in MOOC Forum Discussion

In traditional lectures, instructors use pedagogical interventions based on student behavior observations (Wise & Cui, 2020). However, with larger class sizes and online interactions, it becomes challenging for instructors to observe social cues and intervene effectively (Gillani & Eynon, 2019). Instructor involvement in discussions is considered important for quality online learning (Gillani & Eynon, 2019; Wu et al., 2021), but the specific impact and ways of involvement remain unclear. Some studies suggest a positive association between instructor involvement and student contributions (Gillani & Eynon, 2019; Xie et al., 2020), while others indicate null or negative associations, suggesting potential conversation impediments (Wu et al., 2021). Wise and Cui (2020) emphasize considering both the level of involvement and the manner in which instructors engage in discussions, proposing strategies such as modeling interactive techniques, involving learners as facilitators, and utilizing analytics for monitoring interaction dynamics. Teaching presence encompasses course design, discourse facilitation, and direct instruction (Wise & Cui, 2020; Wu et al., 2021).

1.4 Current Study

Drawing on the existing literature, exploring the dynamics of interactions within MOOC forums holds the potential to provide valuable insights into the pedagogical value and opportunities afforded by MOOCs. Consequently, the primary objective of this study is to examine the effects of instructional interventions on the discussions within a MOOC forum, with a specific focus on investigating potential improvements in student participation and interaction characteristics.

2 Research Design

2.1 Research Questions

Based on the research purpose, two research questions were formulated:

1. To what extent can instructional intervention improve students' participation in the discussion forum?
2. What are the characteristics of instructor-learner interactions and learner-learner interactions resulting from instructional intervention?

2.2 Participants

The present study analyzed the effects of instructional intervention in two instances of an online course on *Internet + College English* (Round 1 and Round 2). The courses were conducted on the Chinese MOOC platform, iCourse (www.icourses.cn), in spring 2020 (Round 1) and autumn 2021 (Round 2) respectively. Round 1 had 3,151 registered students, while Round 2 had 5,202 registered students. In Round 1, students received a carefully designed intervention that incorporated both asynchronous and synchronous online instruction to foster their active participation in the online discussion process (Table 1). In contrast, students in Round 2 did not receive this intervention.

Table 1. Summary statistics for forum participation in Round 1 and Round 2.

Sub-forum	# Parts	# Teachers	# Sub-threads	#Posts	Posts/user
Round 1	380	2	53	601	601/380
Round 2	60	2	2	63	63/61

2.3 Research Instrument

A mixed method of social network analysis (SNA) and inductive qualitative analysis was adopted to analyze the relationships and the underlying interactions they represent in discussions in a EFL MOOC. Ten edgelists were generated to represent learner interactions in the MOOC forum, utilizing Wise and Cui's (2018) five distinct tie definitions for the content-related discussion networks (see Fig. 1).

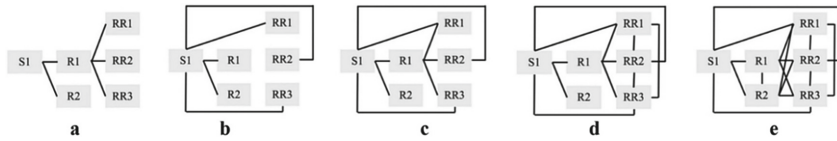


Fig. 1. Wise & Cui's (2018) Ties Extraction Method. Note. (a) Direct Reply; (b) Star; (c) Direct Reply + Star; (d) Limited Copresence; (e) Total Copresence. S = thread starting post; R = reply post; RR = reply to reply post. Solid lines represent ties extracted using this definition.

2.4 Procedure

2.4.1 The Proposed Instruction Intervention Approach on MOOC Forum Discussion

Acknowledging the potential benefits of the instructor's role in online instruction and the significance of well-designed tasks and activities for instructional intervention (Wise & Cui, 2020), this paper proposes a judicious combination of asynchronous and synchronous online instruction to actively involve students in the online discussion process and promote their engagement (see Fig. 2).

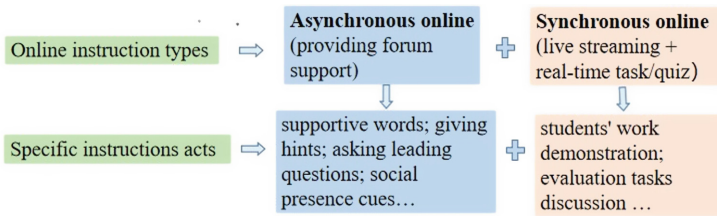


Fig. 2. The proposed intervention in online instructions

Figure 2 depicts actions for asynchronous and synchronous online instruction. In asynchronous instruction, the focus is on providing forum support through feedback that uses encouraging language and provides hints rather than direct answers. Supportive words and social presence cues create a sense of instructor support. Asynchronous discussions guide students in case analysis, evaluation, proposing solutions, and evaluating tasks to improve understanding. Synchronous instruction involves pre-class guidance through live streaming media (e.g., Zoom) and real-time quizzes (e.g., Kahoot!). This creates a virtual classroom for real-time interaction. Instruction intervention includes tasks, oral and written discussions, peer evaluation, and real-time quizzes. This design benefits student engagement and learning performance while enabling teachers to track progress and adjust instruction effectively (Chen et al., 2021a, 2021b, Alonso, et al., 2005).

2.5 Data Collection and Analyses

The teaching materials utilized in this study were extracted from an open online course called “Internet + College English” offered on the iCourse platform in China (<https://www.icourse163.org/>). The course comprised five modules encompassing technology-enhanced listening, speaking, reading, writing, and translating skills. Spanning a duration of twelve weeks, the course specifically focused on leveraging technology to enhance language learning. For the purpose of this study, Module Two, which centered around “technology-enhanced speaking skill drilling,” was selected. Within this module, one specific discussion topic was chosen for analysis: “What will China be like in 30 years? Write a mind-map on the topic and share your answer here. Let’s share, discuss, and learn together.” A total of 55 threads containing 657 discussion posts were collected from 440 learners enrolled in the MOOC.

2.5.1 Thread Classification

Out of the 657 posts that were analyzed, all of them were found to be relevant to the discussion topic. To identify non-content-related posts, we employed a binary classification method based on the approach proposed by Wise et al. (2017). Content-related starting posts encompassed activities such as seeking or providing help, sharing course-related information or resources, asking or responding to subject-related questions, expressing or commenting on subject-related ideas, or sharing external resources. On the other hand, non-content-related posts did not meet these criteria. Following Wise and Cui’s framework (2018), the post types were further categorized as thread-starting, reply, and reply-to-reply. After excluding one repeated post and nine non-content-related posts, the corpus comprised 647 content-related posts from 435 users out of 8,349 MOOC learners. In Round 1, there were 53 thread-starting posts, 295 reply posts, and 397 reply-to-reply posts (see Table 2). In Round 2, there were two thread-starting posts, 61 reply posts, and two reply-to-reply posts (see Table 2).

Table 2. Data information of the present study.

Discussion Topics	Content-related Posts			
	Forum	Thread-starting	Reply	Reply-to-reply
Sharing: My mind map on the topic <i>What will China be like in 30 years?</i>	Round 1	53	295	397
	Round 2	2	61	2

2.5.2 Network Construction and Network Properties

Ten nodelists were generated to represent the network of MOOC forum participants, based on ten edgelists connecting 440 learners’ IDs with 664 content-related posts. These nodelists and edgelists were imported into Gephi 0.9.2 for Windows to construct undirected weighted networks. The Rotate layout algorithm was used for visualization. For

each of the five networks corresponding to the five tie definitions, the number of edges, average node degree, average edge weight, and graph density were computed (Wise et al., 2018).

3 Result and Discussion

3.1 Analysis of Student Engagement in Online Discussion Forums

Student engagement in online discussion forums was analyzed based on the number of participants and the frequency of their participation. Table 3 shows that, for the same original post, in Round 1, both the teacher and the students participated more compared to Round 2. In Round 1, the total number of participants was 382, including 2 teachers, while in Round 2, there were 61 participants, including 2 teachers (see Table 3).

Table 3. Students engagement in terms of the number of participants in the two instances.

Rank	Round 1(N = 382)	Round 2(No = 61)
1	T1 (285)	T1(60)
2	S1 (56)	S2 (2)
3	S2 (41)	S6 (2)
4	S5 (33)	S1 (1)
5	T2 (24)	S3 (1)
6	S3 (20)	S4 (1)
7	S4 (19)	S5 (1)
8	S6 (16)	S7 (1)
9	S13 (13)	S8 (1)
10	S15 (12)	T2 (1)

Among the top 10 lists (ranked by degree) in the Round 1 and Round 2 networks, there were sixteen distinct learners. Apart from T1 and T2, who appeared on the high-degree lists for both networks, the remaining 16 learners exhibited high degrees in either one network or the other, but not in both. Although they were labeled as S1, S2, and so on in both instances, it is important to note that they represent different individuals (see Table 3). This signifies that the top players in the two networks were distinct individuals.

When examining the frequency of participation, the data indicated that in Round 1, the two instructors participated significantly more frequently compared to Round 2. T1 ranked Number 1 with a degree of 285 (see Table 3) in Round 1, and T1 also ranked Number 1 with a degree of 60 (see Table 3) in Round 2. Regarding the high-frequency participants, in Round 1, there were eight students who ranked among the top frequency participants. All of them had degrees higher than 10 (see Table 3), with S1 having a degree of 56, S2 with a degree of 41, and S5 with a degree of 33 (see Table 3). However, in Round 2, the top two participants had a degree of only 2, while the rest of the participants all had a degree of 1.

3.2 Analysis of the Community in Round 1

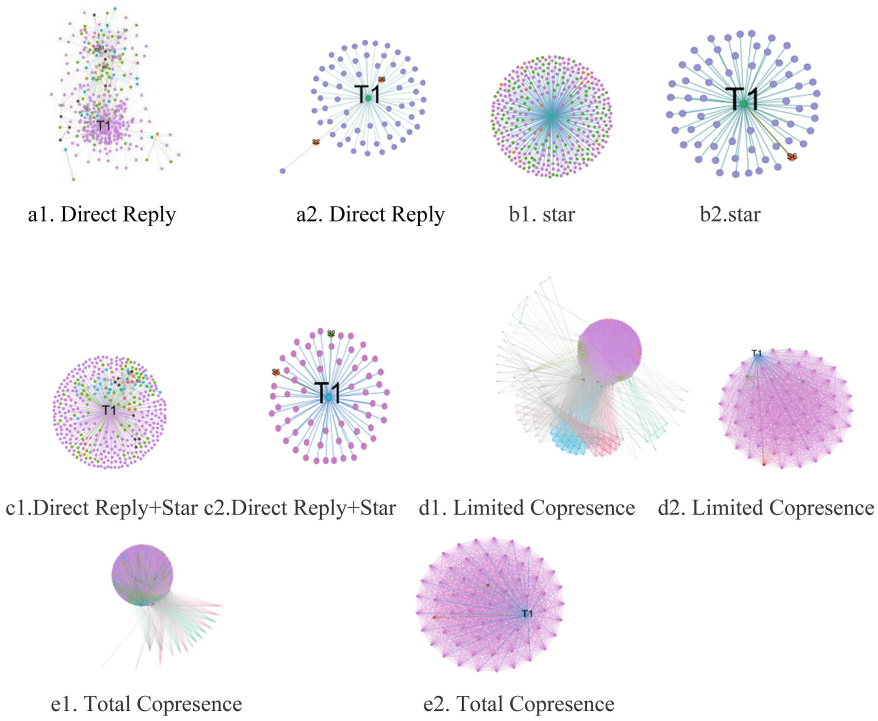


Fig. 3. Visualized graphs of the network

Table 4. Network measure of five networks.

Tie Definition	Round1 (N = 382)				Round 2 (N = 61)			
	#of edges	Avg node degree(SD)	Avg edge weight(SD)	Graph density	#of edges	Avg node degree(SD)	Avg edge weight(SD)	Graph density
DR	522	2.945 (15.14)	1.08 (0.286)	0.007	62	2 (7.683)	1.016 (0.127)	0.032
S	382	3.52 (34.353)	1.764 (0.514)	0.005	62	4 (15.62)	2.032 (0.254)	0.032
DR + S	633	4.971 (34.579)	1.504 (0.554)	0.009	63	4 (15.493)	2 (0.311)	0.032
LC	40231	217.274 (119.686)	1.034 (0.254)	0.55	1953	66.095 (15.743)	1.066 (0.368)	1
TC	67456	357.133 (78.69)	1.014 (0.154)	0.922	1953	66 (15.493)	1.065 (0.358)	1

The analysis of the five types of social network analysis (SNA) ties corresponds to content-related discussion posts derived from a main post, which can be considered a

micro-level analysis within SNA. In Round 1, this main post generated over 50 sub-threads, more than 500 discussions, and engaged over 380 students in the discussion (see Table 4). This indicates a substantial level of participation from both the number of participants and the extent of students' involvement, showcasing their active engagement.

Instructor T1 actively participated in the interactive discussion, providing guidance for forum discussions through synchronous meetings held prior to class. With numerous participants and high frequency of participation, there were both original posts and sub-thread posts. The characteristics of the five types of ties observed in such a large-scale and high-frequency discussion involving multiple participants and posts are described as follows: Direct Reply (DR) serves as a model that objectively illustrates the hierarchical connections among original posts, replies, and replies to replies.

Regarding the Direct Reply tie, the figure clearly illustrates that the overall forum posts revolve around the original post by T1 (see Fig. 3 a1). Subsequently, with an abundance of replies and content, additional top posts have emerged, most of which have received comments from T1. Consequently, many students have followed the instructor's lead and actively participated in replying to the top posts, thereby fostering interaction among students themselves. Notably, the degree of interaction for participants such as ST1 and ST2 (see Table 5) reached 56 and 41, respectively.

Table 5. Subthread with high interaction degree

Subthread(ST)	(T1/T2) Instructor degree	Reply	Participation(%)
ST1	+	56	14.7%
ST2	+	41	10.8%
ST5	+	33	8.7%
ST3	+	20	5.2%

Star: The number of students engaging in interactions with teachers is substantial (refer to Fig. 3 b1), and their distribution is uniformly spread around the teachers, forming a sphere consisting of equidistant points centered on the teachers. This observation indicates that, within this discourse, students predominantly interact with teachers. Moreover, students in closer proximity to T1 exhibit a higher frequency of interactions with teachers, whereas students located further away from T1 demonstrate comparatively fewer interactions with teachers.

Limited Copresence & Total Copresence: Owing to the significant number of participants within the same post, as defined by copresence ties, the graph exhibits a high level of density (refer to Fig. 3 d1, e1). However, this density does not appear to hold substantial relevance for analyzing the network centered around a single original post. Our focus in this research revolves primarily around examining the dynamics of teacher-student and student-student interactions under this particular post. Consequently, these two modes of interaction do not seem to offer distinct advantages in our study.

In this paper, the author contends that when analyzing the aforementioned interactions under a main post, the most effective approach is to employ Direct Reply ties

to illustrate the patterns of interaction. This perspective diverges from Wise & Cui's claim (2018), who argue that DR + Star can adequately depict the nature of interaction. Nonetheless, the data presented in this paper demonstrate that, in comparison to DR + Star, employing the DR tie alone enables a clearer representation of high-frequency participants, their frequency of interaction, the distinctive community characteristics that arise from interactions among these high-frequency participants, as well as the formation of different communities fostered by various high-frequency participants.

3.3 Analysis of the Community in Round 2

The data revealed a hub-and-spoke structure with limited learner-learner interactions, where the instructor served as the central node (see Fig. 3 a2–e2). Despite a diverse student population and no online pre-class guidance provided by the instructors, the main post generated two sub-posts, over 60 discussions, and the participation of more than 60 students, primarily from the larger society rather than the author's university.

The analysis of the Fig. 3 (a2–e2) demonstrated a consistent hub-and-spoke structure, with minimal variation. The tie maps showed limited interaction among participants, with only two direct replies observed. The similarity among the five tie maps indicated that using different types of ties had little significance in analyzing a discussion network with limited interaction. Previous studies also highlighted the instructor-centric nature of learner interactions in MOOCs, resembling the hub-and-spoke structure observed here (Smith et al., 2019; Johnson, 2020, Brooks et al., 2021)). In Round 2, a similar hub-and-spoke structure was confirmed, further emphasizing the influential role of instructors due to their position and status in the course.

In the upcoming research phase, the author plans to conduct a qualitative analysis of the teacher's intervention in the forum. This analysis will explore strategies such as questioning, guidance, and encouragement employed by teachers to promote student participation in online discussion forums.

3.4 Instructors in the Two Communities

3.4.1 Network Structure

Social network analysis revealed differences in interactions surrounding instructors (T1 & T2) between Rounds 1 and 2 (see Table 6). In Round 1, there were six times more nodes and eight times more edges compared to Round 2, indicating a larger and more interconnected community formed around instructor T1. The degree of T1 in Round 1 was over four times higher than in Round 2, suggesting direct interactions with a greater number of learners. Both communities showed a significant proportion of learners solely connected to the instructor. However, Round 1 had higher interconnections among learners compared to Round 2, as indicated by a higher average node degree. Additionally, Round 1 had a slightly higher average edge weight, implying more repeated interactions with the same individuals.

Table 6. Network structures of the two instances (ties based on direct reply).

	Round 1	Round 2
# of nodes	383	60
# of edges	522	62
Graph density	0.007	0.032
Avg node degree (SD)	2.945 (15.14)	2 (7.683)
Avg edge weight (SD)	1.08 (0.286)	1.016 (0.127)
Instructor degree	285	60

3.4.2 Communication Techniques

In Round 1, qualitative analysis identified distinct communication techniques employed by instructors T1 and T2. T1 used words of encouragement, offered hints, asked leading questions, and incorporated social presence cues. They motivated learners with phrases like “Good job!” and “Great! I love your idea of ‘10G’ ~ Great ~” to inspire learners, utilizing hints to guide independent problem-solving. T1 also employed greetings, addressed learners by name, and echoed their words to encourage expression of ideas.

Similarly, in Round 1, T2 frequently provided positive feedback and utilized social presence cues. For example, they praised clear outlines with comments like “Good job! Your outline is quite clear ~”.

To summarize, in Round 1, both instructors made more posts, responded more frequently, and employed distinct communication techniques including supportive words, hints, questions, and social presence cues. In contrast, in Round 2, T1 posted less frequently and used fewer social presence cues. These differences in instructor participation may contribute to the disparities observed in the network structures formed during interactions.

4 Conclusion and Implication

The research findings highlight several key points. Firstly, instructional interventions effectively enhance student participation, promoting both instructor-learner and learner-learner interactions. Secondly, instructional intervention can alter the predominant hub-and-spoke structure in MOOC forums, allowing learners to form their own learning communities. Thirdly, network analysis demonstrates that defining direct reply ties offers the clearest depiction of interaction relationships.

This study contributes to the literature by: (1) Revealing characteristics of instructor-learner and learner-learner interactions in Chinese MOOC discussions. (2) Identifying the correlation between instructional interventions and student engagement. (3) Providing insights for instructors on fostering interactive learning communities in MOOC forums. Teachers and teaching assistants are recommended to participate in targeted

and planned ways to support students' independent learning in the era of intelligent education.

Jung and Lee (2018) emphasized the importance of instructors establishing clear rules for participation and expected outcomes in MOOCs. This involves defining learning goals, discussion topics, assessment criteria, and learning expectations. Wang (2022) suggests that providing clarity on how students should participate in learning activities and what is required to fulfill the learning tasks enhances engagement.

In terms of blended instruction, teachers should assume guiding roles and maintain active involvement in the learning process. They can offer training and support to students in utilizing technology, address obstacles that arise during blended learning, and ensure students stay on the right path towards achieving their learning goals more effectively (Tang et al., 2022).

5 Future Research Directions

The present study serves as a pilot investigation on the impact of instructional interventions, featuring asynchronous and synchronous online instruction, to engage students and enhance interaction. However, it is regrettable that the effects of this intervention on learning outcomes were not available at the time. Future research is recommended to compare the teaching effects of such interventions in distance learning settings. It is important to note that this study primarily aimed to provide objective behavioral validation of the Online Social Engagement (OSE) through students' behaviors rather than solely relying on self-reported data. It does not attempt to measure or claim to measure all learning that occurs in an online course.

Another limitation of the present study is the lack of analysis on individual small learning communities within the data. Consequently, the actual impact of instructional interventions on student engagement was not fully observed in this study.

Acknowledgments. The author would like to express sincere gratitude to the anonymous reviewers and Professor Chen Zexuan from Southern Medical University for their invaluable feedback on earlier versions of this paper. This paper is funded by the Education Science Planning Project in Guangdong Province in 2022 (Higher Education Project) (Grant No. 2022GXJK219).

References

- Alonso, F., Lopez, G., Manrique, D., Vines, J.M.: An instructional model for web-based e-learning education with a blended learning process approach. *Br. J. Edu. Technol.* **36**(2), 217–235 (2005)
- Brooks, C., Thomasson, H., Caverly, D.: The pedagogical role of forums in MOOCs: a study of massive open online course discussion forums. *Internet High. Educ.* **48**, 100777 (2021)
- Chen, Y.H., Hong, J.C., Chiou, S.K.: Effects of flipped classroom and online discussion teaching strategies on college students' learning outcomes, satisfaction, and engagement. *Interact. Learn. Environ.* 1–16 (2021)
- Chen, Z., Jiao, J., Hu, K.: Characteristics of learner interaction and network construction in MOOC forum discussion. In: *ICOIE 2019*, pp. N/A. Hong Kong (2019)

- Chen, Z., Jiao, J., Hu, K.: Formative assessment as an online instruction intervention: student engagement, outcomes, and perceptions. *Int. J. Distance Educ. Technol. (IJDET)* **19**(1), 50–65 (2021)
- Gillani, N., Eynon, R.: Communication, collaboration and community: a critical analysis of the use of social media in higher education. In: Newman, E.M., Eidson, C.T., Henry, R.M. (Eds.) *Communication Across the Life Span*, pp. 177–194. Taylor & Francis (2019)
- Johnson, G.M., Veletsianos, G., Seaman, J.: U.S. faculty and administrators' experiences and approaches in the early weeks of the covid-19 pandemic. *Online Learn.* **24**(2), 6–21 (2020)
- Johnson, N.F.: Inside online learning: examining learner interactions in an asynchronous online discussion forum. *Internet High. Educ.* **44**, 100718 (2020)
- Jones, M., Smith, K.: Exploring the relationship between student motivation, engagement, and retention in a MOOC: a comparative, longitudinal study. *Internet High Educ.* **33**, 60–70 (2017)
- Jung, I., Lee, J.: Effects of different types of interaction on learning achievement, satisfaction, and participation in web-based instruction. *Innov. Educ. Teach. Int.* **55**(3), 338–346 (2018)
- Lee, Y., Choi, I., Kim, T.: Instructor immediacy in massive open online courses (MOOCs): an analysis of learner experiences and perceptions. *Comput. Educ.* **137**, 36–47 (2019)
- Reeve, J.: A self-determination theory perspective on student engagement. In: Christenson, S.L., Reschly, A.L., Wylie, C., (Eds.) *Handbook of Research on Student Engagement*, pp. 149–172. Springer (2012)
- Smith, M.S., Jones-Farmer, A., Williams, B.D., Maben, S.: The impact of prior knowledge on interactions in a MOOC discussion board. *Comput. Educ.* **141**, 103605 (2019)
- Smith, R.L.: Interaction and learning. In: Seo, K.K. (Ed.) *Encyclopedia of Information Science and Technology* (4th ed.), pp. 1731–1740. IGI Global (2018)
- Tang, S., Zhang, X., Chen, Z., Qi, M., Xu, T.: Exploring factors influencing college students' engagement in blended learning: a systematic review and meta-analysis. *Distance Educ.* **43**(4), 540–563 (2022)
- Wang, Q., Woo, H.L., Quek, C.L., Yang, Y., Liu, M.: Using the ARCS model to examine MOOC learner dropout and retention: perspectives of learners and instructors. *Educ. Tech. Res. Dev.* **68**(5), 2201–2222 (2020)
- Wang, Y.: A framework for assessing the quality of MOOCs. In: *Handbook of Research on Cross-Cultural Online Learning in Higher Education*, pp. 120–135. IGI Global (2022)
- Wise, A.F., Cui, Y., Hausknecht, S.: Learning analytics for online discussions: design approaches and tools to reveal patterns of student engagement. *J. Educ. Technol. Soc.* **20**(3), 180–192 (2017)
- Wise, A.F., Cui, Y.: Discourse-centric learning analytics. In: Huang, R., Díaz-Morales, K.J., Ruiz-Shulcloper, J.A. (Eds.) *Intelligent Data Engineering and Automated Learning – IDEAL 2018*, pp. 168–176. Springer (2018)
- Wise, A.F., Cui, Y.: Examining the potential of learning analytics in computer-supported collaborative learning. *Int. J. Comput.-Support. Collab. Learn.* **13**(1), 7–34 (2020)
- Wu, T.Y.: Systematic analysis of online discussion patterns and participant roles in a massive open online course using social network analysis. *Interact. Learn. Environ.* **29**(4), 489–506 (2021)

Personalized and Individualized Learning



Emerging Techniques for Online Learning Analytics

Yidan Wang¹(✉), Xiaoran Huang², Qijing Yu³, and YuXuan Lai¹

¹ School of Computer Science, The Open University of China, Beijing, China
wangyidan@ouchn.edu.cn

² School of Architecture and Arts, North China University of Technology, Beijing, China
xiaoran.huang@ncut.edu.cn

³ Academic Affairs Department, The Open University of China, Beijing, China

Abstract. The rapid growth of online learning platforms presents both challenges and opportunities for educational institutions. Online learning analytics has the potential to extract valuable insights from student learning behaviours, and such insights could offer learners flexible access to educational resources and foster life-long learning opportunities. This paper reviews the research progress and emerging techniques in online learning analytics. We first outline the objectives of learning analytics, emphasising its potential to enhance personalised learning experiences, and examine key challenges faced in such platforms. We also review state-of-the-art techniques and explore emerging trends, natural language processing and multimodal learning analytics in particular. The paper discusses the implications of online learning analytics for educators and instructional designers, identifies open research challenges, and inspires further research and innovation to support learners in their educational journey.

Keywords: learning analytics · student engagement · online learning · natural language processing · multimodal data

1 Introduction

There has been a significant transformation in education due to the rapid rise of online learning, and online platforms offer flexibility and accessibility to both educational institutions and learners. As the digital learning environment continues to evolve, student learning experiences are being enriched in many ways. For instance, data-driven insights obtained from online learning data can assist instructors in predicting student achievements. This paper reviews existing studies and emerging techniques for online learning analytics, exploring how cutting-edge technologies are shaping our understanding, assessment, and enhancement of online education.

The growth of online learning has been accelerated by the urgent need for flexible and remote learning solutions. Consequently, the accumulation of digital footprints within online learning platforms has generated a substantial amount of data for exploration. Learning analytics refers to the systematic analysis of such data, which has emerged as

a significant tool for educators, institutions, and policymakers seeking valuable insights to make informed decisions.

While the foundation of learning analytics has been established over the years, emerging techniques are revolutionising the field. These techniques leverage advancements in artificial intelligence and machine learning, offering unprecedented opportunities to delve deeper into learner behaviours, motivations, and interactions. By applying these innovative methodologies, educators and institutions can uncover hidden patterns, tailor interventions, and ultimately enhance the efficacy of online education.

This paper is organised as follows: we introduce the background and importance of online learning analytics are provided in the rest of this section. The literature review is conducted using CiteSpace and VOSViewer, with visualisations presented in Sect. 2. Section 3 delves into specific techniques, particularly natural language processing and multimodal learning analytics. The challenges associated with the adoption and implementation of these techniques are also examined, along with the directions that require further investigation, discussed in Sect. 4. We conclude this work in Sect. 5.

1.1 Background

The rapid growth of technology has accelerated the integration of online learning platforms into traditional educational settings. This transformation has facilitated data-driven insights and improved students' learning outcomes. Online learning analytics has emerged as a critical tool for enhancing the effectiveness and enriching the experiences of online learning.

Learning analytics aims to comprehend and enhance online learning by measuring, collecting, analysing and reporting learners' data recorded in a digital learning environment (Siemens & Baker, 2012). The roots of online learning analytics can be traced back to the early days of computer-assisted learning in the 1960s and 1970s. However, it wasn't until the widespread application of the Internet and learning management systems (LMS) in the late 20th century that online learning analytics drew the attention of researchers.

The advent of massive open online courses (MOOCs) and the expansion of online degree programs generate substantial data from diverse online learners. Such data enable researchers and instructors to understand how students engage with course content and collaborate with peers. While traditional education primarily relied on limited data to inform instructional practices, the emergence of online learning platforms addressed digital traces including keystrokes, click-streams, and discussion forum posts (Sharma et al., 2020). This shift led to a greater emphasis on evidence-based and data-driven decision-making in education.

Learning analytics initially focused on providing instructors with students' log data that records their login frequency, messages viewed or posted on discussion forums or the number of clicks the student has made for specific resources in LMS (Romero et al., 2013). While data mining algorithms enable us to identify certain patterns from a huge volume of data, researchers and educators sought to improve their understanding of the learning process by discovering patterns from students' log data (Siemens & Gašević, 2012). Consequently, predictions about student success or identification of students at risk of dropping out or underperforming became achievable.

Efficient utilisation of learning analytics makes it possible to identify learner preferences, adjust content delivery, and recommend resources tailored to individual needs. As online education continues to expand, institutions and educators are recognizing the potential of learning analytics to enhance online learning experiences.

1.2 Importance of Online Learning Analytics in Education

As online education gains prominence, understanding and harnessing the power of learning analytics have become imperative for realising the full potential of digital learning environments.

Online learning analytics empowers educators with insights that were previously inaccessible in traditional classrooms. More specifically, learning analytics is critically important in understanding learning and promoting evidence-based practices (Mangaroska & Giannakos, 2018). By analysing student interactions with online content, educators can gain a comprehensive understanding of how learners engage with materials, which resources are most effective, and where learners encounter challenges. Instructors can also tailor their teaching strategies, modify content delivery, and identify areas that require additional support based on the derived insights. This personalised approach enhances the effectiveness of instruction and leads to improved learning outcomes.

The diversity of learners in online education requires a more adaptable and personalised approach to instruction. Learning analytics enables the creation of tailored learning pathways that cater to individual strengths, weaknesses, and preferences.

Institutions and administrators benefit significantly from online learning analytics as well. The data-driven insights provided by analytics platforms guide strategic decision-making at the institutional level. By aggregating and analysing data on course success rates, program effectiveness, and learner demographics, institutions can optimise resource allocation, refine curriculum design, and identify areas for improvement.

2 Literature Review Using VOSviewer and CiteSpace

VOSviewer and CiteSpace were employed to conduct a systematic literature review of online learning analytics. As tools for bibliometric analysis and visualisation, both VOSviewer and CiteSpace facilitate the exploration of scholarly publication networks and create insightful visualisations that highlight the interconnectedness of research topics as well as emerging trends.

2.1 Data Collection

The literature review was conducted by importing a collection of relevant scholarly publications obtained from Web Of Science database. The search queries encompassed a wide range of keywords and phrases related to online learning analytics, including educational analytics, adaptive educational systems, personalised learning, and educational data mining. The research area was limited to computer science or educational research. We further filtered the query's outcome and refined the publication year from 2014 to 2023. As a result, a total of 3,486 publications were included as the most relevant publications for analysis and visualisation purposes.

2.2 Visualisation of Analysis Using VOSviewer and CiteSpace

2.2.1 Research Clusters and Themes

The co-occurrence network unveils distinct research clusters by analysing the co-occurrence of keywords in the literature. The visualisation of the network uncovers prevalent themes and topics within the field. The derived clusters enable users to have a structured understanding of the multidimensional facets of online learning analytics. Figure 1 shows the co-occurrence of keywords from the literature, and it is observed that four clusters have been identified in previous studies.

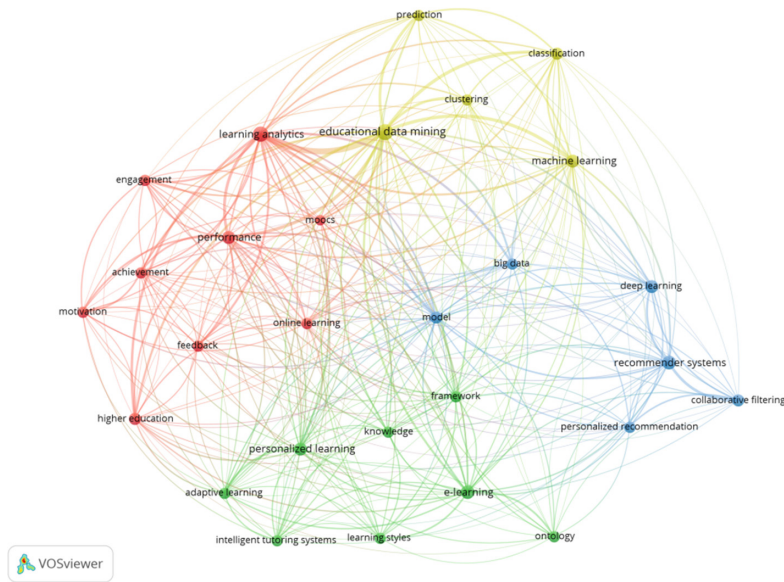


Fig. 1. The co-occurrence network of keywords in online learning analytics literature

Specifically, education data mining is one of the most popular topics in the field, while clustering, classification and other machine learning algorithms are being widely applied to explore the valuable insights from educational data. Meanwhile, learning analytics, closely related to education data mining, has drawn extensive attention from researchers. Online learning behaviour can be considered as a critical indicator of student engagement, learning outcomes and performance. MOOCs are key sources of educational data, and learning analytics benefits online learning and higher education in numerous ways. It is worth noting that students' feedback should be addressed in the analysis of educational data. Deep learning has also received extensive study to enhance personalised recommendation systems. Additionally, personalised learning, adaptive learning, and profiling student learning styles are themes that are closely associated with e-learning systems.

2.2.2 Influential Authors and Research Organisations

The co-authorship network unveils authors and research organisations that have collaborated extensively or significantly contributed to the advancement of the field. Figure 2 provides an overview of the co-authorship network of authors. Through analysis of the co-authorship network, we have identified several influential authors and prolific research groups. The colours assigned to authors indicate scores based on the average publication year of their studies.

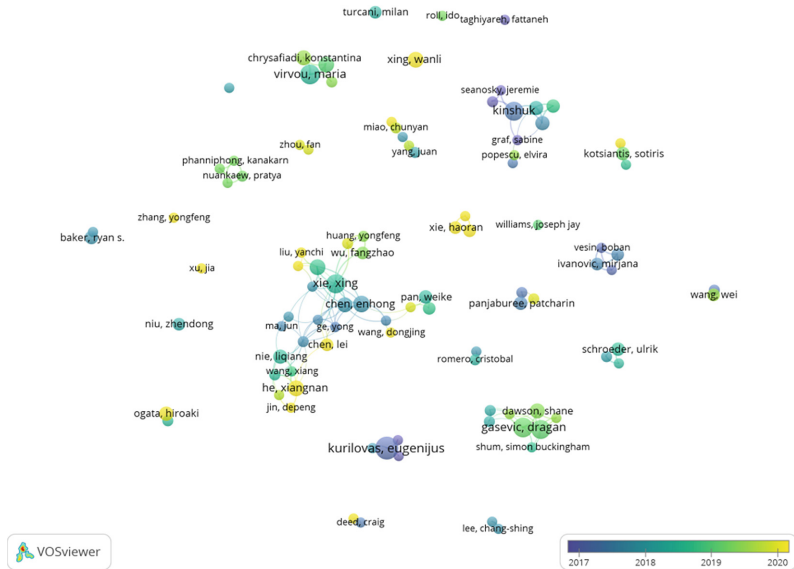


Fig. 2. The co-authorship network of authors in online learning analytics literature

2.2.3 Emerging Trends

By analysing the keyword co-occurrence in the literature spanning from 2014 to 2023, we have identified emerging trends and shifting research emphasis in online learning analytics over time. Figure 3 presents the top 20 keywords with the strong citation bursts, highlighting the evolution of key concepts and themes within the field. While learning styles, intelligent tutoring systems and data mining were recognized as trends driving increased attention and citation before 2018, researchers have increasingly explored the application of deep learning in online learning systems since 2021. Additionally, there is a growing interest in sentiment analysis, reinforcement learning, graph neural networks and personalised recommendation systems.

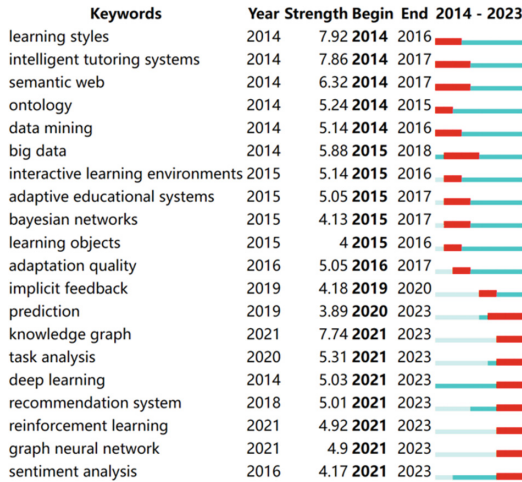


Fig. 3. Top 20 Keywords with the strongest citation bursts

3 Emerging Techniques for Online Learning Analytics

The swift expansion of online learning has accelerated the exploration of innovative techniques within the field of learning analytics. By providing deeper insights into student behaviour, engagement, and performance, these emerging approaches could potentially enrich students’ online learning experiences. This section delves into the forefront of emerging techniques and explores their practical implementations in the domain of online learning analytics.

3.1 Natural Language Processing

Natural language processing (NLP) aims to enable computers to understand natural language texts and conduct relevant tasks with the given texts. It gives insights from textual interactions within digital environments. By leveraging NLP’s capabilities in online learning analytics, researchers and instructors could potentially understand students’ sentiments, estimate engagement levels, comprehend their online learning experiences, and facilitate automated essay scoring.

Through the examination of students’ posts in discussion forums of MOOCs, positive or negative emotions could be identified (Kumar et al., 2020). Such emotions extracted from students’ sentiments are key indicators of their attitudes toward the course and overall success (Crossley et al., 2016, Robinson et al., 2016). Click-stream logs in MOOCs record students’ online behaviours, including the time or counts that students interact with certain contents. These data logs are useful in understanding students’ completion and engagement in MOOCs (Boyer & Veeramachaneni, 2015, He et al., 2015, Sharma et al., 2015, Seaton et al., 2014). While a single source of data may fail to capture students’ differences, Crossley et al. (2016) studied click-stream data and combined NLP indices to predict MOOC completion. Moreover, Robinson et al. (2016) developed a model that predicts students’ achievement using NLP from unstructured text responses.

Students' posts in discussion forums could also enable the analysis of interactions among students or enhance the communication between instructors and students. Specifically, Wang and Xu (2023) established collaborative relationship networks and investigated the impact of social, topic and attribute relationships of learners according to the data indicating their behaviours and generative texts. Besides, NLP-driven chatbots empower online learning platforms by offering students instant assistance or guidance. The personalised nature of communication with such chatbots makes the learning environment more engaging. For instance, González-González et al. (2023) proposed a chatbot architecture and integrated it into a Gamified Learning System to engage students.

Furthermore, NLP ensures fairness and objectivity in essay scoring by analysing the structure, coherence, grammar and content of writing assessment. Personalised feedback is also available to assist learners in improving their writing. NLP plays an essential role in providing efficient and consistent assessment of writing assignments (Gardner et al., 2021). Pribadi et al. (2017) investigated short answer scoring methods, while Taghipour and Ng (2016) explored neural network models in automated essay-scoring systems.

NLP transforms online learning analytics by facilitating sentiment analysis, comprehending collaborative relationships and automated assessment. In particular, online learning experiences could be further enriched by crafting personalised learning journeys with the continued incorporation of NLP in digital learning environments.

3.2 Multimodal Learning Analytics

While various types of data are available in an online learning environment, including click-streams, keystrokes, or posts in forums, researchers and educators extend conventional analytics by integrating diverse data modalities. This holistic approach captures students' learning experiences from different perspectives and addresses their interaction with varied content formats (Mitri et al., 2018).

With the development of technologies capturing data from human activity, tracking learners' behaviour through biosensors, wearable cameras, eye tracking or online activity logging has become possible. Comprehensive insights could be obtained from these data collection and sensing techniques, describing learners' activities from multiple dimensions (Blikstein & Worsley, 2016). Sharma et al. (2020) considered eye-tracking data to profile students and predict student achievements. Moreover, several studies investigated eye movement or eye contact in learning processes to motivate students or understand cooperative relationships (Sharma et al., 2019, D'Angelo & Begel, 2017, Schneider et al., 2018, Ding et al., 2017).

Apart from learning behaviour data and facial expression data, researchers also explored physiological data during the learning process. Specifically, Mills et al. (2017) detected cognitive load by measuring electroencephalograms during learning. In terms of emotion recognition, studies have been conducted by monitoring skin temperature, blood volume pulse or breath respiration (Yin et al., 2017, Mu et al., 2020).

The wide range of modalities provides instructors with profound insights into learners' cognitive processes, attention patterns and learning preferences. While multimodal learning analytics can be considered an essential component in devising individualised

and effective instructional strategies, addressing the fusion of multimodal data is crucial to facilitating the development of multimodal learning analytics (Mu et al., 2020).

4 Challenges and Future Directions

As analytics delve deeper into learners' data, it is critical to ensure data privacy and ethical considerations become paramount. Striking a balance between data collection and respecting individual privacy issues is essentially required. Moreover, there is a challenge in effectively integrating emerging techniques into pedagogical practices. Educators need to understand how to interpret analytics insights and translate them into actionable strategies that improve learning outcomes.

Implementing sophisticated analytics techniques at scale requires significant computational resources, infrastructure, and expertise. Institutions should address resource limitations while ensuring that analytics initiatives can accommodate larger student populations without compromising the quality of insights.

Future directions in online learning analytics involve enhancing the transparency and interpretability of AI-driven insights. Developing models that provide explanations for their predictions can empower educators to make informed decisions and understand the rationale behind analytics recommendations.

Coupled with adaptive learning systems, real-time insights can enable on-the-fly adjustments to content, pacing, and interventions in online education, providing students with personalised and dynamic learning experiences.

As online education becomes more prevalent, longitudinal analysis tracking learners' journeys over time is gaining importance. This includes understanding how online learning experiences influence lifelong learning, career trajectories, and continuous skill development.

5 Conclusion

In this paper, we review the research progress and emerging techniques of online learning analytics. Keyword co-occurrence networks and research trends are visually presented using CiteSpace and VOSViewer. We also delve into natural language processing and multimodal learning analytics and discuss their application and implementation in learning analytics. The integration of emerging techniques in online learning analytics is challenging. Challenges are addressed related to data privacy, integration with pedagogy, and resource constraints. We discuss the implications of online learning analytics for educators and instructional designers, identify open research challenges, and inspire further research and innovation to support learners in their educational journey.

Acknowledgement. The authors acknowledge the support of the 2022 Open University of China Youth Research Grant (Q22A0009).

References

- Blikstein, P., Worsley, M.: Multimodal learning analytics and education data mining: using computational technologies to measure complex learning tasks. *J. Learn. Analytics* **3**(2), 220–238 (2016)
- Boyer, S., Veeramachaneni, K.: Transfer learning for predictive models in massive open online courses. In: *Artificial Intelligence in Education: 17th International Conference, AIED 2015, Madrid, Spain, June 22–26, 2015. Proceedings 17*, pp. 54–63. Springer International Publishing (2015)
- Crossley, S., Paquette, L., Dascalu, M., McNamara, D.S., Baker, R.S.: Combining click-stream data with NLP tools to better understand MOOC completion. In: *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, pp. 6–14 (2016)
- D’Angelo, S., Begel, A.: Improving communication between pair programmers using shared gaze awareness. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 6245–6290 (2017)
- Di Mitri, D., Schneider, J., Specht, M., Drachsler, H.: From signals to knowledge: a conceptual model for multimodal learning analytics. *J. Comput. Assist. Learn.* **34**(4), 338–349 (2018)
- Ding, Y., Zhang, Y., Xiao, M., Deng, Z.: A multifaceted study on eye contact based speaker identification in three-party conversations. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 3011–3021 (2017)
- Fan, Y., Matcha, W., Uzir, N.A.A., Wang, Q., Gašević, D.: Learning analytics to reveal links between learning design and self-regulated learning. *Int. J. Artif. Intell. Educ.* **31**(4), 980–1021 (2021)
- Gardner, J., O’Leary, M., Yuan, L.: Artificial intelligence in educational assessment: ‘breakthrough? Or buncombe and ballyhoo?’. *J. Comput. Assist. Learn.* **37**(5), 1207–1216 (2021)
- González-González, C.S., Muñoz-Cruz, V., Toledo-Delgado, P.A., Nacimiento-García, E.: Personalized gamification for learning: a reactive chatbot architecture proposal. *Sensors* **23**(1), 545 (2023)
- He, J., Bailey, J., Rubinstein, B., Zhang, R.: Identifying at-risk students in massive open online courses. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 29, No. 1 (2015)
- Kumar, A., Srinivasan, K., Cheng, W.H., Zomaya, A.Y.: Hybrid context enriched deep learning model for fine-grained sentiment analysis in textual and visual semiotic modality social data. *Inf. Process. Manage.* **57**(1), 102141 (2020)
- Li, K.C., Wong, B.T.M.: Trends of learning analytics in STE(A)M education: a review of case studies. *Interact. Technol. Smart Educ.* **17**(3), 323–335 (2020)
- Mangaroska, K., Giannakos, M.: Learning analytics for learning design: a systematic literature review of analytics-driven design to enhance learning. *IEEE Trans. Learn. Technol.* **12**(4), 516–534 (2018)
- Mills, C., Fridman, I., Soussou, W., Waghay, D., Olney, A.M., D’Mello, S.K.: Put your thinking cap on: detecting cognitive load using EEG during learning. In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, pp. 80–89 (2017)
- Mu, S., Cui, M., Huang, X.: Multimodal data fusion in learning analytics: a systematic review. *Sensors* **20**(23), 6856 (2020)
- Pribadi, F.S., Adji, T.B., Permanasari, A.E., Mulwinda, A., Utomo, A.B.: Automatic short answer scoring using words overlapping methods. In: *AIP Conference Proceedings*, vol. 1818, No. 1. AIP Publishing (2017)
- Robinson, C., Yeomans, M., Reich, J., Hulleman, C., Gehlbach, H.: Forecasting student achievement in MOOCs with natural language processing. In: *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, pp. 383–387 (2016)

- Romero, C., López, M.I., Luna, J.M., Ventura, S.: Predicting students' final performance from participation in on-line discussion forums. *Comput. Educ.* **68**, 458–472 (2013)
- Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., Pea, R.: Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. *Int. J. Comput.-Support. Collab. Learn.* **13**, 241–261 (2018)
- Seaton, D.T., Bergner, Y., Chuang, I., Mitros, P., Pritchard, D.E.: Who does what in a massive open online course? *Commun. ACM* **57**(4), 58–65 (2014)
- Sharma, K., Dillenbourg, P., Giannakos, M.: Stimuli-based gaze analytics to enhance motivation and learning in MOOCs. In: 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT), vol. 2161, pp. 199–203. IEEE (2019)
- Sharma, K., Giannakos, M., Dillenbourg, P.: Eye-tracking and artificial intelligence to enhance motivation and learning. *Smart Learn. Environ.* **7**(1), 1–19 (2020)
- Sharma, K., Jermann, P., Dillenbourg, P.: Identifying styles and paths toward success in MOOCs. In: International Educational Data Mining Society (2015)
- Siemens, G., Gašević, D.: Special issue on learning and knowledge analytics. *Educ. Technol. Soc.* **15**(3), 1–163 (2012)
- Siemens, G., Baker, R.S.D.: Learning analytics and educational data mining: towards communication and collaboration. In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, pp. 252–254 (2012)
- Spikol, D., Ruffaldi, E., Dabisias, G., Cukurova, M.: Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *J. Comput. Assist. Learn.* **34**(4), 366–377 (2018)
- Taghipour, K., Ng, H.T.: A neural approach to automated essay scoring. In: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp. 1882–1891 (2016)
- Wang, C., Xu, Y.: Who will work together? Factors influencing autonomic group formation in an open learning environment. *Interact. Learn. Environ.* 1–19 (2023)
- Wong, B.T.M., Li, K.C.: A review of learning analytics intervention in higher education (2011–2018). *J. Comput. Educ.* **7**(1), 7–28 (2020)
- Yin, Z., Zhao, M., Wang, Y., Yang, J., Zhang, J.: Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. *Comput. Methods Programs Biomed.* **140**, 93–110 (2017)
- Zawacki-Richter, O., Marín, V.I., Bond, M., Gouverneur, F.: Systematic review of research on artificial intelligence applications in higher education—where are the educators? *Int. J. Educ. Technol. High. Educ.* **16**(1), 1–27 (2019)



Prediction of At-Risk Students Using Learning Analytics: A Literature Review

Kam Cheong Li, Billy T. M. Wong^(✉), and Hon Tung Chan

Institute for Research in Open and Innovative Education, Hong Kong Metropolitan University.
Ho Man Tin, Kowloon, Hong Kong SAR, China
{kcli, tamiwong, hotchan}@hkmu.edu.hk

Abstract. Early identification of at-risk students has been recognised as being of the utmost importance for the provision of timely support. Despite the prevalent use of learning analytics in this regard, little attention has been paid to systematically surveying and summarising relevant latest work. To address the literature gap, this paper reviews the use of learning analytics to support prediction of at-risk students. The study covers 138 research articles published between 2013 and 2022 which were collected from the Scopus database. Through a content analysis approach, the relevant work was examined in terms of the prediction objectives, the types of data collected, the techniques used for prediction, and the metrics employed for evaluation of prediction performance. The findings reveal a strong scholarly interest in the prediction of students' learning performance to identify those who are potentially at risk. The common types of data collected are related to students' previous academic performance, socio-demographics, and learning behaviours, particularly those on learning management systems. The most frequently used techniques for prediction are decision trees, neural networks, and Bayesian networks. The results also show the widespread use of classification accuracy, recall, sensitivity, and true positive rate as the performance metrics. The findings contribute to advancing our understanding about the potential of learning analytics for at-risk student prediction, as well as informing the provision of timely and proper support for specific student groups.

Keywords: At-risk students · prediction · learning analytics · student support

1 Introduction

The importance of identifying at-risk students to provide them with timely and appropriate learning support has been recognised in contemporary education. Early identification of at-risk students has been greatly supported by the development of learning analytics in the past decade. Learning analytics refers to “the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” (Long & Siemens 2011, p. 34). With the data related to student learning being systematically managed and analysed, institutions can make predictions about student learning and find out those

who are likely to fail in or drop out from a course. This type of precautionary work plays an important role in assisting instructors in tailor-making targeted and evidence-based interventions for students in need of additional learning support.

Considerable scholarly attention has been devoted to using learning analytics to support teaching and learning. For example, van Leeuwen et al. (2015) investigated the impacts of learning analytics for collecting and processing information about students' cognitive activities to facilitate teachers' diagnoses and interventions. Cerezo et al. (2016) analysed students' interactions on Moodle and their learning achievements using data mining techniques. Kim et al. (2018) examined the use of learning analytics to support self-regulated learning in asynchronous online courses. Herodotou et al. (2020) explored the extent to which learning analytics could inform the design of motivational interventions and their impacts on student retention. Cloude et al. (2021) studied the use of game-based learning analytics to support student learning. Zamecnik et al. (2022) focused on the impacts of learning analytics dashboards on students' course performance. These studies have highlighted the importance of learning analytics for teachers and students in terms of providing pedagogical insights into the areas and ways to optimise teaching and learning.

Systematic reviews have also been conducted to analyse what has been done and found in the field of learning analytics (e.g., Alban & Mauricio 2019; Alwarthan et al. 2022; López-Zambrano et al. 2021; Srivastava & Ahmad 2022). While these reviews have informed the ways to further advance related studies and optimise teaching and learning, little attention has been paid to systematically surveying and summarising relevant latest work on the prediction of at-risk students through learning analytics. The present study therefore aims to review relevant work and, in particular, address the following research questions:

- What types of prediction have been made in relation to identifying students who are potentially at risk?
- What data and techniques have been used to predict students who are potentially at risk?
- What performance metrics have been adopted to evaluate the predictions of at-risk students?

2 Related work

Previous review studies on learning analytics have addressed a broad range of areas. For example, Wong (2017) and Wong et al. (2018) examined the overall developments and trends of learning analytics practices in higher education, and Li et al. (2018) focused particularly on the implementation of learning analytics in Asia. On specific areas of development, Li and Wong (2020) surveyed the trends of learning analytics in STEM and STEAM education. Wong et al. (2022) analysed how learning analytics has facilitated personalised learning. Li et al. (2023) presented the use of predictive analytics to support student admission.

Reviews related to the use of analytics techniques to support at-risk students have also covered various areas. They have addressed, for example, student success, drop-out and retention, (Li et al. 2022; Shafia et al. 2022; Tamada & de Magalhães Netto 2019),

and student performance (Alamri & Alharbi 2021; Albreiki et al. 2021; Alwarthan et al. 2022; López-Zambrano et al. 2021; Srivastava & Ahmad 2022).

There is a range of work which has examined the prediction of at-risk students. For example, an early study by Stuenkel (2006) investigated the use of standardised examinations, course grades, and entrance criteria to identify at-risk students, and found that these were effective predictors for achieving the purpose. The findings serve to inform when to provide early interventions for students potentially at risk, such as the period of admission to a programme when entrance criteria could be used for advising students, and the period of entry to a study major when students show a low GPA or low grades in their prior courses. Hawken et al. (2014) examined the use of check-in and check-out interventions for students at risk of emotional and behavioural disorders. The authors observed that this type of intervention was, in general, helpful in reducing students' problem behaviours and increasing appropriate social and academic behaviours. Kremer et al. (2015) studied the effects of after-school programmes on externalising behaviours and school attendance with at-risk students. They found that simply implementing these programmes may not be enough to have positive impacts on desirable outcomes without building in specific mechanisms to impact those outcomes.

The use of learning analytics has become increasing popular in recent work for predicting students at risk. For example, Choi et al. (2018) utilised student response systems to collect data related to students' learning status to identify those who may have encountered difficulties in learning. They also comprehensively analysed various intervention approaches and proposed a mechanism of systematic proactive intervention to deal with various types of at-risk student at different times during a course. Fahd et al. (2022) surveyed the use of various machine learning models, evaluation metrics, and criteria for assessing students' academic performance, risk, and attrition in higher education, and identified the most frequent adoption of classification models, regression, and clustering. Their results have provided insights into identifying the characteristics of analytics techniques for at-risk student prediction and future research directions.

Despite the various studies on the application of learning analytics for at-risk student prediction, little attention has been paid to reviewing them and summarising their patterns. Among the limited related studies, Na and Tasir (2017) examined how at-risk students could be identified by analysing students' online learning behaviours. In terms of research purpose, they found that a majority of relevant publications focus on identifying the attributes or indicators of at-risk students. As regards the data employed for identifying at-risk students, learning level and network data were found to be more frequently used than other types, such as learning emotion. Regarding the types of analytic technique used, logistic regression and decision trees were found to be most common. As for the strategies to help at-risk students, interventions related to course designs, instructional materials, pedagogical recommendations, and teaching methods are mostly recommended. However, this study was restricted to the context of online learning and only to publications up until 2017.

To address the literature gap, the present study aims to extend the line of inquiry by surveying and analysing what has been done and found in relation to the recent work on the prediction of at-risk students through the use of learning analytics. The results

provide us with the most up-to-date account of the field, including its latest research developments and future research directions.

3 Methodology

3.1 Data Collection

For the present study, relevant articles were collected from Scopus—a publication database commonly used for literature reviews. The set of keywords [“learning analytics” and (“predict” or “prediction”)] was used for searching relevant articles, and the period of publication was set from 2013–2022. An initial search yielded a total of 232 results. Each was checked and screened based on the following selection criteria: (i) the article reports on an empirical study on the use of learning analytics for at-risk student prediction, (ii) it was written in English, and (iii) it was available in full text. Articles which failed to meet any of these criteria were excluded from the present study. Finally, a total of 138 articles were selected for review and analysis.

3.2 Data Analysis

The selected articles were reviewed using a content analysis approach. Each article was first skimmed to gain an overall understanding of the work, and relevant information in the article related to the research questions of this study was identified, including (i) the types of prediction made for identifying students who are potentially at risk, (ii) the data and techniques used to predict at-risk students, and (iii) the performance metrics adopted to evaluate the predictions. The identified information was coded and categorised first by a researcher and then checked by one of the authors of this paper. Divergence in the judgement of coding was discussed until an agreement was reached. Based on the coding results, the patterns related to the use of learning analytics for at-risk student prediction were analysed.

4 Results

4.1 Number of Publications

Figure 1 shows the number of the publications between 2013 and 2022. There has been an overall increasing trend in the amount of relevant work throughout the period. The number remained low in the early years, with no more than three publications in each year between 2013 and 2015. Then there was an increase in the number, particularly since 2019. The number of publications reached 35 in 2020, which was the largest throughout the entire period. These results reveal the rapid development in the use of learning analytics techniques for prediction of at-risk students in recent years.

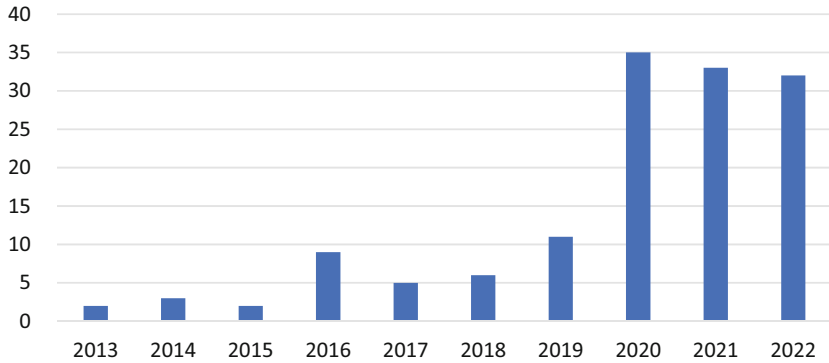


Fig. 1. Number of publications between 2013 and 2022

4.2 Areas of Prediction

Figure 2 presents the areas of prediction made for identification of at-risk students in the studies reviewed. Four areas of prediction were identified. Students' academic performance, such as quiz and examination scores, course grades, and GPA, was most commonly predicted as an indicator of their risk level, covering 72.1% of the studies. This is followed by students dropping out of their courses/programmes, which accounts for 26.4% of the studies. Only a very small proportion of studies focused on prediction of student engagement and assignment submission, i.e., 0.7% for each of these two areas. The results show that at-risk students were most frequently defined and identified through their academic performance and dropout rate.

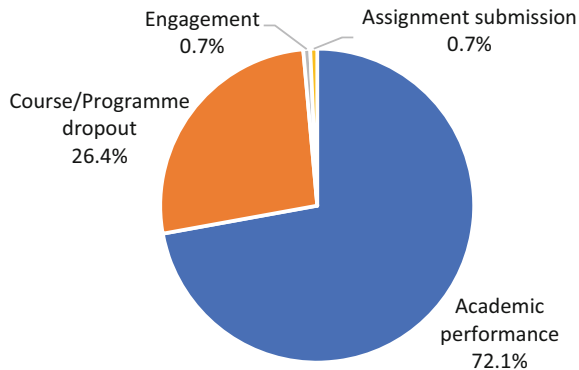


Fig. 2. Areas of prediction made in the studies reviewed

4.3 Data Used for Prediction

Figure 3 shows the data used to predict at-risk students in the studies. Five types of data were identified. Data related to students' previous academic performance (e.g.,

course grades and test scores) was most frequently used, i.e., 62.3% of the studies. This is followed by socio-demographic data (50.7%) such as gender, age, and family income, data about activities on learning management systems (44.9%) such as time and frequency of participating in learning activities, and academic records (35.5%) such as admission test scores and number of course activities taken. Data about learning habits and behaviours (e.g., participation in after-school tutoring courses) only accounts for 7.2% of the studies.

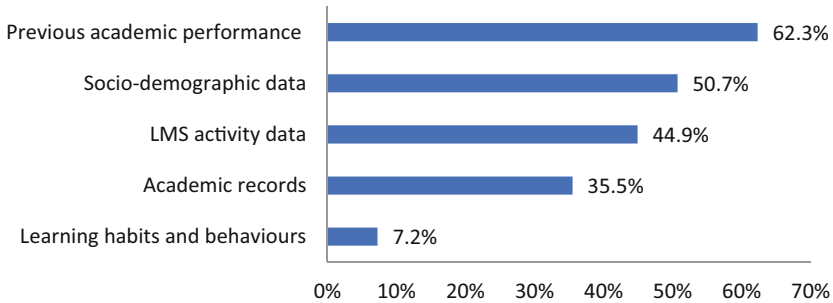


Fig. 3. Data used to predict at-risk students

4.4 Techniques Used for Prediction

Figure 4 illustrates the analytics techniques used for prediction. A total of 13 techniques were identified. More than half of the studies (52.9%) adopted decision trees as the prediction technique. This is followed by Bayesian networks and neural networks, with each accounting for 34.8% of the studies, and then support vector machines (30.4%), logistic regression (26.8%), and random forests (26.8%). It is worth noting that 84 of the 138 studies reviewed (i.e., 60.9%) employed more than one technique for prediction.

4.5 Performance Metrics Used to Evaluate Predictions

Figure 5 presents the top 10 performance metrics used to evaluate the predictions of at-risk students. Only about 88% of the publications report the performance metrics used. Accuracy is the most frequently used performance metric, which was adopted in 63% of the studies. This is followed by recall/sensitivity/true positive rate (44.2%), precision (36.2%), and F-measure/F-score/F1-measure/F1-score (29.7%). Similar to the results of the analytics techniques above, a majority of the studies (63.8%) adopted more than one performance metric for evaluation.

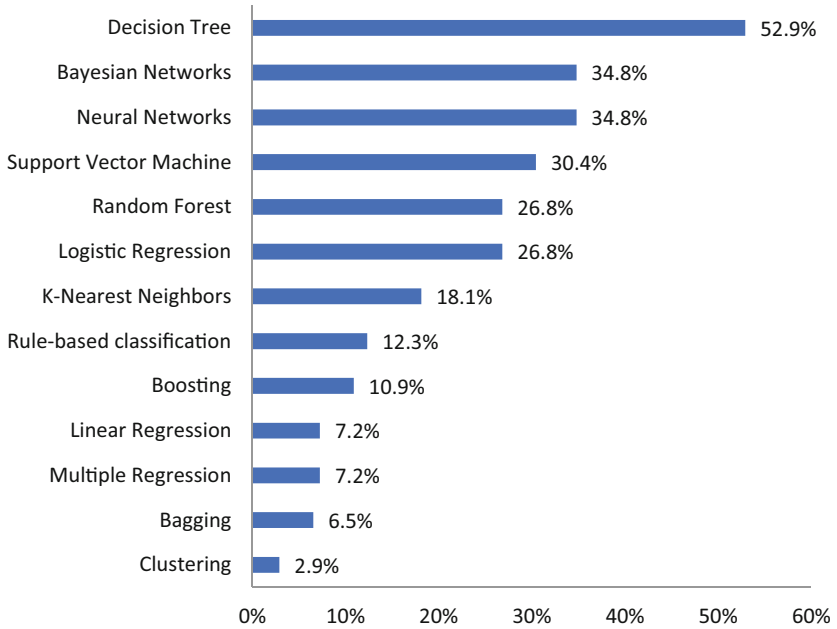


Fig. 4. Techniques used to predict at-risk students

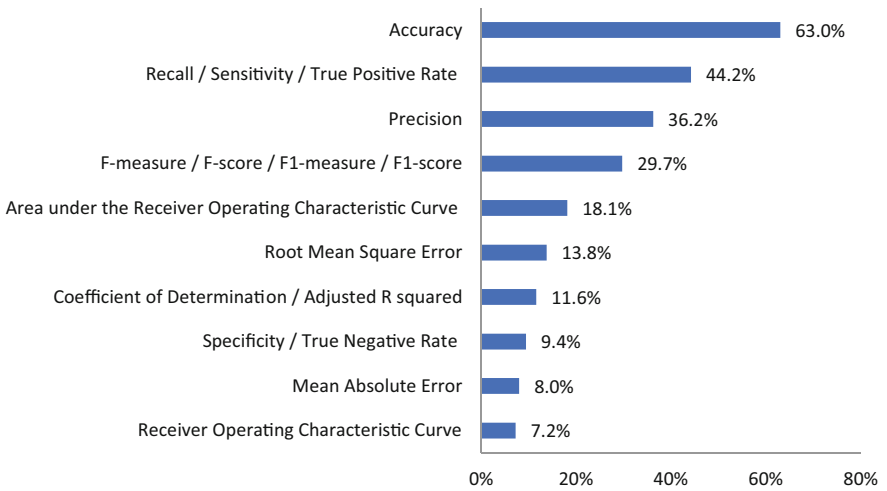


Fig. 5. Top 10 performance metrics used to evaluate predictions

5 Discussion and Conclusion

The findings of this study present an overview of the use of learning analytics for prediction of at-risk students. Regarding the areas of prediction, most of the studies focused on identifying students who are likely to fail a course or have poor academic performance. Academic failure is also closely related to student dropout, and its prediction helps determine the various learning difficulties that students face (Bedregal-Alpaca et al. 2020). Early detection of students' possible academic failure is therefore important for addressing their difficulties in a timely manner and reducing the likelihood of their dropping out. In terms of the data used for prediction, the common use of students' previous academic performance, socio-demographic data, academic records, and LMS data could be attributed to their easy access from institutional databases, as well as the strong association of students' academic performance (Alturki & Alturki 2021).

The broad range of analytics techniques used in the prediction studies and the common pattern that most studies employed multiple techniques reveal that no single technique could consistently show the best prediction performance among the available techniques. As indicated by Priyasadie and Isa (2021), each technique performs differently depending on the attributes and data set used. This suggests a need to experiment and compare various techniques to determine the most suitable ones in every implementation case.

A majority of the studies used accuracy as a performance metric for evaluation of prediction results. Nevertheless, it should be noted that the accuracy metric may not be sufficient on its own to reliably show the performance of a prediction technique. In cases where the data are not evenly distributed, the evaluation results from this metric can be misleading (Barros et al. 2019). This also suggests the need to adopt multiple performance metrics to provide comprehensive and reliable evaluation results of prediction performance.

These findings contribute to informing the potential future work to be conducted. Given the rapid development of artificial intelligence in recent years, future research could examine the use of the latest relevant artificial intelligence techniques for prediction of at-risk students. Moreover, the factor of academic disciplines could be studied to find out whether there are differences between the patterns of at-risk students studying in various academic disciplines. Further analysis should also be done to uncover the relationships among the various areas of prediction, the data and techniques used, and the performance metrics, all of which would be helpful for identifying the suitable combination of them in different contexts for prediction.

Acknowledgments. The work described in this paper was partially supported by grants from Hong Kong Metropolitan University (2021/011 and CP/2022/04).

References

- Ablan, M., Mauricio, D.: Predicting university dropout through data mining: a systematic literature. *Indian J. Sci. Technol.* **12**(4), 1–12 (2019)

- Alamri, R., Alharbi, B.: Explainable student performance prediction models: a systematic review. *IEEE Access* **9**, 33132–33143 (2021)
- Albreiki, B., Zaki, N., Alashwal, H.: A systematic literature review of students' performance prediction using machine learning techniques. *Educ. Sci.* **11**, 552 (2021)
- Alturki, S., Alturki, N.: Using educational data mining to predict students' academic performance for applying early interventions. *J. Inf. Technol. Educ. Innov. Pract.* **20**, 121–137 (2021)
- Alwarthan, S.A., Aslam, N., Khan, I.U.: Predicting student academic performance at higher education using data mining: a systematic review. *Appl. Comput. Intell. Soft Comput.* 8924928 (2022)
- Barros, T.M., Neto, P.A.S., Silva, I., Guedes, L.A.: Predictive models for imbalanced data: a school dropout perspective. *Educ. Sci.* **9**(4), 275 (2019)
- Bedregal-Alpaca, N., Cornejo-Aparicio, V., Zarate-Valderrama, J., Yanque-Churo, P.: Classification models for determining types of academic risk and predicting dropout in university students. *Int. J. Adv. Comput. Sci. Appl.* **11**(1), 266–272 (2020)
- Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, Núñez, J.C.: Students' LMS interaction patterns and their relationship with achievement: a case study in higher education. *Comput. Educ.* **96**, 42–54 (2016)
- Choi, S.P.M., Lam, S.S., Li, K.C., Wong, B.T.M.: Learning analytics at low-cost: at-risk student prediction with clicker data and systematic proactive interventions. *Educ. Technol. Soc.* **21**(2), 273–290 (2018)
- Cloude, E.B., Carpenter, D., Dever, D.A., Azevedo, R., Lester, J.: Game-based learning analytics for supporting adolescents' reflection. *J. Learn. Anal.* **8**(2), 51–72 (2021)
- Fahd, K., Venkatraman, S., Míah, S.J., Ahmed, K.: Application of machine learning in higher education to assess student academic performance, at-risk, and attrition: a meta-analysis of literature. *Educ. Inf. Technol.* **27**, 3743–3775 (2022)
- Hawken, L.S., Bundock, K., Kladis, K., O'Keeffe, B., Barrett, C.A.: Systematic review of the check-in, check-out intervention for students at risk for emotional and behavioral disorders. *Educ. Treat. Child.* **37**(4), 635–658 (2014)
- Herodotou, C., Naydenova, G., Boroowa, A., Gilmour, A., Rienties, B.: How can predictive learning analytics and motivational interventions increase student retention and enhance administrative support in distance education? *J. Learn. Anal.* **7**(2), 72–83 (2020)
- Kim, D., Yoon, M., Jo, I-H., Branch, R.M.: Learning analytics to support self-regulated learning in asynchronous online courses: a case study at a women's university in South Korea. *Comput. Educ.* **127**, 233–251 (2018)
- Kremer, K.P., Maynard, B.R., Polanin, J.R., Vaughn, M.G., Sarteschi, C.M.: Effects of after-school programs with at-risk youth on attendance and externalizing behaviors: a systematic review and meta-analysis. *J. Youth Adolesc.* **44**, 616–636 (2015)
- Li, C., Herbert, N., Yeom, S., Montgomery, J.: Retention factors in STEM education identified using learning analytics: a systematic review. *Educ. Sci.* **12**, 781 (2022)
- Li, K.C., Wong, B.T.M.: Trends of learning analytics in STE(A)M education: a review of case studies. *Interact. Technol. Smart Educ.* **17**(3), 323–335 (2020)
- Li, K.C., Wong, B.T.M., Chan, H.T.: Predictive analytics for university student admission: a literature review. In: Li, C., Cheung, S.K.S., Wang, F.L., Lu, A., Kwok, L.F. (eds.) *Blended Learning: Lessons Learned and Ways Forward*, pp. 250–259. Springer (2023). https://doi.org/10.1007/978-3-031-35731-2_22
- Li, K.C., Wong, B.T.M., Ye, C.J.: Implementing learning analytics in higher education: the CASE of Asia. *Int. J. Serv. Stand.* **12**(3/4), 293–308 (2018)
- Long, P., Siemens, G.: Penetrating the fog: analytics in learning and education. *Educause Rev.* **46**(5), 30–40 (2011)
- López-Zambrano, J., Torralbo, J.A.L., Romero, C.: Early prediction of student learning performance through data mining: a systematic review. *Psicothema* **33**(3), 456–465 (2021)

- Na, K.S., Tsair, Z.: Identifying at-risk students in online learning by analysing learning behaviour: a systematic review. In: The 2017 International Conference on Big Data and Analytics, pp. 118–123. Kuching, Malaysia (2017)
- Priyasadie, N., Isa, S.M.: Educational data mining in predicting student final grades on standardized Indonesia Data Pokok Pendidikan data set. *Int. J. Adv. Comput. Sci. Appl.* **12**(12), 212–216 (2021)
- Shafia, D.A., Marjani, M., Habeeb, R.A., Asirvatham, D.: Student retention using educational data mining and predictive analytics: a systematic literature review. *IEEE Access*, **10**, 72480–72503 (2022)
- Srivastava, N., Ahmad, J.: A review on the learners' performance prediction techniques in MOOC courses through data mining. In: The 5th International Conference on Advances in Science and Technology, pp. 314–317. Mumbai, India (2022)
- Stuenkel, D.L.: At-risk students: Do theory grades + standardized examinations = success? *Nurse Educ.* **31**(5), 207–212 (2006)
- Tamada, M.M., de Magalhães Netto, J.F.: Predicting and reducing dropout in virtual learning using machine learning techniques: a systematic review. In: The 2019 IEEE Frontiers in Education Conference, pp. 1–9. Covington, KY, USA (2019)
- van Leeuwen, A., Janssen, J., Erkens, G., Brekelmans, M.: Teacher regulation of cognitive activities during student collaboration: effects of learning analytics. *Comput. Educ.* **90**, 80–94 (2015)
- Wong, B.T.M.: Learning analytics in higher education: an analysis of case studies. *Asian Assoc. Open Univ. J.* **12**(1), 21–40 (2017)
- Wong, B.T.M., Li, K.C., Cheung, S.K.S.: An analysis of learning analytics in personalised learning. *J. Comput. High. Educ.* (2022). <https://doi.org/10.1007/s12528-022-09324-3>
- Wong, B.T.M., Li, K.C., Choi, S.P.M.: Trends in learning analytics practices: a review of higher education institutions. *Interact. Technol. Smart Educ.* **15**(2), 132–154 (2018)
- Zamecnik, A., et al.: Team interactions with learning analytics dashboards. *Comput. Educ.* **185**, 104514 (2022)



Towards the Design of a Personalized Vocabulary Learning System for University Graduates Based on Context Readability and Task Diversity

Gary Cheng¹(✉) and Di Zou²

¹ Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR, China

chengks@eduhk.hk

² Centre for English and Additional Languages, Lingnan University, Hong Kong SAR, China

dizoudaisy@gmail.com

Abstract. This study aimed to enhance vocabulary acquisition among university graduates using a personalized system that incorporates two strategies based on the involvement load hypothesis: context readability and task diversity. Context readability involves understanding learners' interests and suggesting learning tasks within their preferred contexts, while task diversity focuses on tracking learners' progress and recommending various learning tasks to ensure comprehensive training for each target word. To evaluate the effectiveness of the proposed system, 30 university graduates were divided into three groups: a control group, a group with only task diversity, and a group with both context readability and task diversity. The learning performance of participants was assessed through pre- and post-vocabulary tests. Results show that integrating both strategies is most effective in enhancing vocabulary acquisition for university graduates. Future research could refine the system and explore additional strategies for diverse learner populations.

Keywords: vocabulary acquisition · personalized learning · context readability · task diversity

1 Introduction

The advent of online resources like mobile apps and MOOCs has revolutionized language learning, offering a lifeline to university graduates who lack access to traditional educational platforms (Fang et al. 2022). However, this wealth of resources can be a double-edged sword, leading to information overload that hampers effective vocabulary acquisition (Bawden & Robinson 2009). The complexity of mastering various aspects of vocabulary adds another layer of difficulty (Kalyuga & Kalyuga 2008). For university graduates, who often juggle limited study time with work pressures, these challenges are intensified (UBS 2015). While self-paced learning offers flexibility, it can also result in a narrow focus on a particular aspect of word usage (Zou et al. 2014) and a decrease in learning motivation (Shen et al. 2009).

Against this backdrop, this study aims to improve vocabulary acquisition for university graduates by introducing a personalized learning system that employs two strategies: context readability and task diversity, based on the involvement load hypothesis (Laufer & Hulstijn 2001). Context readability tailors learning tasks to individual interests, while task diversity offers varied exercises for comprehensive word training. The study also evaluates the system's effectiveness on the learning performance of participants. The results could inform the design of personalized language learning systems, especially for university graduates, who are frequently overlooked in existing research (Xie et al. 2019).

2 Related Work

2.1 Vocabulary Acquisition

Vocabulary acquisition is a vital aspect of second language development, impacting both reading and writing skills (Schmitt 2008). It involves not just learning new words but also understanding their usage in different contexts. Researchers in vocabulary acquisition focus on two main areas: measuring vocabulary knowledge and understanding the learning process. The former includes assessing the range of words a learner knows (breadth) and their proficiency in using those words (depth), often through vocabulary-level tests and word association tests (Bahar & Hansell 2000; Read 2007). The latter explores how vocabulary is learned, either intentionally through focused study or incidentally through natural exposure (Schmitt 2008). Some studies suggest that a combination of both intentional and incidental learning is most effective (e.g. Sok & Han 2020). Additionally, the advent of technology has transformed the landscape of vocabulary learning through the introduction of digital platforms such as games and mobile apps. These platforms not only make learning more engaging but also allow for personalized, adaptive learning experiences (Zhang & Perez-Paredes 2021).

2.2 Context Readability and Task Diversity

Prior studies have explored various methods for creating personalized vocabulary learning systems that recommend specific materials and activities to individual learners (e.g. Huang et al. 2012; Tsai et al. 2016). However, these systems often lack a solid grounding in linguistic theories for making these recommendations. One such theory, namely "involvement load hypothesis", suggests that the effectiveness of vocabulary learning is influenced by the level of learner involvement in the task (Laufer & Hulstijn 2001). This involvement is determined by three key factors: need, search, and evaluation. Two strategies, "context readability" and "task diversity", align well with this hypothesis. Context readability focuses on recommending tasks that are relevant to the learner's interests, thereby increasing their motivation or 'need' to learn new words. Task diversity, on the other hand, aims to improve vocabulary proficiency by offering a variety of tasks that challenge different cognitive aspects like 'search' and 'evaluation'. Despite the promise of these strategies, there is a research gap concerning their combined effectiveness, particularly for adult learners like university graduates. The current study aims

to fill this gap by investigating the effect of integrating both context readability and task diversity into personalized vocabulary learning systems, specifically focusing on learner performance in vocabulary acquisition.

3 Method

3.1 Study Context

This research is a part of a larger university-funded initiative named “A personalized task recommendation system for vocabulary learning based on context readability and task diversity for university graduates”. The research took place from September 2022 to June 2023 at a Hong Kong university with a specialization in teacher education. The main objective was to assess the effectiveness of a personalized vocabulary learning system that integrates both context readability and task diversity. University graduates, who participated in this study, were asked to engage in a two-week, self-directed online vocabulary learning using a personalized vocabulary learning system.

3.2 Participants

Participants in this study included 30 university graduates in Hong Kong who were interested in improving their vocabulary knowledge. Before the study began, all participants gave their written consent and filled out a questionnaire that collected basic demographic information, such as gender, age, and year of graduation. Among the participants, 18 were males and 12 were females, with ages ranging from 24 to 28 years ($M = 25.4$, $SD = 1.1$). They had graduated within the past 1 to 3 years, and they all represented the academic discipline of Information and Communication Technology (ICT).

3.3 Design of the Personalized Vocabulary Learning System

In the personalized vocabulary learning system, learners are recommended to undertake specific vocabulary learning tasks. As previously discussed, the system employs two strategies: (i) context readability, in which task recommendations are based on the learner’s prior familiarity with certain domains—such as recommending vocabulary tasks related to computer science for learners who are familiar with computer science articles; and (ii) task diversity, where learners who have mainly engaged in one type of exercise are subsequently advised to undertake alternative types of exercise. For example, like cloze exercises, are advised to explore alternative task types, such as sentence writing exercises.

The personalized vocabulary learning system used in this study focuses on teaching 20 target words, as listed in Table 1. These words are not among the 9,000 most frequently used words in Davies’ (2012) Corpus of Contemporary American English (COCA). Given that the average person’s English vocabulary size ranges from 6,700 to 9,000 frequent words, it is expected that our participants lack familiarity with these target words.

Regarding context readability, tasks were created in two distinct contexts: one in a general context and the other specifically focused on ICT. In terms of task diversity, four

types of tasks were developed. These include cloze exercises, where learners are asked to identify the correct word choice to fill in the blanks; sentence writing, where learners are asked to use the target word to create an original sentence; and reading comprehension, where learners are given an excerpt from an article and must write down the semantic meaning of the target words within that excerpt. Some representative tasks from these categories within general and specific contexts are given in Tables 2 and 3, respectively. Learners would receive the system's feedback on each completed task. For the tasks of sentence writing and reading comprehension and inferencing, ChatGPT generates feedback on the quality of the learner's writing. Figure 1 illustrates the overall design of the system and Fig. 2 shows the screenshot of the user interface.

Table 1. List of the target words.

Word	Part of Speech	Rank order in COCA
Abscond	Verb	30,537
Iconoclast	Noun	32,195
Ebullient	Adjective	27,715
Obfuscate	Verb	23,438
Dilettante	Noun	32,445
Hapless	Adjective	15,863
Inculcate	Verb	21,481
Ennui	Noun	27,759
Recumbent	Adjective	35,855
Bifurcate	Verb	35,710
Incongruity	Noun	24,200
Quizzical	Adjective	27,307
Impugn	Verb	25,514
Juggernaut	Noun	18,415
Pernicious	Adjective	17,104
Ingratiate	Verb	30,247
Intrepidity	Noun	61,217
Erudite	Adjective	23,446
Admonish	Verb	14,368
Munificence	Noun	51,902

3.4 Measure

To measure their vocabulary proficiency of participants, they were given both a pre-test and a post-test before and after using the personalized vocabulary learning system. In

Table 2. List of sample tasks with general context.

Task type	Task description
Cloze exercise	Upon realizing that the police were on their way, the thief decided to _____ with the stolen jewels
Sentence writing	Write an original sentence using the target word ‘iconoclast’
Reading comprehension	‘During the town’s annual festival, Lucy’s ebullient spirit was impossible to miss. She danced through the streets, laughing and singing, bringing smiles to the faces of everyone she passed. Her radiant energy was infectious, making even the most reserved individuals join in the festivities with renewed vigor.’ What is the meaning of the target word ‘ebullient’ in this passage?

Table 3. List of sample tasks with specific context.

Task type	Task description
Cloze exercise	After discovering a security vulnerability in the company’s software, the programmer decided to _____ with sensitive user data without notifying his superiors
Sentence writing	Write an original sentence using the target word ‘iconoclast’ within the context of information technology
Reading comprehension	‘In the software team, many engineers were busy with their work. But when something good happened or a problem was fixed, Kevin’s ebullient spirit was obvious. He was really happy about any progress, which made everyone feel better and work harder.’ What is the meaning of the target word ‘ebullient’ in this passage?

each test, participants were required to write down the meaning and create an original sentence for each target word. The scoring system, based on Zou’s (2017) methodology, was applied to each of the 20 target words covered in the system. A score of 0 was given if the meaning was completely incorrect, 1 if it was a semantically acceptable equivalent of the target word, and 2 if it was comparable to the meaning of the target word. Similarly, for sentence writing, a score of 0 was assigned if the sentence had a completely inappropriate semantic context for the target word, 1 if the sentence had an appropriate semantic context but the use of the target word was grammatically incorrect, and 2 if the sentence had an appropriate semantic context and the use of the target word was grammatically correct. Therefore, the total possible score ranged from 0 to 80.

3.5 Procedure

The procedure involved three major steps. The first step was to examine participants’ understanding of the target words using a pre-test before the start of the study. This was

done to ensure their homogeneity in vocabulary proficiency. In the second step, participants were randomly assigned into one of three treatment groups, each containing 10 members. Group A did not receive any recommendations on learning tasks, and they selected and attempted vocabulary learning tasks at their own pace. Group B received recommendations on learning tasks from the system, with only the feature of task diversity included. Group C received recommendations on learning tasks from the system with both features of context readability and task diversity. The duration of this second step was approximately two weeks. The final step was to administer a post-test to these three groups. Differences in the post-test results among the groups were evaluated to examine the effectiveness of the proposed system. Figure 3 provides a visual representation of the research procedure.

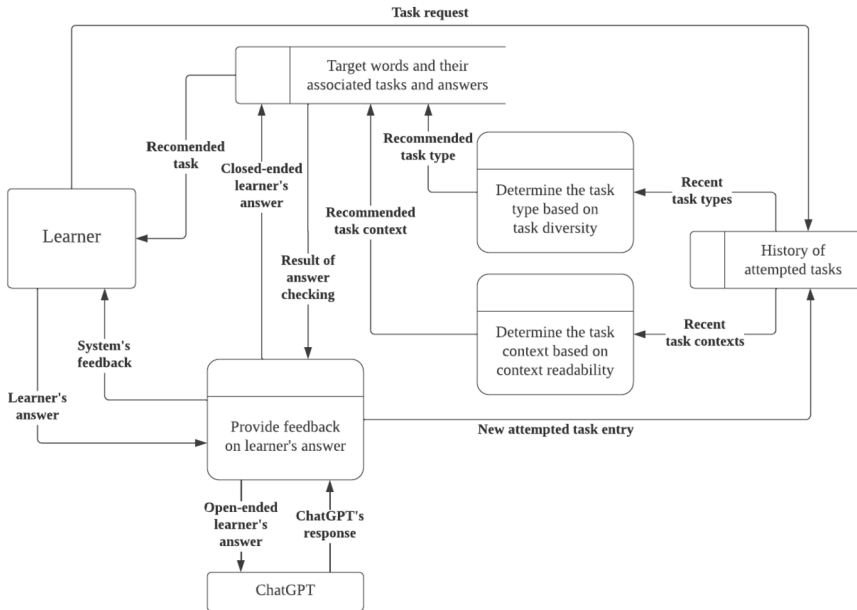


Fig. 1. Overall design of the personalized vocabulary learning system

Personalized Task Recommendation System

Use ICT context

Task diversity

Select a task type

Clozed exercise Sentence writing Reading comprehension

System message

The IT manager admonished the team for not regularly updating their passwords, emphasizing the importance of maintaining strong security practices. [Write down the meaning of the target word: admonish]

User input

Type your answer here

Submit Answer New Task

Fig. 2. Screenshot of the user interface

4 Results and Discussion

Figure 4 presents a box plot that delineates the distribution of the post-test scores across the three groups. This graphical representation serves to provide an overview of the data's central tendency and variability. In terms of statistical assumptions, the post-test scores for each group were subjected to the Shapiro-Wilk test to assess normality. The results confirmed that the data for Group A ($W = 0.895$, $p = 0.194$), Group B ($W = 0.920$, $p = 0.359$), and Group C ($W = 0.874$, $p = 0.113$) were normally distributed. Additionally, the Levene's test was employed to verify the homogeneity of variances across the groups, yielding an F-value of 0.387 and a p-value of 0.683. The fulfillment of both data normality and variance homogeneity conditions validates the use of one-way Analysis of Variance (ANOVA) and subsequent post hoc analyses for comparing performance scores among the three groups.

As shown in Table 4, the results of the one-way ANOVA revealed a significant difference between the groups [$F(2, 27) = 24.90$, $p < 0.001$]. Post hoc comparisons using Tukey's HSD test indicated that Group B significantly outperformed Group A ($p = 0.01$) and that Group C significantly outperformed Group A ($p < 0.001$) and Group B ($p = 0.002$). The effect size for the test, as measured by η^2 , was 0.648. This large effect size suggests that approximately 64.8% of the variability in performance scores can be attributed to the different group memberships. Such findings indicate that the intervention applied to Group C was particularly effective in enhancing students' performance scores.

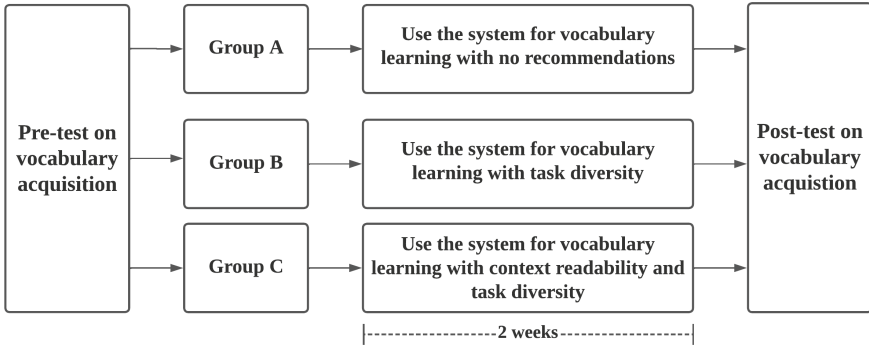


Fig. 3. Research procedure

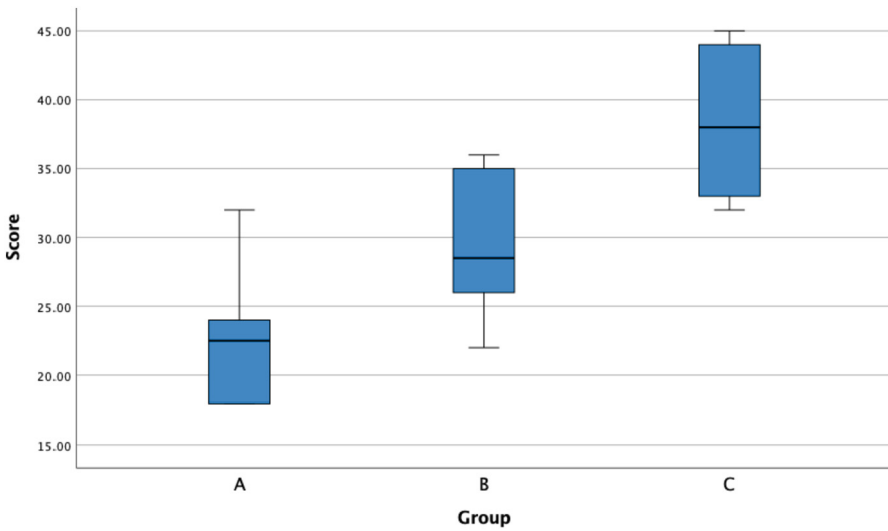


Fig. 4. Distribution of the post-test scores across the three groups

Table 4. Results of the one-way ANOVA.

	Sum of Squares	df	Mean Square	F	Sig
Between Groups	1174.200	2	587.100	24.901	<0.001
Within Groups	636.600	27	23.578		
Total	1810.800	29			

5 Concluding Remarks

The present study aimed to investigate the effectiveness of a personalized vocabulary learning system that recommends vocabulary learning tasks to university graduates. The study was designed with three groups: Group A, which received no recommendations; Group B, which received recommendations based on task diversity; and Group C, which received recommendations based on both context readability and task diversity. The findings suggest that the system's recommendations, especially those that incorporate both context readability and task diversity, could make a substantial positive impact on vocabulary learning outcomes.

While the results are promising, there are several avenues for future research. First, the study duration was relatively short, lasting only two weeks. Longer-term studies could provide more insights into the sustained effects of the system on vocabulary acquisition. Second, the sample size was limited to 30 participants from the same discipline. Future studies could benefit from a more diverse sample to generalize the findings. Additionally, the study only examined two features—context readability and task diversity—for task recommendations. Future research could explore other features, such as learners' cognitive styles, to assess their impact on learning outcomes. Subsequent research could also employ qualitative methods, such as interviews or open-ended questionnaires, to gain a more comprehensive understanding of how learners interact with the system and how it impacts their motivation and engagement.

Acknowledgements. The research has been supported by the Interdisciplinary Research Scheme of the Dean's Research Fund 2021/22 (FLASS/DRF/IDS-3 2022) of The Education University of Hong Kong.

References

- Bahar, M., Hansell, M.H.: The relationship between some psychological factors and their effect on the performance of grid questions and word association tests. *Educ. Psychol.* **20**(3), 349–364 (2000)
- Bawden, D., Robinson, L.: The dark side of information: overload, anxiety and other paradoxes and pathologies. *J. Inf. Sci.* **35**(2), 180–191 (2009)
- Davies, M.: *Corpus of American contemporary English* (2012). <http://www.americancorpus.org/>
- Fang, J.-W., Hwang, G.-J., Chang, C.-Y.: Advancement and the foci of investigation of MOOCs and open online courses for language learning: a review of journal publications from 2009 to 2018. *Interact. Learn. Environ.* **30**(7), 1351–1369 (2022)
- Huang, Y.M., Huang, Y.M., Huang, S. H., Lin, Y.T.: A ubiquitous English vocabulary learning system: evidence of active/passive attitudes vs. usefulness/ease-of-use. *Comput. Educ.* **58**(1), 273–282 (2012)
- Kalyuga, M., Kalyuga, S.: Metaphor awareness in teaching vocabulary. *Lang. Learn. J.* **36**(2), 249–257 (2008)
- Laufer, B., Hulstijn, J.: Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Appl. Linguis.* **22**(1), 1–26 (2001)
- Read, J.: Teaching and learning vocabulary: bringing research into practice. *Stud. Second. Lang. Acquis.* **29**(1), 128–129 (2007)

- Schmitt, N.: Reviewed article: Instructed second language vocabulary learning. *Lang. Teach. Res.* **12**(3), 329–363 (2008)
- Shen, L., Wang, M., Shen, R.: Affective e-learning: using emotional data to improve learning in pervasive learning environment. *Educ. Technol. Soc.* **12**(2), 176–189 (2009)
- Sok, S., Han, Z.: A study of L2 vocabulary acquisition under incidental and intentional conditions. *Vigo Int. J. Appl. Linguis.* **17**, 113–140 (2020)
- Tsai, P.S., Tsai, C.C., Hwang, G.H.: The effects of instructional methods on students' learning outcomes requiring different cognitive abilities: context-aware ubiquitous learning versus traditional instruction. *Interact. Learn. Environ.* **24**(7), 1497–1510 (2016)
- UBS. Prices and earnings 2015: Do I earn enough for the life I want? Switzerland: Chief Investment Office WM (2015). <https://www.ubs.com/microsites/prices-earnings/>
- Xie, H., Chu, H.-C., Hwang, G.-J., Wang, C.-C.: Trends and development in technology-enhanced adaptive/personalized learning: a systematic review of journal publications from 2007 to 2017. *Comput. Educ.* **140**, 103599 (2019)
- Zhang, D., Perez-Paredes, P.: Chinese postgraduate EFL learners' self-directed use of mobile English learning resources. *Comput. Assist. Lang. Learn.* **34**(8), 1128–1153 (2021)
- Zou, D.: Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: extending the evaluation component of the involvement load hypothesis. *Lang. Teach. Res.* **21**(1), 54–75 (2017)
- Zou, D., Xie, H., Li, Q., Wang, F.L., Chen, W.: The Load-based learner profile for incidental word learning task generation. *Adv. Web-Based Learn. LNCS* **8613**, 190–200 (2014)



Theoretical Framework of Personal Learning Environments: SPET Model

Xiaoshu Xu¹(✉), Yilin Sun², Jie Weng¹, and Yunfeng Zhang³

¹ School of Foreign Studies, Wenzhou University, Wenzhou, China

Lisaxu@wzu.edu.cn

² Seattle Colleges, Seattle, US

Yilin.Sun@seattlecolleges.edu

³ Faculty of Language and Translation, Macao Polytechnic University, Macao, China

zhangyunfeng@mpu.edu.mo

Abstract. The lack of a compelling educational theory underlying the advancement of Personal Learning Environments (PLEs) has an impact on PLE implementation in Higher Education (HE). The purpose of this article is to introduce the concepts and theoretical framework of PLEs in the context of higher education. Constructivism, Connectivism, and Cognitivism have all been used in previous studies to explain the principles and implementations of PLEs. However, in order to promote the development of PLEs in higher education, a comprehensive theoretical framework is required. We developed the Society, Policy, Education, and Technology (SPET) framework to explain the environmental factors involved in accelerating the promotion of PLEs, beginning with redefining the value of Higher Education. Furthermore, we created a PLE framework that incorporates four major perspectives in the educational field to explain the internal factors of PLEs. These are academic administrative, learner and learning, teacher and teaching, and information communication technology (ICT) perspectives. All stakeholders involved in promoting PLEs in higher education can benefit from the two PLE frameworks.

Keywords: personal learning environments · PLEs · theoretical framework · higher education

1 Introduction

Two decades have passed since the first recorded use of the term “Personal Learning Environment” in a session titled at the 2004 Joint Information Systems Committee/Center for Education Technology & Interoperability Standards Conference in the UK, while the implementation of PLEs was still in its early stages, slow, and fraught with challenges (Ricardo Torres-Kompen 2015). As Attwell (2021) pointed out, recent developments in the PLEs field have a tendency to drift into tech-centrism; that is, too much time has been spent on the new technology rather than looking at the pedagogy of how PLEs might be used.

The reasons for the slow progress in the PLEs field were complex and varied, involving a lack of consensus on PLEs among all stakeholders; conservative policy;

the rigid Higher Education (HE) system; the inherent attitudes of teachers, learners, and administrators; an overemphasis on the social development function of HE, and, most importantly, a lack of PLEs theory and a theoretical framework.

The purpose of this research is to create a wholistic theoretical framework for PLEs that includes the external Society, Policy, Education, and Technology (SPECT) framework as well as the internal framework that includes four perspectives: academic administration, learner and learning, teacher and teaching, and information communication technology (ICT).

We hope that the proposed framework will give all key stakeholders an in-depth and comprehensive understanding of the methods and theories that can be used to promote the implementation of PLEs in higher education.

2 The Concept of Personal Learning Environments (PLEs)

First, consider PLEs as platforms. PLEs appeared as “a new construct in the e-learning literature that finds its support on social media and steadily gains ground in the e-learning field as an effective platform for student learning” (Dabbagh & Kitsantas 2012). PLEs have been described as “a type of e-learning system that is structured on a model of e-learning itself rather than a model of the institution” (Kalz et al. 2011), as well as multidimensional spaces (van Harmelen 2006) or central nodes of a network (Downes 2010, p. 30).

Second, PLEs as source aggregation: such as Fiedler & Pata (2009), Amine (2009), Dolores Reig (el caparazón) 2009), or Henri et al. (2008) who define PLEs as a self-defined collection of resources, services, tools, and devices that can assist teachers and students in shaping their personal learning and knowledge networks.

Third, PLEs as an innovative educational methodology: PLEs were considered a means of understanding both how students learn and how educators teach (Attwell, 2007; Adell & Castañeda 2010; Cabero 2012). PLEs were essentially a social-pedagogical approach to teaching and learning with technology. (Attwell 2021). PLE “focuses on the students and the decisions they make to personalize and self-regulate their learning.”

As can be seen from the above definitions of PLEs, the implementation of PLEs was not only about technology reform but also about a philosophy or a way of working, as asserted by Sebba et al. (2007). The essence of PLEs was a vision to empower learners to take more ownership of their learning as well as to foster learner autonomy and lifelong learning (Ballard & Butler 2011). More recently, the concept of PLEs has tended to drift into techno-centrism, usually including references to social media and the possibilities for the creation and sharing of knowledge that ultimately relate to the control of learning by the student (Rahimi et al. 2015).

Dabbagh & Castañeda (2020) claimed that PLEs could be characterized as a techno-social reality that embodies the socio-material entanglement with which people are able to learn, with an approach that enacts contemporary ideas about how people learn. The PLE’s learning ecology and ability to connect formal, non-formal, and informal learning experiences within a framework that addresses the challenges of continuous learning and empowers learners to develop lifelong learning skills.

In the current study, PLEs are defined as new digital learning literacy, conceptual space, pedagogical processes, and social networks that enable and support learners in

achieving their lifelong learning objectives. PLEs are regarded as dynamic, interconnected environments with an ever-changing community of learners, instructors, tools, and content.

We envision the PLE as the heart of a diverse, personalized, social, adaptive, integrated, and transparent learning activity ecosystem, enabling the creation of a network of learning that supports students as peers, creators, entrepreneurs, and agents of their own learning.

3 The Features of Learning in PLEs

The PLEs paradigm occurs in a social context in which knowledge is shared globally and students act as problem solvers. Faculty members provide motivational and evaluative coaching, as well as the facilitation of relationships in ad hoc, peer-to-peer networks. The assumption is that students can self-actualize by utilizing the proposed educational system to fully realize their human potential. Learning in PLEs is self-education in which the learner, rather than the teacher, determines the learning program's goals, learning procedures, resources, and evaluation decisions (Moore 1984).

Our understanding of PLEs is comprised of five parts. First, it entails self-directed learning processes in which learners must set goals, make appropriate plans, and implement the plans. To do so, learners require the space, environment, and guided support to integrate their needs, life experiences, and culture into goal-setting and action plans.

Second, it entails a collaborative process via networked learning or support groups. It should include guiding principles to ensure that all key stakeholders involved in the collaborative process understand the importance of equity and access in order for all learners to develop a strong sense of belonging.

Third, it is a “knowledge construction” process in which learners simultaneously build knowledge through theoretical learning and real-world practice. This process improves skills, task mastery, confidence, and a sense of belonging, as well as preparing the mind and body for intellectual rigour in mindset and skills, which leads to purposeful and spontaneous learning.

Fourth, by integrating the curriculum with the practical application of knowledge and teamwork, these processes necessitate and strengthen human trust and reciprocity (Humphris 2007).

Finally, there is the changing value and status of Information and Communication Technologies (ICTs), which are no longer merely instruments or tools but are shaping sociocultural practices and human behavior at all levels (Adell 2018; Selwyn 2017; Veletsianos and Moe 2017).

To summarize, PLEs are processes of collaborative, cooperative, and collective inquiry; knowledge construction and transfer; trusting relationships and supportive settings; a sense of belonging and capacity realization; and convivial technologies. PLEs help people, learning sites, and actions, as well as ideas, resources, and solutions, connect across time, space, and media.

4 Why does Higher Education need Personal Learning Environments?

The growing demand for lifelong learning opportunities to keep up with social, economic, and technological changes drives the demand for more accessible alternatives to traditional real-time, campus-based instruction. Technology advancements are reshaping education by changing when, how, and where students learn, as well as empowering them at every stage of their journey.

Technology has the potential to tailor education to individual preferences and requirements. As a learner, he or she can establish formal or informal learning objectives. There are three options for achieving these personalized learning goals.

First, through formal higher education, which benefits from standardization, systematization, institutionalization, and specialization; second, through difficult and uncertain self-directed learning; and third, through enterprise-supported learning tools and platforms, which are distinguished by personalization, cutting-edge technology, adaptability, and flexibility, as shown in Fig. 1.

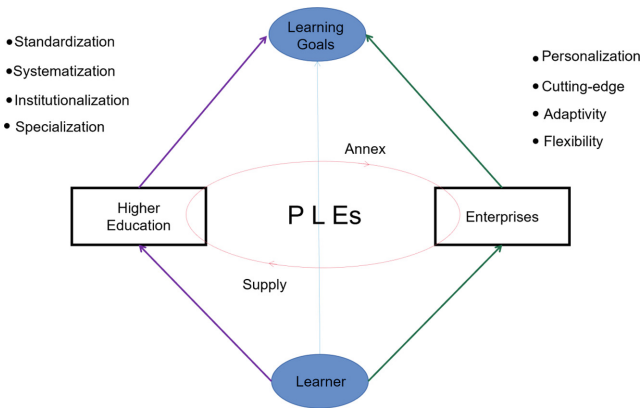


Fig. 1. Three ways to reach the educational ideal goal

With the advancement of Information and Communication Technology (ICT), higher education (HE) is no longer the sole source or path to achieving the educational ideal. Knowledge claims are no longer made solely by universities; knowledge production is increasingly being built in businesses, non-governmental organizations, and the so-called third sector, with local, inter, and supranational actors involved (Soysal & Baltaru 2021).

Learners can use third-party learning platforms like Zoom or Google Classroom, as well as apps like Duolingo, to make learning anywhere more efficient and enjoyable.

Meanwhile, enterprises create interactive online tools such as the Google Slides add-on “Pear Deck,” course content sharing “Edmodo,” and an advanced LMS tool “Blackboard,” among others, to assist learners in making the best decisions about which applications best suit their various online subjects and learning needs.

For example, the pandemic has created an opportunity to assess the suitability of third-party education delivery systems. All of these learning tools, platforms, Open Educational Resources (OER), and technologies, such as data mining and learning analytics, enable personalized learning. Learners can create their own learning paths based on their preferred learning style, prior knowledge, and skills and interests.

With the trend toward more personalized learning, there is an urgent need for PLEs to be integrated into HE. For one thing, formal higher education lacks flexibility and responsiveness, which causes its pedagogy, curriculum design, and teaching methods to lag far behind in the rapidly changing digital information age. These decreases learning efficiency and increases the gap between learners' needs, expectations, and achievements. Self-paced, personalized learning, on the other hand, can aid in the development of learners' creativity, critical thinking, and problem-solving abilities. Furthermore, many students are unhappy with their current academic performance.

Many higher education institutions are attempting to improve their outputs and demonstrate their social and economic worth while ignoring potentially more worthwhile institutional missions and human pursuits. To meet these challenges, higher education must reform. Personal Learning Environments (PLEs), which represent a learning paradigm shift, have emerged as the reform path's top priority.

On the one hand, educational providers other than universities or colleges, as competitors, are devoted to developing their own commercial PLEs or similar personalized online learning platforms for learners or HE institutions, such as Symbaloo, Protopage, Dingding, Zoom, and so on.

To maintain a competitive advantage, HE must either annex the PLEs developed by other educational providers or purchase their service. Many higher education institutions are now creating their own PLEs with open-source social networks like Elgg (www.elgg.com), OSSN, DEMO, and others. Meanwhile, for some institutions that are not yet ready to develop their own PLEs, enterprises can provide a solution.

The sections that follow will focus on the external framework—the Society, Policy, Education, and Technology (SPET) framework—as well as the internal framework—Academic Administration, Teacher, Learner, and ICT (ATLI).

The external framework refers to environmental factors that, if not present, will prevent the paradigm shift toward PLEs from having a broad and long-term impact on HE. The internal framework refers to key stakeholders within HE who are driving the implementation of PLEs in HE.

5 The Society, Policy, Education, and Technology (SPET) Framework for PLEs

PLEs enable the development of learners' personal communicative and formative scenography to enhance learning that is halfway between formal and informal, encouraging them to organize and take responsibility for their own learning.

However, fully implementing PLEs in HE requires not only revolution, collaboration, and compromise among all key stakeholders as internal factors, but also the combined efforts of external environmental factors—Society, Policy, Education, and Technology (SPET), as shown in Fig. 2.

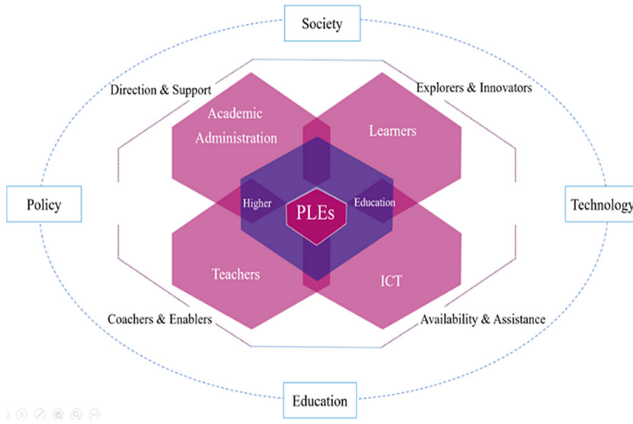


Fig. 2. The theoretical framework of PLEs

5.1 Society

The class system, cultural values, power structure, balance of individual freedom and social control, and degree of urbanization and industrialization all have a significant impact on any society's education system.

In retrospect, our society has come a long way, from the hunting society (Society 1.0), agricultural society (Society 2.0), industrial society (Society 3.0), and information society (Society 4.0), to the future super-smart society. Among them is the Society 5.0 vision, which is a human-centered imagination society in which humans, nature, and technology create a sustainable balance that is enhanced by super-smart AI data systems.

Economic growth in the current Society 4.0 or information society is knowledge-driven and innovation-driven, with the goal of achieving sustainable growth, more efficient resource use, and the formation of an “intelligent society” (Valentina 2017). Education is essential for social change.

For example, HE assists British society in transition from a colonial power to a welfare state and a globally oriented knowledge-based economy. To generate a new paradigm of social growth in the transition from Society 4.0 to Society 5.0, education should be transformed into smart education using smart technologies (technologies based on interaction and experience exchange). In terms of the paradigm shift from traditional HE to PLEs, society should indeed make changes to facilitate HE reform.

First, universities have long been regarded as a socioeconomic source of human capital and economic growth. Not only are knowledge, research, and degrees measured as the economic and societal value in the twenty-first century (Schofer et al. 2020), but academic sciences have become “economic engines” on which both the market and the state act.

However, the growing specialization of educational provision and the rapid expansion of universities are being discussed as the consumerization of higher education and the erosion of the university's public mission (Wright & Shore 2018).

The global race to excellence in higher education, for example, which is widely perceived and promoted as a driver of economic growth and competitiveness, has accelerated

(European Commission 2003; OECD 2008). To implement PLEs in higher education, the emphasis on HE's socioeconomic value must be reduced. For example, higher education should place a greater emphasis on its ability to foster desired states of being that allow individuals to pursue a meaningful future life course, to be proactive and societally engaged, and to carry the imprints of their historical legacies (Tomlinson 2021; Soysal and Baltaru 2021). Second, from a social-cultural standpoint, social structures, social norms, and social value systems should change.

Both formal and informal social structures influence individuals' thoughts, values, and behaviors. Individuals' opinions, beliefs, and behaviors can be shaped by formal organizations (Prell et al. 2010). According to organizational theory, shared values in formal organizations would evolve into organizational norms, which would then become concrete guidelines outlining the appropriate types of behavior for employees (Hill & Jones 2000).

The cultures of educational institutions can guide administrators', teachers', students', and parents' views, opinions, and subsequent actions regarding the implementation of PLEs in HE and the redefining of institutional logic and purpose, beliefs, approaches, and so on.

In terms of informal social structure, "social networks" refer to the similarity among individuals, including views, beliefs, and behavior (Lazarsfeld & Merton 1954), that would form a social attraction or a tie among individuals (McPherson et al. 2001), also known as the "homophily effect" in the social network literature. Individuals' views and behavior toward PLE implementation in HE would be guided or constrained by the heterogeneous population's understanding of PLEs.

According to Bicchieri (2006, p. ix), social norms are the "grammar" of social interactions. Norms are customary or typical behaviors that people engage in because they believe that is what others do or how you should behave (Reno et al. 1993). Changing social norms may be more important than changing behavior (Bicchieri & Xiao 2009; Krupka & Weber 2015).

Social values are idealistic criteria shared by the majority of a society that regulate and organize daily life. Sociological values are ideally unique to one group at a single point in time, according to sociology (Doan 2011). General and specific social values exist. General social values, for example, include universally accepted human values such as social equality, justice, and women's empowerment. Social values can be used to evaluate social norms both subjectively and objectively. Social norms are founded on social values that are justified by moral or aesthetic standards. Individuals' social values are built into their personalities and guide their thoughts and actions. Social norms can also have an impact on an individual's attitudes.

The development of PLEs in higher education necessitates a shift in social norms and social values regarding educational beliefs and values, particularly among parents, students, teachers, and administrators.

They should be aware of the PLEs paradigm from a macro-perspective, such as the acceptance of informal education and vocational education, the redefinition of talents, shifts in learning appraisal from score-oriented to competency-based or capability-realization, the prevalence of open education and open science, and credit recognition across universities or even countries.

Teachers, in particular, must arm themselves with new pedagogical concepts such as competency-based education, personalized learning, and so on; learn new teaching approaches such as teaching through enterprise projects and real work with local employers (Tyumaseva et al. 2020); form community partnerships; employ asset-based pedagogical approaches such as Culturally Responsive instruction, and so on; and shift their roles from instruction to mentoring, advising, and consulting; acquire new skills in the effective use of information technologies, mentoring, advising, and consulting; encourage self-motivation to unlock learners' potentials; build their academic confidence, sense of belonging; and ensure they can take ownership and responsibility for their own learning (Miliband 2006; Sun and Xu, forthcoming) to lead a life they believe is meaningful and valuable.

Administrators must support attainment- or mastery-based assessment, criteria-based, flexible, and continuous assessment of sub-competences, and the long-term use of e-portfolios for formative assessment to ensure assessment for learning and learning success.

To summarize, PLEs reform necessitates the acceptance of social norms. For example, if the social norm promotes educational equity by providing individuals with an environment conducive to their overall development, respecting education as a basic human right, and supporting open education and open data, more people will advocate for HE reforms toward PLEs.

5.2 Policy

To successfully implement PLEs in higher education, the policy must be redesigned from the education system to the HE organization infrastructure to reflect the value and goal of higher education.

First, full-time classroom-based education must be replaced by capacity-realization mode by integrating formal education, professional training, and entrepreneurship into PLEs and giving every single learner the opportunity to be the best they can be, regardless of talent or background (Miliband 2004). Learners can gain, improve, and retain the knowledge, competencies, skills, instincts, instruments, tools, and other resources needed to expand their capabilities. The underlying philosophy is that the social function of education has progressed through three stages, beginning with illiteracy eradication and ending with human development. With the rapid development of ICT, HE is entering a capacity-realization phase that corresponds to a higher level of human development. Learners are empowered to pursue alternative paths to self-actualization. Policymakers must now shift their attitude and mindset away from time-based education and toward competency-based and capacity-realization education. In other words, traditional programs require fixed units of time (semesters or terms) and a minimum number of courses or credits to complete the program. This single-model approach prevents many students from achieving their full potential. Time and space are much more flexible in competency-based or capacity-realization education for learners, who can learn through a variety of activities at their own pace and schedule.

Second, the infrastructure of the HE organization should be modified accordingly. For example, rankings have aided the current model of the university system and the organizational development that has surrounded it (Sauder & Espeland 2009; Baltaru &

Soysal 2018). To promote equity, diversity, distinctiveness, and uniqueness, we propose that universities be organized around disciplines and further subdivided by discipline attributes. We believe that grouping disciplines can optimize the layout of disciplines in higher education, promote diverse development, and avoid repetitive construction due to the pursuit of large and comprehensive disciplines. When rankings are made within disciplinary groups rather than university rankings, each discipline has access to resources based on its own characteristics rather than being sidelined by utilitarian factors in a university, which is especially true in the humanities. Students may also select a university or college based on discipline rankings rather than the university's overall ranking. Ranking universities, on the other hand, would increase inequity and utilitarianism.

For example, in China, the Double First-Class Initiative (China's largest education development scheme to date, aimed at increasing global recognition of China's university system by 2049) seeks to use the performance assessment approach to drive reform and strengthen the development of key disciplines. However, it will exacerbate the regional imbalance in higher education because the elected universities will receive more financial resources, human resources, material resources, and policy support (Liu et al. 2019).

In addition to the above-mentioned fields of policy, the following areas are also crucial to implementing PLEs in HE: first, structures to promote privacy and security standards, regulations, and legal protection should be established in HE since Artificial Intelligence applications will raise ethical issues and pose potential threats to the right to education; second, the method of evaluation and appraisal of HE should be more dynamic, focusing on HE's potential to enable individuals to pursue a meaningful future life (Tomlinson 2021); third, financial support is vital, which should invest more in digital learning through the development of online learning platforms; and a, tailored support and resources offered by instructional designers and technology coaches to faculty and staff in the form of consultation; training; and workshops (Archer-Kuhn et al. 2020; Morley & Clarke 2020).

5.3 Education

To fully implement PLEs in HE, we should redefine education's goals and values to get it back on track. As shown in Fig. 3, we believe that the value of education, from K–12 to higher education, is presented in three areas: social development, human development, and knowledge development.

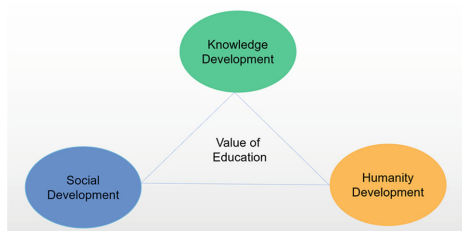


Fig. 3. The value of education

Aron (1965) defines social development as the “increasing moralization of human behavior and thought.” This is linked to educational systems. For example, the system of the twentieth century was largely based on a “one size fits all” model, similar to factory models of the previous century. Since the turn of the century, there has been a government push to create social systems that prioritize citizens, challenging the “old” model and creating systems that are more responsive to individual needs. The educational trend is shifting away from the industrial economy and toward a knowledge economy, where citizens must be equipped with a diverse set of personal, social, thinking, and learning skills. Education is thought to be inextricably linked to processes such as social mobility, community development, national integration, urbanization, and population change.

Education is essential for the recovery of human and social capital. Learning declines have a long-term negative impact on productivity and economic growth (Global Partnership for Education 2020). Higher education planning must include a strategic plan to meet societal needs and create job opportunities. The pandemic, for example, changed the entire concept of education by emphasizing community well-being and engagement. As a result, the entrepreneurial university model emerges, adding a strong third mission to Higher Education Institutions (HEIs) (Stolze 2020).

Meanwhile, education has equitably become the focal point of the recovery, with an emphasis on strengthening educational systems (Global Partnership for Education 2020). Reducing inequities in education will be the norm, and reimagining our educational systems will have a positive societal impact (Garca & Weiss 2020a, b).

To summarize, higher education is critical for increasing youth employment opportunities, reducing societal disparities, ensuring the inclusion of vulnerable and marginalized groups, and conducting impact research that generates long-term socioeconomic returns.

Knowledge development is the accumulation and transfer of knowledge. For one thing, in terms of knowledge accumulation, HE is seen as the pinnacle of “powerful knowledge”—a concept that has recently gained traction (Wheelahan 2012; Young & Muller 2013). HE also allows students to reflect on and evaluate various bodies of knowledge, as well as their own relationships to dominant scientific and socio-political regimes of truth. Another benefit of higher education is that it motivates and inspires people to explore the unknown world in search of knowledge and truth. Knowledge transfer can be accomplished by involving innovation partners, delivering innovative products (Păunescu & McDonnell Naughton 2020), sharing knowledge, conducting interdisciplinary research, and collaborating with colleagues from other fields.

In terms of humanity development, the model uses the Merriam-Webster definition: “compassionate, sympathetic, or generous behavior or disposition: the quality or state of being humane.” Higher education promotes both personal and social growth. Individually, higher education empowers students by increasing their agential freedom. Higher education provides a set of resources from which a person can make positive decisions about the course of their lives and how to cultivate this further. The dispositions or “qualities” identified by Barnett (2018) as central to the concept of contemporary university education—willingness to learn, engagement, openness to experience, and resilience—entail processes of ontological change that are now potentially as valuable as formal

knowledge acquisition. For the development of social humanity, higher education promotes, among other things, citizenship education, human rights, democracy, cultural diversity, sustainability, national history, security, and moral values.

According to Tomlinson (2021), current HE is valued in three ways: first, the continued movement toward systemic massification; second, the movement toward marketization and the entrepreneurial and transactional orientation of academics, students, and senior managers; and third, the graduate employment rate in a more flexible and uncertain labor market. All three methods of evaluation center on the monetary benefits of higher education, which drives institutions to be organized along market logic in order to meet market-related desires. The value-for-money metric tells us little about how HE's transformational potential manifests itself or how and in what ways HE affects our future lives. Educational research has also contested and critiqued these value framings (Crawford & van der Erve 2015; Downs 2017; Tomlinson 2018). The value of higher education is centered on its ability to foster desired states of being that allow individuals to pursue a meaningful future life course.

5.4 Technology

With increased digitalization, connectivity, and the use of Artificial Intelligence (AI), collective intelligence technologies, and knowledge management systems, it is recognized that the digital world is becoming smarter than Big Data technologies.

Meanwhile, the concepts of "knowledge society" and "smart society" persisted after the "information society," as evidenced by the number of FTTH network subscribers, download speed, broadband access quality, and so on.

When it comes to Society 5.0, technology is supposed to help build innovative mechanisms for a smart society where society can self-regulate through smart technologies and trained citizens.

To fully implement PLEs, the role, function, and value of technology in education must be altered.

First, technology's role must shift. Technology is no longer an afterthought in education; rather, it should be a key decision-making factor. Technology has the potential to improve data collection and analysis in order to facilitate educational decision-making and practices. For example, during COVID-19, some countries were able to respond quickly to the pandemic due to their prior stance on technology access for education (see, for example, the International Council of Education Advisers Report 2018–20, published by the Scottish Government in December 2020). For another, as technology becomes more prevalent in daily life, everyone will need to "access, manage, understand, integrate, communicate, evaluate, and create information" in order to participate in the working, social, and political spheres.

Second, the role of technology must evolve. In the digital age, technology and the internet will become a source of water and electricity, as well as a service for all as a human right. Technology has evolved into a powerful tool for overcoming inequalities, which could be realized by utilizing technology to serve disadvantaged populations while also ensuring its accessibility and affordability for all. Technology transforms the organization of HE institutions (e.g., the cross-fields institute, transcending disciplines,

etc.), curriculum design, and pedagogical mode in the HE field (e.g., virtual reality, augmented reality, personalized learning) at the micro level.

Third, the value of technology must change. Technology is currently transforming our perceptions and expectations of education, with concepts such as lifelong learning, mutual credit recognition, personalized learning, and so on taking root in our minds. In other words, sustainability has infiltrated higher education, necessitating seamless, agent-based, and technology-driven intelligent learning (Menon & Suresh 2020).

Technology helps to develop a lifelong learning system that extends beyond the boundaries of educational institutions and areas and organizes lifelong learning opportunities in a more flexible manner, especially at the local level (Mazzucato 2018). As a result, HE can take advantage of digitalized systems and technology-assisted processes, and strategies for demonstrating flexibility, sustainability orientation, and impact achievement are developed.

People, things, and systems will all be connected in cyberspace in the coming “Society 5.0,” and the massive amount of data from sensors in physical space will be analyzed by artificial intelligence (AI) and fed back to humans in physical space. Self-actualization through education becomes possible when people have the freedom and technical support to fully express their talents and create values for society. The new value created by innovation will, in turn, eliminate regional, age, gender, and language gaps, allowing for the provision of products and services finely tailored to diverse individual needs and latent needs.

6 The Four Perspectives of PLEs Scales

Successfully implementing PLEs in higher education necessitates the collaboration of four major stakeholders or perspectives in the educational field, namely Academic Administration, teachers, students, and ICT. A set of PLE rubrics that ensure quality control and assurance for the use of PLEs in higher education is necessary. The four scales are known as the PLEs Teacher and Teaching Scale (PLEs-TTS), PLEs Learner and Learning Scale (PLEs-LLS), PLEs Academic Administration Scale (PLEs-AAS), and PLE Technology Scale (PLEs-TS).

We believe that the current slow adoption of PLEs in higher education is primarily due to a lack of consensus among various stakeholders, which has resulted in a demand and supply gap between users and developers. The following sections explain the roles and relationships of the four PLE perspectives. The roles of the four perspectives are shown in Fig. 4.

As shown in Fig. 4, the implementation of PLEs in higher education necessitates cooperation and collaboration among the four perspectives. Academic administration is defined as “the activities are undertaken by academic workers to achieve effective teaching, research, and community service in accordance with the institution’s prescribed regulations” (KOKO 2011). Academic administration’s primary role is to set clear expectations and criteria for the other three perspectives, which include, student recruitment and services, faculty development and assessment, policy and planning, financial and budgeting, and purchasing goods and equipment.

In the context of PLEs, instructors, among other things, serve as designers of learning experiences and facilitators of learning (Francom 2014; Reigeluth & Karnopp 2013).

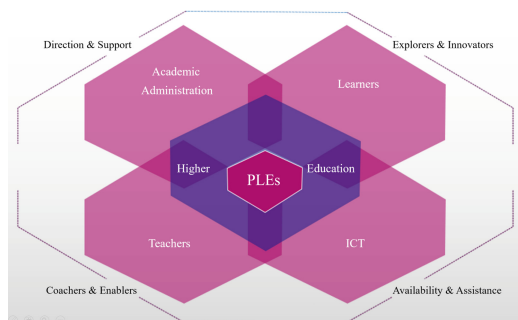


Fig. 4. The roles of the four perspectives in PLEs

Teachers, as designers, create sequences of learning tasks to provide an appropriate level of challenge for students and determine how to personalize and differentiate tasks based on students' prior knowledge. They must treat learners as explorers and innovators and use asset-based transformative pedagogical principles to ensure that students, especially those from historically and systemically underserved backgrounds, feel a part of the academic community and have the confidence and skills to successfully navigate PLEs and achieve their learning goals and potentials (Sun and Xu forthcoming). They must learn skills in designing and managing online collaborative strategies for appropriate online communication, as well as developing and/or reviewing current technologies integrated into their teaching area to enhance learning (Segrave et al. 2005).

More importantly, as facilitators, they must explain the learning task and demonstrate how different parts of the task should be completed; recommend appropriate resources; provide coaching and feedback to learners as they work on learning tasks; and assist students in learning on their own.

Learners in PLEs become masters of their own learning, transitioning from passive receivers of information to self-directed and collaborative learners, as well as active knowledge explorers (Reigeluth 2016; Reigeluth & Karnopp 2013). They investigate and direct their learning, solve problems, and conduct self-analysis, self-control, and self-evaluation of their learning process.

They develop into collaborative learners who teach and learn from one another through discussions, teamwork, presentations, peer critique, and other activities (Reigeluth & Karnopp 2013). PLEs may also aid in the development of higher-order thinking (Elfeky 2018).

In summary, learners gain the ability to plan their educational and self-educational activities, manage their time effectively, work productively with educational materials, and monitor the outcomes of their work.

In terms of ICT, the primary role is to provide access to and support for PLEs used by teachers, students, and administrators. In terms of accessibility, the growing use of data for Learning Analytics and the use of AI in education (e.g., automated feedback systems, recommender systems, and content creation systems) offer significant potential for the development of PLEs by providing feedback, opportunities for reflection, and recommendations based on learner data. In terms of assistance, studies and PLE

workshops and projects have repeatedly demonstrated that students lack confidence and competence in using technology to create their own PLEs.

In fact, most students’ digital practice is informal, and they are more accustomed to using apps than applications (JISC 2009). Effective PLEs are complex applications, and many students and teachers require structured assistance in using, let alone developing, PLEs.

To sum up, as the Organization for Economic Co-operation and Development (OECD) suggested,

Education systems will need to adapt to the changes brought on by automation, teaching children and teenagers the skills necessary to fully benefit from the present wave of technological implementation. This involves both cognitive and social intelligence, as well as the abilities required to function effectively in a digital environment, both as specialists and as users of digital technology.

Academic Administration establishes rules, principles, and regulations, as well as supports and procedures to ensure that such policies, standards, and regulations are followed by learners and teachers via ICT-integrated assessment and appraisal. In turn, learners and teachers follow those regulations and requirements in order to fulfill their duties and accept administrative supervision. In PLEs, learners and teachers collaborate to achieve learning. Teachers can use technology to boost their productivity, implement useful digital tools to expand students’ learning opportunities, and increase student support and engagement. It also enables teachers to improve their instructional methods and personalize learning. Learners use ICT to create their own learning environment, access and analyze information, interpret and transform that information into their own personal knowledge (Ertmer & Ottenbreit-Leftwich 2013), and develop their learning confidence and competence as lifelong learners. ICT provides support and access to academic administration, learners, and teachers.

From an educational standpoint, perhaps the most significant aspect is that the paradigm shift to PLEs must include more than just a change in the tools or resources used, but must also promote a change in attitude in which students have the opportunity to take an active role in the teaching and learning process, of which they are an essential part through self-regulated learning. The teacher, in turn, possesses the various skills and abilities required to modify the design and planning of instructional situations in the context of teaching. Meanwhile, academic administrators accept the competency-based

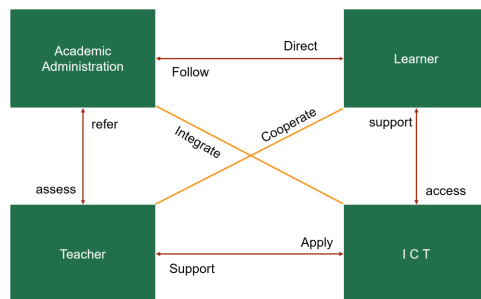


Fig. 5. The relationship between the four perspectives in PLEs

or capacity-realization mode of education in order to establish appropriate policies, principles, and standards. And the ICT department understands the principle of PLEs, makes sure ICT resources are accessible, and provides proper support for learners, teachers, and administrators. The relationship between the four perspectives is shown in Fig. 5.

7 Conclusion

The pandemic has fueled public debate about the importance of higher education in general, and specific models of higher education in particular. According to some researchers, the current higher education system is out of date in terms of cost-return and breadth (Psacharopoulos 1972; Bigliardi et al. 2005).

Meanwhile, higher education institutions are in flux, as are the expectations and practices of teachers, students, and parents. The one-size-fits-all approach is fading, and the idea of PLEs is gaining traction. The most recent advancements in information technology and digital content creation allow all students to benefit from the PLE approach.

In the context of Personalized Learning Environments, the purpose or value of higher education is to facilitate the realization of learners' capabilities through learner-driven and learner-designed education. However, the development of the PLEs agenda is not about quick fixes but about strengthening the educational system's capacity to meet the learning needs of all learners, regardless of age. Competency-based or capability-realization education necessitates learners understand their current capabilities through objective cloud-based assessment and then rationally gain knowledge, experience, and attitude in a self-driven paradigm to awaken or transcend their potential.

The evaluation of successful learners, as well as the productivity and effectiveness of education, is based on competency development and personal capability fulfillment rather than criterion-referenced measurement. Nonetheless, our Policy, Education, Society, and Technology (PEST) must all evolve in order to facilitate the implementation of PLEs in higher education. PLEs seek to re-establish the value of higher education in terms of social development, knowledge development, and human development.

References

- Adell, J.: Más allá del instrumentalismo en tecnología educativa. Cambiar los contenidos, cambiar la educación. Madrid: Morata (2018)
- Adell, J., Castañeda Quintero, L.J.: Los Entornos Personales de Aprendizaje (PLEs): una nueva manera de entender el aprendizaje (2010)
- Almenara, J.C.: Tendencias para el aprendizaje digital: de los contenidos cerrados al diseño de materiales centrado en las actividades. El Proyecto Dipro 2.0. Revista de Educación a Distancia (RED), (32) (2012)
- Amine, M.: Mohamed Amine Chatti's ongoing research on Knowledge and Learning (2009). <http://mohamedaminechatti.blogspot.com/2009/04/ple-pkn.html>
- Archer-Kuhn, B., Ayala, J., Hewson, J., Letkemann, L.: Canadian reflections on the Covid-19 pandemic in social work education: from Tsunami to innovation. Soc. Work. Educ. **39**(8), 1010–1018 (2020)
- Aron, B.: Le Developpement Social. Congress et Colloques (1965)

- Attwell, G.: Personal learning environments-the future of eLearning. *Elearning Papers* **2**(1), 1–8 (2007)
- Attwell, G.: Personal learning environments: looking back and looking forward. In: Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'21), pp. 522–526 (2021)
- Ballard, J., Butler, P.: Personalised learning: developing a vygotskian framework for e-learning. *Int. J. Technol. Knowl. Soc.* **7**(2), 21 (2011)
- Baltaru, R.D., Soysal, Y.N.: Administrators in higher education: organizational expansion in a transforming institution. *High. Educ.* **76**(2), 213–229 (2018)
- Barnett, R.: *The Ecological University: A Feasible Utopia*. Routledge (2017)
- Bicchieri, C.: *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press (2005)
- Bicchieri, C., Xiao, E.: Do the right thing: but only if others do so. *J. Behav. Decis. Mak.* **22**(2), 191–208 (2009)
- Bigliardi, B., Petroni, A., Dormio, A.I.: Organizational socialization, career aspirations and turnover intentions among design engineers. *Leadersh. Org. Dev. J.* **26**(6), 424–441 (2005)
- Dabbagh, N., Castaneda, L.: The PLE as a framework for developing agency in lifelong learning. *Educ. Tech. Res. Dev.* **68**, 3041–3055 (2020)
- Dabbagh, N., Kitsantas, A.: Personal learning environments, social media, and self-regulated learning: a natural formula for connecting formal and informal learning. *Internet High. Educ.* **15**(1), 3–8 (2012)
- Dolors, R.: Entornos personalizados de aprendizaje (2009). <http://www.slideshare.net/dreig/ple-1340811>
- Downes, S.: New technology supporting informal learning. *J. Emerg. Technol. Web Intell.* **2**(1), 27–33 (2010)
- Downs, Y.: Furthering alternative cultures of valuation in higher education research. *Camb. J. Educ.* **47**(1), 37–51 (2017)
- Elfeky, A.I.M.: The effect of personal learning environments on participants' higher order thinking skills and satisfaction. *Innov. Educ. Teach. Int.* (2018)
- Ertmer, P.A., Ottenbreit-Leftwich, A.: Removing obstacles to the pedagogical changes required by Jonassen's vision of authentic technology-enabled learning. *Comput. Educ.* **64**, 175–182 (2013)
- Fiedler, S., Pata, K.: Distributed learning environments and social software: in search for a framework of design. In: *Social Computing: Concepts, Methodologies, Tools, and Applications*, pp. 403–416. IGI Global (2010)
- Francom, G.M.: Educational Technology Use among K-12 Teachers: What Technologies Are Available and What Barriers Are Present?. Online Submission (2016)
- Henri, F., Charlier, B., Limpens, F.: Understanding ple as an essential component of the learning process. In: *EdMedia+ Innovate Learning*, pp. 3766–3770. Association for the Advancement of Computing in Education (AACE) (2008)
- Hill, C.W., Jones, G.R.: *Cases in Strategic Management* (5th ed.). Houghton Mifflin Company (2000)
- Humphris, D.: Multiprofessional working, interprofessional learning and primary care: a way forward? *Contemp. Nurse* **26**(1), 48–55 (2007)
- Kalz, M., Schön, S., Lindner, M., Roth, D., Baumgartner, P.: *Systeme im Einsatz-Lernmanagement, Kompetenzmanagement und PLE* (2011)
- Krupka, E.L., Weber, R.A.: Identifying social norms using coordination games: why does dictator game sharing vary? *J. Eur. Econ. Assoc.* **11**(3), 495–524 (2013)
- Lazarsfeld, P.F., Merton, R.K.: Friendship as a social process: a substantive and methodological analysis. *Freedom Control Mod. Soc.* **18**(1), 18–66 (1954)

- Liu, Q., Turner, D., Jing, X.: The double first-class initiative in China: background, implementation, and potential problems. *Beijing Int. Rev. Educ.* **1**(1), 92–108 (2019)
- Mazzucato, M.: Mission-oriented research & innovation in the European: A problem-solving approach to fuel innovation-led growth (2018)
- McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: homophily in social networks. *Ann. Rev. Sociol.* **27**(1), 415–444 (2001)
- Menon, S., Suresh, M.: Synergizing education, research, campus operations, and community engagements towards sustainability in higher education: a literature review. *Int. J. Sustain. High. Educ.* **21**(5), 1015–1051 (2020)
- Miliband, D.: *Personalised Learning : Building a New Relationship with Schools* (2004)
- Miliband, D.: Choice and voice in personalised learning. *School. Tom. Personalising Educ.* 21–30 (2006)
- Moore, M.: On a theory of independent study. In: *Distance Education*, pp. 68–94. Routledge (2020)
- Morley, C., Clarke, J.: From crisis to opportunity? innovations in Australian social work field education during the COVID-19 global pandemic. *Soc. Work. Educ.* **39**(8), 1048–1057 (2020)
- Prell, C., Reed, M., Racin, L., Hubacek, K.: Competing structure, competing views: the role of formal and informal social structures in shaping stakeholder perceptions. *Ecol. Soc.* **15**(4) (2010)
- Psacharopoulos, G.: Rates of return to investment in education around the World. *Comp. Educ. Rev.* **16**(1), 54–67 (1972)
- Rahimi, E., van den Berg, J., Veen, W.: Facilitating student-driven constructing of learning environments using Web 2.0 personal learning environments. *Comput. Educ.* **81**, 235–246 (2015)
- Reigeluth, C.M.: Instructional theory and technology for the new paradigm of education. *Revista de Educación a Distancia (RED)*, (50) (2016)
- Reigeluth, C.M., Karnopp, J.R.: Reinventing schools: It's time to break the mold. *R&L Education* (2013)
- Reno, R.R., Cialdini, R.B., Kallgren, C.A.: The transsituational influence of social norms. *J. Pers. Soc. Psychol.* **64**(1), 104 (1993)
- Sauder, M., Espeland, W.N.: The discipline of rankings: tight coupling and organizational change. *Am. Sociol. Rev.* **74**(1), 63–82 (2009)
- Schofer, E., Ramirez, F.O., Meyer, J.W.: The societal consequences of higher education. *Sociol. Educ.* **94**(1), 1–19 (2021)
- Sebba, J., Brown, N., Steward, S., Galton, M., James, M.: *An Investigation of Personalised Learning Approaches used by Schools*. DfES Publications, Nottingham (2007)
- Segrave, S., Holt, D., Farmer, J.: The power of the 6 three model for enhancing academic teachers' capacities for effective online teaching and learning: benefits, initiatives and future directions. *Australas. J. Educ. Technol.* **21**(1) (2005)
- Selwyn, N.: *Education and Technology: Critical Questions* (2017)
- Soysal, Y.N., Baltaru, R.D.: University as the producer of knowledge, and economic and societal value: the 20th and twenty-first century transformations of the UK higher education system. *Europ. J. High. Educ.* **11**(3), 312–328 (2021)
- Stolze, A.: A meta-ethnography on HEIs' transformation into more entrepreneurial institutions: towards an action-framework proposition. *Ind. High. Educ.* **35**(1), 14–27 (2021)
- Tomlinson, M.: Conceptions of the value of higher education in a measured market. *High. Educ.* **75**(4), 711–727 (2017). <https://doi.org/10.1007/s10734-017-0165-6>
- Tomlinson, M.: Missing values: engaging the value of higher education and implications for future measurements. *Oxf. Rev. Educ.* **48**(1), 46–62 (2022)
- Torres Kompen, R.: *Personal Learning Environments Based on WEB 2.0 Services in Secondary and Higher Education* (2016)

- Tyumaseva, Z.I., et al.: Psychological and social aspects of innovations and standardization in education. *Bull. Natl. Acad. Sci. Republic Kazakhstan* **4**, 274–283 (2020)
- van Harmelen, M.: Personal learning environments. In: *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)*, pp. 815–816 (2006)
- Van Harmelen, M.: Design trajectories: four experiments in PLE implementation. *Interact. Learn. Environ.* **16**(1), 35–46 (2008)
- Veletsianos, G., Moe, R.: The rise of educational technology as a sociocultural and ideological phenomenon. *Educause Rev.* **4** (2017)
- Voronkova, V., Kyvliuk, O.: Philosophical reflection smart-society as a new model of the information society and its impact on the education of the 21st century. *Future Hum. Image* **7**, 154–162 (2017)
- Wheelahan, L.: *Why Knowledge Matters in Curriculum: A Social Realist Argument*. Routledge (2012)
- Young, M., Muller, J.: On the powers of powerful knowledge. *Rev. Educ.* **1**(3), 229–250 (2013)



Construction and Application of “AI + RPA + X” Intervention Field for College Students’ Learning Adaptation

Yi Qiu^(✉), Zhiyang Peng, and Yi Lu

School of Educational Information Technology, South China Normal University, No. 55, West of Zhongshan Avenue, Tianhe District, Guangzhou, China
etqiuyi@163.com

Abstract. The research focuses on the construction and application of learning adaptation intervention field for college students in the age of intelligence, uses research methods of survey research, literature research, theoretical deduction, and design-based research, under the guidance of field theory, combined with the development direction of digital transformation, proposed the concept of college students’ learning adaptation intervention field construction of factor connectivity, link routing and data integration; applied technologies such as AI and RPA to build an “AI + RPA + X” intervention field that highlights the characteristics of human-machine synergy, and clarified its hierarchical structure and operating mechanism; carried out a one-semester practical application for 49 undergraduate normal students majoring in educational technology in the 2021 grade of H University, and analyzed the application effect of “AI + RPA + X” intervention field. The results show that the “AI + RPA + X” intervention field of college students’ learning adaptation can provide theoretical and practical reference for constructing and optimizing the existing field and interfering with college students’ learning adaptation.

Keywords: College students’ learning adaptation · Intervention field · AI + RPA + X · Construct · Apply

1 Introduction

The new infrastructure construction, represented by big data center and artificial intelligence(AI), has gradually built a “digital pedestal” for the reform of enabling education and teaching. The mutual empowerment of scene-based learning and “intelligence +” learning environment constantly enriches college students’ learning experience, continuously improves college students’ learning self-efficacy, and promotes college students to maintain high learning motivation in the process of participating in learning activities, in order to reduce their learning inadaptability. Field refers to a relational framework endowed with specific gravity (Bourdieu 1992), and the intervention field of college students’ learning adaptation is the external environment framework with specific gravity that supports the object of intervention (college students) and implements intervention

activities. Most of the existing studies on college students' learning adaptation intervention field discuss the construction method and teaching innovation application of general learning field, but its structure and operating mechanism are not clear. Based on this, the study proposes the following three research questions:

RQ1: What concepts should be followed in constructing the intervention field of college students' learning adaptation under the digital transformation?

RQ2: What is the hierarchical structure and operation mechanism of the "AI + RPA + X" intervention field of college students' learning adaptation?

RQ3: What is the effectiveness of the "AI + RPA + X" intervention field for college students' learning adaptation?

The rest of this article is structured as follows. Section 2 reviews the progress of related research, further clarifies the work to be done in this study and emphasizes its significance. Section 3 introduces the process and methods of the research. Section 4 proposes the concept of constructing the intervention field of college students' learning adaptation with factor connectivity, link routing, and data integration. Section 5 constructs the intervention field of "AI + RPA + X" (AI + Robotic Process Automation + Scenario Linkers), which is characterized by human-machine synergy, and clarifies its hierarchical structure and operation mechanism. Section 6 carries out a one-semester practical application for 49 undergraduate normal students majoring in educational technology in the 2021 grade at the University of H, and analyzes the application effect of "AI + RPA + X" intervention field.

By solving the above three problems, the study provides theoretical references and practical lessons for the construction and optimization of the existing field to intervene in college students' learning adaptation.

2 Related Work

The study summarizes the literature on the intervention field of college students' learning adaptation and the construction and application of learning field supported by RPA. Retrieving 1093 articles from 2000 to 2023 through "Web of Science" and other database, excluding the less relevant ones, it is found that with the accelerated evolution of scientific and technological revolution, college students' learning adaptation intervention field is developing towards the direction of dataization and intelligence. However, it is still necessary to further clarify its hierarchical structure and operating mechanism, so as to provide theoretical and practical guidance for the construction of intervention field that reflects the full connectivity of learner data and resource data, provides personalized and customized services for learners.

2.1 The Present Situation of Intervention Field Research on College students' Learning adaptation

The main focus of research on college students' learning adaption is to explore how to effectively implement interventions domestically and internationally. Researchers have explored intervention strategies, methodological models and model building. For example, T.-z et al. (2011) conducted a questionnaire survey on 160 college students, and proposed methods and strategies to improve students' learning adaptability in the network

environment. Huanqing Qiu (2010) proposed an integrated intervention model based on the three major factors of college students’ online learning effectiveness, combined with the self-management learning representation model. Taking vocabulary learning in college English courses as an entry point, Xu D et al. (2021) constructed a multimedia-based interactive English teaching adaptation intervention model. In terms of field construction and application, Scepanovic S et al. (2012) connected the data of students’ learning styles to the adaptive learning system, in order to provide students with more suitable courses. Aiming at the limitations of rule-based adaptation, Zhao T et al. (2017) further proposed a framework for the generation and evolution of software adaptation rules based on reinforcement learning. Nguyen A et al. (2018) integrated learning analysis, proposed a game analysis framework specific to people with intellectual disabilities, so as to create adaptive learning environments. Synthesizing the existing research, it is found that most of the fields of intervention for college students’ learning adaptation can provide functional support for the implementation of intervention activities, and their internal components, structure, and operating mechanism are not clear enough.

2.2 The Present Situation of Construction and Application of Learning field Supported by RPA

RPA refers to the simulation of various rule-based and repeatable processes without real-time creativity or judgment, “bringing robots out of humans”, with the characteristics of repeatability, determinism, and clarity (Zhang L. F. 2018, p. 303). At present, the research on learning fields supported by RPA mainly focuses on management governance, teaching practices, and scenario applications. In terms of management governance, Mosteanu et al. (2022) assisted the students’ enrollment registration process through RPA, and helped professors conduct student evaluation processes to improve the effectiveness of teaching, research and study programs, for the purpose of improving academic study programs. Gupta A et al. (2023) developed RPA-driven Admission Management System (R.P.A.A.M.S) to manage and verify student information throughout the process. In terms of teaching practice, Sutipitakwong, S. et al. demonstrated that RPA8 can be used as a teaching resource to support the teaching and learning process, which has been used as part of students’ visual and verbal interactions. In terms of scenario application, Somasundaram M et al. (2022) described a reporting tool developed using UiPath, which was successfully implemented in the Department of Electronics and Communications (ECE) of R.M.K., in order to generate reports that provide student information based on the principles of AI and automatically send them to each student’s email for program improvement. Synthesizing the existing studies, it is found that RPA has already taken part of the automation work in different scenarios of education and teaching, but accessing it to AI and existing teaching tools to build a digital-intelligence-integrated learning field needs to be further explored.

Compared with the existing research, this study not only focuses on the new trend of field construction brought by technological development, but also pays more attention to the integration of college students’ learning adaptation and field theory, trying to clarify the composition of “AI + RPA + X” intervention field of college students’ learning adaptation from the structure, providing direction for optimizing the existing field. At the same time, the study focuses on the function of intervening learning adaptation to

help college students adapt to university learning quickly in the digital age and improve learning effectiveness.

3 Research Process and Methods

This study utilizes a combination of research methods such as survey research, literature research, theoretical deduction, technical realization and design-based research to carry out the study, and the specific research process and methodology are shown in Fig. 1.

During the phase of current situation survey and literature analysis, through the methods of survey research and literature research we investigated the existing intelligent learning fields domestically and internationally, clarified the points of convergence between learning field and learning adaptation, then constructed a hierarchical model of learning adaptation intervention fields for college students in the intelligence era. During the phase of field conduction, through the methods of theoretical deduction and technical realization, combined with AI assistants, RPA, and Scenario Linkers, selected existing fields such as cyberspace, hybrid space, and ubiquitous space, focusing on “core concepts, key technologies, supporting tools, operating mechanism”, we constructed the “AI assistant + PRA + X” intervention field for college students learning adaption. During the phase of field conduction, through the methods of theoretical deduction and technical realization, we carried out a one-semester practical application for 49 undergraduate normal students majoring in educational technology in the 2021 grade of H University, collected and analyzed the data of CSLAli scale measurements and learning satisfaction, so as to verify the effectiveness of the “AI + PRA + X” intervention field.

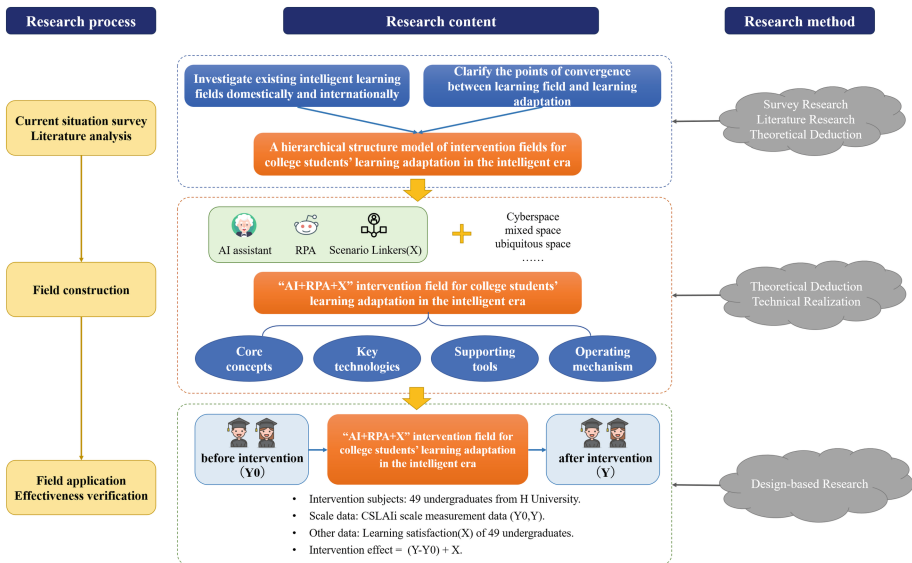


Fig. 1. Research process and methods

4 Construction of “AI + RPA + X” Intervention Field for College Students’ Learning Adaptation

4.1 Concept of Construction

Pierre Bourdieu pointed out that a field refers to a network or a structure of objective relations between locations that are objectively defined, and each field is defined by a market that connects producers and consumers in the field (Li Q 2002). The learning field, i.e., the evolution of the concept of field towards autonomy in the field of education, and the fundamental tendency to make learning activities manifest. The intervention field described in the study, as a specific type of learning field, has more artificial, unnatural and explicit features, which points to the goal of “intervention on the learning adaptation of college students”. The field of intervention on college students’ learning adaption is the external environment with a specific gravitational structure that supports the target (college students) and implements the intervention activities. In short, the field should not only reflect the inherent nature of the learning field, but also meet its specific goals. Based on the concepts of factor connectivity, link routing, and data integration, the study constructed an intervention field for college students’ learning adaptation.

Factor Connectivity

With the deepening of “Internet +”, the boundary between online space and offline space is gradually blurred, and blended learning is becoming the new normal in universities. Connectivity, as an essential feature of “Internet +”, extends the boundaries of the learning field and strengthens the links between learners, learning resources, service providers, managers and data.

Link Routing

The learning field can be regarded as a learning-themed flow network formed through the connection of factors. In order to fully utilize the effectiveness of the learning field, it is especially important to build links that point to factors and dynamically allocate them. Drawing on the idea of network layer routing in the OSI architecture, the learning field also needs to rely on the form of link routing to support its efficient operation, and the link routing shows the rule-based and automated characteristics, which clarifies the direction for the construction of the learning field.

Data Integration

Digitalization is an inevitable trend of educational reform and development, and it provides a material basis for highlighting data as an emerging factor of production. We should promote multi-modal data integration in multiple dimensions and throughout the process to realize the value of data. Based on the high reliability and credibility of structured data, semi-structured data, and unstructured data, the unified processing is carried out with specific standards, in order to realize the “fusion” between data, lay the foundation for data integration, which promote the full business of data circulation and enhance its interpretability.

4.2 Supporting Technology

With the rapid development of a new generation of information technology represented by generative artificial intelligence, AI is reshaping the production mode of society and creates a huge impact on people's way of life and cognition. The intervention field for college students' learning adaptation consists of AI-centered technology systems, mainly including key supporting technologies such as AI assistants/AI agents, RPA, and scenario linkers(X). AI assistants/AI agents, as the main body of the service, can generate personalized learning content for students, provide customized learning strategies and resource push, and answer students' questions in the learning process, i.e., partially fulfilling the roles of producers, peers, service providers and managers. RPA technology simulates human operating steps and automates business processes to realize the collaborative processing of different tasks between humans and machines, so that the learning field has the best efficiency. The Scene Linker (X) technology realizes data connection between applications through the API between different applications.

4.3 Hierarchical Structure

In line with the development direction of digital transformation of education, we take the factor connectivity, link routing and data integration as the construction concepts, highlight the role of data elements, integrate the system engineering and divide-and-conquer ideas, and rely on AI, RPA and other technologies to form a hierarchical structure model of intervention field for college students' learning adaptation in the age of intelligence, as shown in Fig. 2; The model consists of four levels: carrier layer, connection layer, link layer and session layer.

The carrier layer is a large-scale data set composed of multi-source and multi-modal data, which is the "digital base" of the whole intervention field. In this layer, structured data such as student status, course selection and grades, semi-structured data such as behavior logs generated by students in the learning process, and unstructured data such as documents, text, pictures, audio and video, together constitute a "digital portrait" describing students' own characteristics, behavior habits, learning style and so on, which provides material basis for the upper layer to carry out the connection between applications, and provide evidence source for the implementation of scale personalized learning adaptation intervention.

The connection layer is an application connection flow composed of multi-application and multi-interface connections, which is the supporting link of machine feedback in the whole intervention field. In this layer, the application connector standardizes the platform, resources and tools to establish the application connection flow for data aggregation, exchange and transfer in the carrier layer, so as to achieve a specific intervention process. Different application connection flows correspond to different intervention processes, and can be generated and adjusted dynamically, which provide process support for automatic intervention at the upper level, and give an operation plan for the implementation of scale personalized learning adaptation intervention.

The link layer is an automatic execution network generated by multi-rule and multi-process mapping, which is the core component of supporting human-machine synergistic

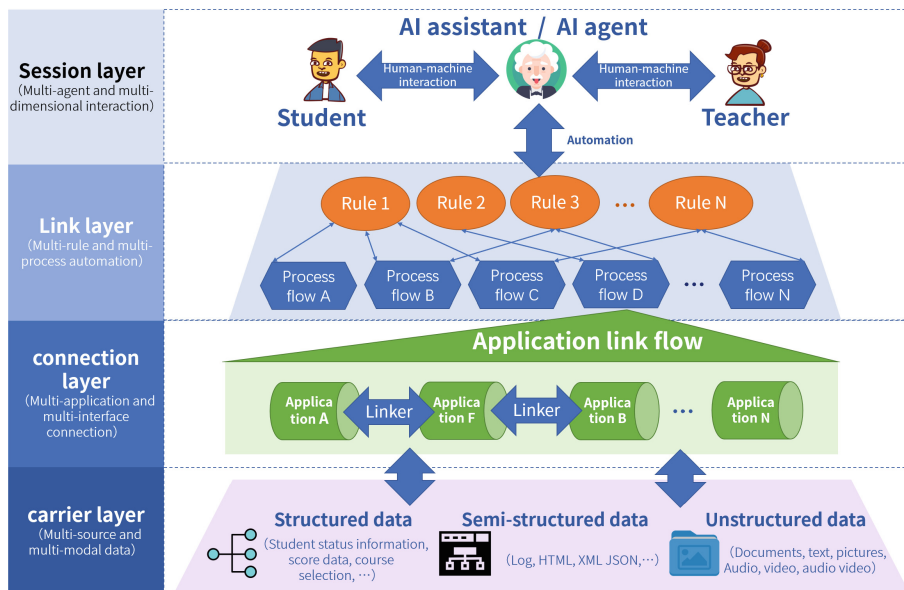


Fig. 2. “AI + RPA + X” intervention field hierarchy model for college students’ learning adaptation in the intelligent age.

intervention in the whole intervention field. This layer is a multimapping network composed of N rules and N processes, and automatic execution is realized by RPA (Robotic Process Automation). RPA can act as a fast executor of various artificial intelligence algorithm models, and establish a linkage channel among AI assistants, AI agents, and AI algorithm models, so that data at the carrier layer can flow continuously and accurately between applications at the link layer, generate credible decision-making suggestions and provide a connectivity channel for implementing individualized learning adaption interventions at scale.

The session layer is a value experience ring formed by multi-agent and multi-dimensional interaction, and it is the part of the whole intervention field that shows the process of human-machine synergistic intervention and feedback the effectiveness of intervention. This layer is composed of student groups, teacher groups, and artificial intelligence agents or assistants. Artificial intelligence agents or assistants can assist students and teachers to establish effective connections, and provide students and teachers with appropriate feedback and appropriate guidance according to the specific situation of human-machine interaction, provide an entrance and entry point for implementing individualized learning adaption interventions at scale.

4.4 Operating Mechanism

According to the hierarchical structure model of college students’ learning adaptation intervention field in the age of intelligence, combined with the core concepts, key technologies and supporting tools of “AI + RPA + X” intervention field construction, around

the generation, circulation and application of data points, the research proposed the intelligent era's college students' learning adaptation field operation mechanism, which characterized by refining key multi-dimensional data points, binding multi-application connection flow, establishing automatic execution rule chain and forming a session set for human-machine interaction, as shown in Fig. 3.

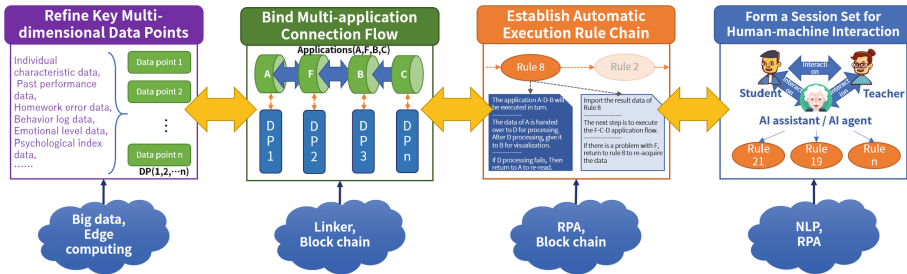


Fig. 3. The operating mechanism of college students' learning adaptation intervention field in the intelligent era.

Refine Key Multi-dimensional Data Points

Refining the key multi-dimensional data points provides the data basis for the connection of multi-elements between the whole intervention fields. By refining and analyzing the multidimensional data of college students' individual characteristics, past achievements, wrong homework questions, behavior logs, emotional level and psychological index, we can find out a number of key data points, and analyze which learning application can generate corresponding data, so as to provide a basis for establishing application connection flow. Using big data to mine the association and implied information between data, and using edge computing distributed processing can improve the granularity of data, in order to provide a more credible basis for the construction of application connection flow.

Bind Multi-application Connection Flow

Binding multi-application connection flow is to establish a channel for data flow, thus providing a trusted routing network for data exchange. On the basis of multi-dimensional data points, find matching teaching and learning applications and carry out semantic binding, and establish connection flows between applications, which can provide support for the automatic execution of corresponding intervention rules. In this link, the standardized exchange of data between applications can be realized by using connectors, and the consistency in the process of data exchange can be ensured by using blockchain, thus providing an available path for establishing an automated execution rule chain.

Establish Automatic Execution Rule Chain

Establishing automatic execution rule chain provides association rules and transfer dependencies for the connection flow between different applications, thus providing a

basis for automatic combination and mutual call between applications. Different application connection flows are encapsulated into rule classes with various functional attributes, and mutual call and data exchange are realized through interfaces. In this link, RPA can be used to realize the automatic call, chain call and adaptive adjustment of rules, and blockchain can ensure the consistency of data among various rules, providing intelligent attributes for the formation of human-machine interaction session sets.

Form a Session Set for Human-Machine Interaction

Forming a session set for human-machine interaction provides an entry point and an exit point for human-machine synergistic intervention. On the one hand, AI agents or AI assistants can provide students with corresponding decision-making suggestions and learning resources, provide intervention methods and related cases for teachers, and guide effective interaction between students and teachers. On the other hand, while automating the execution of preset rules, AI agents or AI assistants carry out network tuning of rules, cascading calls of applications and visual feedback of data according to the feedback of students and teachers. Constantly adjust the effectiveness and credibility of data points, application flows, rule chains and session sets to ensure the effectiveness of the intervention.

5 The Application of “AI + RPA + X” Intervention Field in College Students’ Learning adaptation

5.1 Design

The study selected 49 undergraduate normal students in the 2021 grade of H University as the research subjects, took the 18-week teaching period of the “Instructional System Design” course in the second semester of the 2022–2023 academic year as the research cycle, adopted the design-based research method, applying the “AI + RPA + X” intervention field in this process. We used the College Student Learning Adaptation in the age of Intelligence (CSLAiI) scale (Xie Y et al. 2023) as a measurement tool, analyzed the effectiveness of the “AI + RPA + X” intervention field by combining the feedback of teaching and learning activities and student satisfaction.

Firstly, relying on the CSLAiI scale, we conducted a pre-test on the learning adaptation of the research object to obtain the result (Y0). Then, the optimization tool based on the “AI + RPA + X” intervention field hierarchy model was used to intelligently transform the existing field. We selected a certain teaching week as an important observation point, got the results(X) from learning satisfaction, etc. Finally, the results (Y) were obtained after the post-test on the research subjects, and based on the pre and post multidimensional measurements, the application effect of the intervention field was determined to be $C = (Y - Y_0) + X$.

5.2 Implement

The study took the key observation class “strengthening first-class curriculum design of ideology and politics” as an example, which aims to require students to understand the

connotation and characteristics of curriculum ideology and first-class courses, and to master the method of designing first-class courses with strengthened curricular ideology and politics. Students applied the “AI + RPA + X” field to participate in learning activities, including “Independent judgment: Please judge whether the content shown in the following pictures reflects the ideology and politics”, “Human-machine synergistic analysis: What are the characteristics of a first-class curriculum?” and “Collaborative group discussion: How to design a first-class course that strengthens the ideology and politics?” etc. The class relied on the National Higher Education Intelligent Education Platform, which connects students’ learning data inside and outside the classroom by carrying out blended teaching, providing solid support for the implementation iterations of intervention and field adjustments.

Automatic Acquisition of Push Information Data

Before class, aiming at the online classroom data in the National Higher Education Intelligent Education Platform, we used the RPA technology to realize the automation of the whole process of learning data collection, processing and pushing, and then help teachers to accurately design teaching activities according to the email content. After class, the classroom feedback results obtained from the pre-designed satisfaction questionnaire were automatically sent to the teacher within 2–3 days after the end of a classroom lesson, and the feedback results are categorized according to different conditions for timely processing and adjustment.

Carry Out Human-machine Synergistic Teaching Based on Mathematical Intelligence

The generative AI tool provided in the classroom activated the human-machine synergistic channel. In the process of interactive dialogue between students and AI assistants, it dynamically generated and recommended valuable viewpoints and content for students, customized precise and intelligent teaching services one-to-one, formed the problem solutions based on subjective dialogues (as shown in Fig. 4) to solve problems arising during the completion of learning tasks. It expanded the category of subjects in the learning field, effectively stimulated the internal drive of students’ participation, guided students to reach the established learning objectives without probation, and facilitated the process of deep interaction and blended learning.

Improvement of Dynamic Tracking, Evaluation and Diagnosis

Based on the iData data collection and circulation platform, the one-click questionnaires on course satisfaction of Instructional System Design were sent to students after class to track the study longitudinally and focus on the process dynamically, so as to understand the students’ learning experience and satisfaction in terms of learning tools, teaching content, teaching activities and learning inputs, etc.. Then we gave full play to the guiding, diagnosing and feedback functions of tracking and evaluation data, to automatically transfer the resultant data to teachers in a timely manner, and classify the feedback results according to different conditions for processing and adjustment, so that teachers can judge students’ learning situation and improve teaching accurately.

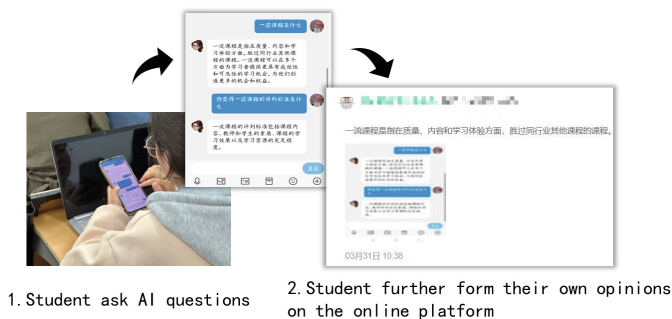


Fig. 4. Learning activities based on AI conversation

6 The Effect of “AI + RPA + X” Intervention Field on College Students’ Learning Adaptation

6.1 The Index of Learning Adaptation

The results of pre and post test of intervention field application show that the total index and each dimension index of college students’ learning adaptation almost increased in different degrees, among which the total index of college students’ learning adaptation increased from 5.0521 to 6.3804, indicating that positive changes have taken place in college students’ learning adaptation as a whole, and the application of “AI + RPA + X” intervention field has achieved certain positive results. However, by further carrying out the paired sample T test, it is found that there are only significant changes ($P < 0.05$) in the dimensions of “learning tool” and “learning community”. The specific results are shown in Table 1.

According to the test results, it can be seen that the total index of college students’ learning adaptation is $t(49) = -0.165$, $p = 0.870 > 0.05$, which shows that even though college students’ learning adaptation has improved ($K = 0.2\%$), it has not changed significantly. The study discussed this separately and concluded that relying on the intervention field of “AI + RPA + X”, teachers adjust teaching activities and integrate AI assistants, which is the main reason for the improvement of “cognitive ability index” ($K = 3.3\%$, $p = 0.135$), “thinking ability index” ($K = 2.0\%$, $p = 0.455$) and “emotional attribution index” ($K = 3.4\%$, $p = 0.181$). At the same time, according to Sewart’s continuous attention theory, the after-class “learning satisfaction” survey provides a channel for students’ feedback, which may be the main factor for the relative increase of emotional attribution index. In addition, most of the existing research subjects are digital aborigines, who have certain advantages in using learning platforms for digital learning, using digital learning tools to adjust their own learning process, and carrying out community learning; however, due to the students’ vague cognition of the overall learning goals, their self-regulation is easy to deviate from the direction of self-regulation and need guidance to better achieve self-adaptive adjustment, this is not only an important reason why the indexes such as “learning platform” ($K = 2.7\%$, $p = 0.392$) and “learning resources” ($K = 2.6\%$, $p = 0.376$) have not changed significantly, but also a key aspect of further optimization of “AI + RPA + X” intervention field. Even in the “AI + RPA +

Table 1. Results of T-test of paired samples for college students' learning adaptation.

Index name	Pre/Post-Intervention Test	N	M	K	SD	df	t	p
Learning Adaptation Total Index	before intervention	49	5.052	0.2%	0.781	48	-0.165	0.870
	after intervention	49	5.064		0.716			
Cognitive Ability Index	before intervention	49	0.148	3.3%	0.032	48	-1.521	0.135
	after intervention	49	0.153		0.029			
Thinking Ability Index	before intervention	49	0.348	2.0%	0.076	48	-0.753	0.455
	after intervention	49	0.355		0.074			
Emotional Attribution Index	before intervention	49	0.322	3.4%	0.068	48	-1.359	0.181
	after intervention	49	0.333		0.054			
Learning Platform Index	before intervention	49	0.148	2.7%	0.026	48	-0.864	0.392
	after intervention	49	0.152		0.020			
Learning Resource Index	before intervention	49	1.059	2.6%	0.200	48	-0.893	0.376
	after intervention	49	1.087		0.140			
Learning Community Index	before intervention	49	0.042	316.7%	0.061	48	-12.852	0.000
	after intervention	49	0.175		0.047			
Learning Tools Index	before intervention	49	0.302	6.0%	0.055	48	-2.390	0.021
	after intervention	49	0.320		0.044			

*Percentage of changes in M before and after intervention.

X” intervention field, there are no significant differences in many dimensions of learning adaptation before and after the intervention, but the development of accurate teaching activities based on AI services has significantly improved the indexes such as “learning tools” ($K = 316.7\%$, $p = 0.0$) and “learning community” ($K = 6.0\%$, $p = 0.021$). In short, from the perspective of learning adaptation index, the intervention field has achieved good results and can promote the interaction among subjects.

6.2 Learning Satisfaction

The students’ satisfaction questionnaire shows that the average score of classroom teaching by the students participating in the study is 4.5 (the full score is 5). Specifically, 43% of students gave the course a score of 4 and 5, while the remaining 14% gave the course a score of 4.5. In addition, students show positive feedback such as “comfortable teaching environment”, “rich resources of course team”, “good paper sharing”, “comprehensive and easy to understand practice”, and the possibility of recommending courses to friends or other students is higher than 3.5. And in the overall experience of the course, most learners agree that “I can speak actively in the online or offline community”, “I think both online and offline learning goals are in line with learning needs”, and “The mixed learning evaluation in the classroom can better reflect my learning situation and urge me to adjust my study flexibly” and so on.

From the above analysis, it can be seen that the “AI + RPA + X” intervention field has varying degrees of effect in learning adaptation index and learning satisfaction, which proves the effectiveness of the intervention field. However, due to the limitations of the indirect port of the platform, security authentication and so on, it is not clear that the independent or combined application of AI, RPA, X and other technologies will play a specific role and effect on college students’ learning adaptation. Therefore, the study has some limitations, and further research is still needed.

7 Conclusion

Learning field plays an important role in supporting, guiding and promoting students’ learning activities. The research focuses on the construction and application of learning adaptation intervention field for college students in the age of intelligence, uses research methods of survey research, literature research, theoretical deduction, and design-based research method, under the guidance of field theory, combined with the development direction of digital transformation, proposed the concept of college students’ learning adaptation intervention field construction of factor connectivity, link routing and data integration; applied technologies such as AI and RPA to build an “AI + RPA + X” intervention field that highlights the characteristics of human-machine synergy, and clarified its hierarchical structure and operating mechanism; carried out a one-semester practical application for 49 undergraduate normal students majoring in educational technology in the 2021 grade of H University, and analyzed the application effect of “AI + RPA + X” intervention field. The results show that the “AI + RPA + X” intervention field of college students’ learning adaptation can provide theoretical and practical reference for constructing and optimizing the existing field and interfering with college students’ learning adaptation.




References

- Gupta, A., Soneji, P., Mangaonkar, N.: Robotic process automation powered admission management system. In: Smys, S., Kamel, K.A., Palanisamy, R. (eds.) *Inventive Computation and Information Technologies. Lecture Notes in Networks and Systems*, vol. 563, pp. 169–178. Springer, Singapore (2023). https://doi.org/10.1007/978-981-19-7402-1_13
- Hu, T.Z., Chen, F.-R.: Research on improving learning adaptability of educational technology students in network environment. In: 2011 International Conference on Electrical and Control Engineering, pp. 6732–6735 (2011). <https://doi.org/10.1109/iceceng.2011.6056948>
- Li, Q.: A brief analysis of Bourdieu's field theory. *J. Yantai Univ. (Philos. Soc. Sci. Ed.)* (02), 146–150 (2002). <https://doi.org/10.13951/j.cnki.issn1002-3194.2002.02.005>
- Mosteanu, N.R.: Machine learning and robotic process automation take higher education one step further [Article]. *Rom. J. Inf. Sci. Technol.* **25**(1), 92–99 (2022). <Go to ISI>://WOS:000775912300007
- Nguyen, A., Gardner, L.A., Sheridan, D.: A framework for applying learning analytics in serious games for people with intellectual disabilities [Article]. *Br. J. Edu. Technol.* **49**(4), 673–689 (2018). <https://doi.org/10.1111/bjjet.12625>
- Pierre Bourdieu, L.J.D.W.: *An Invitation to Reflexive Sociology [Practice and reflection: the guidance of reflective Sociology]*. University Of Chicago Press (1992)
- Qiu, H.: *A probe into the Integrated intervention Model of College students' Web-based Learning adaptability [Master , Hunan University]* (2010)
- Scepanovic, S., Debevc, M.: (2012, 2012 Nov 19-21). adaptation of learning objects in virtual learning environment to learning styles of students. In: [5th International Conference of Education, Research and Innovation (iceri 2012)]. 5th International Conference of Education, Research and Innovation (ICERI), Madrid, Spain (2012)
- Somasundaram, M., Sumitra, V., Vardan, R.S.P.V., Sakthipriya, B., Pavithra, K., Nivedha, M.K.: Monitoring and facilitating students programming skill development using robotic process automation (RPA) and artificial intelligence (AI)-a case study. In: 2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), pp. 1–6 (2022). <https://doi.org/10.1109/icerec256837.2022.10060772>
- Sutipitakwong, S., Jamsri, P.: The effectiveness of RPA in fine-tuning tedious tasks. In: 2020 6th International Conference on Engineering, Applied Sciences and Technology (ICEAST), pp. 1–4 (2020) <https://doi.org/10.1109/iceast50382.2020.9165452>
- Xie, Y., Qiu, Y., Luo, W., Zhang, R.: Research on remodeling of college students' learning adaption in the age of intelligence: connotation, model and measurement. *e-Educ. Res.* **44**(03), 13–20 (2023). <https://doi.org/10.13811/j.cnki.eer.2023.03.002>
- Xu, D., Tsai, S.-B.: A study on the application of interactive english-teaching mode under complex data analysis [article]. *Wirel. Commun. Mob. Comput.* 2675786 (2021). <https://doi.org/10.1155/2021/2675786>
- Zhang, L.F.: Maturity appraisal of RPA application in China. *AEBMR-advances in economics business and management research [Proceedings of the third international conference on economic and business management (feb2018)]*. In: 3rd International Conference on Economic and Business Management (FEBM), Hohhot, Peoples R China (2018)
- Zhao, T., Zhang, W., Zhao, H., Jin, Z.: A reinforcement learning-based framework for the generation and evolution of adaptation rules. In: [2017 IEEE International Conference on Automatic Computing (icac)] 14th IEEE International Conference on Autonomic Computing (ICAC), Ohio State Univ, Columbus, OH (2017, 2017 July 17–21)

Smart Learning Environment



Utilization of Digital Human Model in Design Education: A Design Students' Learning Survey and Case Study

Hung-Hsiang Wang¹  and Chih-Ping Chen²  

¹ Department of Industrial Design, National Taipei University of Technology, Taipei, Taiwan

² Department of Product Design, Ming Chuan University, Taoyuan, Taiwan

ms1101841@ms1.mcu.edu.tw

Abstract. This investigation explores the utilization of the digital human model (DHM) in design education, including a survey and a case study of design students' learning results. The survey comprises an anthropometric investigation of 106 students of a university and an investigation of design students who use DHM. The study examines the design students apply DHM and product design to design scenes. The survey results show that the students' design works can be divided into three types: (A) adjusting 3D human models (61.3%), (B) no adjusting 3D human models (23.6%), and (C) no importing 3D models (i.e., only cut-paste human shape images; 15.1%). Most of the students are able to use the low design threshold approach to generate good design works. In addition, the case study focuses on the students' design works of door handle devices in public places during the COVID-19 epidemic. The results show that the simulated sensory feedback the virtual environment gives to the students lacks sufficient accuracy. Although designers could simulate the user's operation in the early design stage, however, applying DHMs does not ensure the generation of optimal designs. In product design experiments, physical prototypes are needed to simulate the user's working force, to which the optimal angles and shapes can be determined.

Keywords: Digital human model · Design education · Product design · Human-machine interactive · Anthropometry · Door handle auxiliary device · COVID-19

1 Introduction

Due to the rise of the wave of the metaverse, the virtual digital human model (DHM) has become indispensable. The development of digital human models (DHMs) in the human factors engineering field has a long history, mainly in industrial engineering, simulation, and the application of assembly efficiency on the production line. Recently, the application of DHMs is not limited to the field above and plays a vital role in education, industry, entertainment, and design. New applications of DHMs have become an important academic research field [1]. However, how DHMs can assist design education is worth considering. This article studies the case of door handle auxiliary devices under

the epidemic as the theme. It will combine digital human body models, anthropometric measurements, and product design to discuss the use of combined DHMs and design education to enter the virtual world to assist in the design process and ascertain the feasibility of DHMs to assist design education.

In addition, there are obvious differences between Taiwan's domestic industries and foreign industries. Taiwan has a few heavy industries but focuses on small and medium-sized enterprises, including consumer electronics, hand tools, components, medical equipment, etc. However, foreign countries apply DHM to the automobile industry, factories, production line planning, medical devices or workstations, etc. DHM is rarely used in combination with product design in Taiwan. Many enterprises are suitable for applying DHM in product design, such as furniture design, shoe design, input devices design, etc., which are worthy of in-depth discussion. Also is worth introducing DHMs in human factors engineering courses for students to learn the application of this new technology in product design.

Therefore, this article considers the importance of users in the human factor design with daily life. In addition, it combines DHMs and human-centered design. So humans are no longer invisible characters in product design drawings, but play a critical role in having a close interaction with the product. On the other hand, applying DHMs can simulate user operation behaviors, enhancing students' creativity in product design, and solve complex real-life design problems. This paper introduces the course of DHMs in the human factors engineering course at the university. We aim to attract more students' interest in combining DHMs and product design through DHMs studying courses.

2 Literature Review

The article's literature review focuses on the significance of DHMs in product design, the impact of CAD on design education, and the development of door-handle auxiliary devices to address public health concerns during the COVID-19 pandemic.

2.1 Digital Human Model

Advances in digital human models (DHMs) and computer-aided design have revolutionized the field of product design, with applications in clothing design, furniture design, and automotive design. Design prototypes can be simulated and virtually tested using 3D analysis software, which assists in understanding the interaction between the designed product and relevant body regions [2]. In addition, digital human modeling of design has been widely used in biomechanics, education, and healthcare [3]. Also, University of Michigan researchers need to improve existing DHM to better serve as effective ergonomic analysis and design tools and effective posture and motion prediction models for various human populations [4].

On the other hand, virtual reality has been widely used in various manufacturing industries, and the DHM is crucial for virtual manufacturing applications. The DHM solution is based on morphological, behavioral, and product modeling. Also, based on multiple knowledge engineering operations, the available knowledge is transformed into

a higher level [5]. Simply put, DHMs are increasingly used in the early stages of product design.

The DHM is considered the best method to capture all the detailed movements, correct wrong usage, and to facilitate improvement in the process. The goal is to simulate the operation use of real-life people. Relevant cases include a method for analyzing the ergonomics of equipment specifically designed to house refrigerated display units, the study describes the method, including selected virtual persons, refrigerated units, and products handled [6]. A research lab at the University of Oregon proposed an ergonomic approach for incorporating Human Factors Engineering (HFE) guidelines early in the design process. The method utilized virtual reality (VR), computer-aided design (CAD), and participants in the conceptualization phase before building a functional prototype. It was used to evaluate human performance during a Boeing 767 cockpit fire emergency [7]. In addition, related applications and research of DHM include electric bicycles [8], traffic lights [9], threshing machines [10], car driving seats [11], and wheelchairs [12].

2.2 CAD's Impact on Design Education

Research by Brink showed that educators have various encounters with the expected learning results when the computer-aided design is utilized in instructing, and four progressive classes arose: (a) Dealing with the product, (b) Utilizing instant models, (c) Manufacturing and making printed models, and (d) Planning. School teachers and educators additionally need to offer the most significant benefit for understudies in their computer-aided design course [13]. Ye's research mentions the question about CAD courses: what should be included in such a computer-aided design educational curriculum [14]?

Different researchers have different opinions on whether CAD can stimulate students' creativity. Altay mentions that the virtual techniques suggest that CAD has practically zero worth as a stimulus for ideas [15]. However, Mustaal's research mentions that from the analysis and discoveries, the study recommends a connection between the emergence of imaginative ways of behaving and the use of CAD in designing [16]. Thus, Gelmez's research mentions a model associating the knowledge types with learning results as far as their intricacies were developed for CAD courses [17]. This means that the design of CAD courses needs to consider different factors to achieve the best results for students and creativity stimulation.

2.3 Door Handle Assist Design During the Epidemic

The COVID-19 virus may remain on various surfaces for up to 72 h. Thus, avoiding direct contact with surfaces is a critical way to reduce contamination during a pandemic, especially in public and healthcare centers [18]. Some components, such as buttons, door handles, switches, etc., can be viral spreaders because they interact with many people, especially in public places, such as hospitals, buildings of communities, and transportation stations. Therefore, a series of handle modifications have been developed to allow alternative mechanisms for door opening without direct skin-to-surface contact [19].

These doors handle auxiliary devices can be pressed by elbows, hooked by fingers, stepped on by feet, or presented as additional designs. Relevant literature includes Taipei Veterans General Hospital, which proposed a new device to enhance existing door handles, to avoid direct contact with hands and reduce the transmission of microorganisms through healthcare workers [20]. In addition, researchers report their experiences of designing, producing, and dispatching three 3D printed components designed to reduce the risk of COVID-19 contamination by limiting direct contact: (a) fixed hand-free door openers, (b) door hooks and (c) button pushers [21].

There are also additional door handle-mounted designs, with the most commonly used door handles revised to develop proposals that can be applied to existing models. The goal is to avoid the need for their complete replacement. Through the interplay between engineering, design, and ergonomics, an assisted hands-free door-opening device is developed to provide greater ergonomic comfort [22]. On the other hand, 3D printing also plays an important role. For example, one hand-free door handles auxiliary device is a gadget that allows the user to open the door by elbow or foot. Also can be applied 3D printing to reduce direct contact with the door handle, contact and avoid potential contamination points [23].

3 Methods

This article outlines a research framework consisting of three main components: survey and analysis, case study, and discussion. It focuses on the application of Digital Human Models (DHMs) in product design, with a particular emphasis on door handle auxiliary device design in response to the COVID-19 pandemic.

3.1 Research Framework

The framework of this study consists of three parts: survey and analysis, case study, and discussion. In the survey phase, there are two investigations. The first one is an anthropometric survey of university students and then create 3D of DHMs from the survey data. The second one is followed by a survey of design students learning the DHM to design scenes. The research framework is presented in Fig. 1.

In addition, the case study applies the DHM to aid door handle auxiliary device design. Many people attract attention to the door handle auxiliary device design, due to the Covid-19 pneumonia epidemic that has impacted the world in recent years. People have begun to emphasize the need to maintain a safe distance from each other and reduce unnecessary contact. Therefore, door handles that we often encounter in our lives have become one of the problems that need to be urgently solved. Many products have appeared on the market. The handle device allows people to open the door without touching the handle. However, the shape currently on the market is too limited in its application, and considering products with human factors is required. Thus, we plan to design a model that is more in line with the design of the ergonomic door handle auxiliary device through the DHM, design experiments, and analysis of human factors.

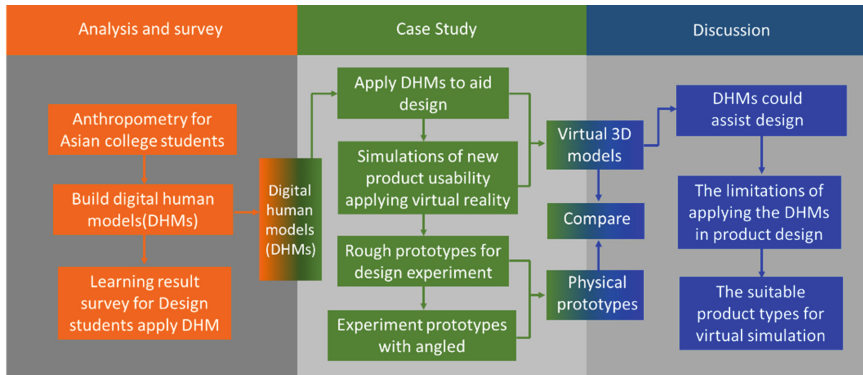


Fig. 1. The Research Framework.

The last part of the research framework is the discussion. Applying the virtual 3D DHMs method and physical models of product design, the study will compare the differences between the two. Also, we discussed the advantages and limitations of the virtual model application of the DHM.

3.2 Create a Process Combining DHM and Product Design

The 3D software for creating DHMs is MakeHuman, and the software for integrating product design and DHMs is Blender. Afterward, 3D rendering software is used for rendering, and preparations for implementing VR/AR are provided. The detailed creation procedure is as follows.

1. Apply MakeHuman software to create a DHM. Next, adjust age, gender, body shape, skin color, and other characteristics. Then set the export 3D file format, MHX2 or MPFB.
2. Use Blender to receive the 3D file format of the MakeHuman or apply the MPFB file format to input. In Blender, you can adjust the posture of the received DHM and integrate it with product 3D models by other designed software.
3. Apply product design software to design products 3D, such as Rhino or Creo, and set the output 3D file format.
4. Use Blender to integrate product design 3D models and DHMs. And then complete scene rendering. In addition, it can also use other rendering software, such as Keyshot.
5. Convert the 3D modeling combining DHMs and product design to virtual reality (VR). Software such as Simlab Composer or Unity allows the participants to simulate the usage situation of the product.

4 Survey and Analysis

The study implements two surveys regarding DHMs. The first one conducts a survey of anthropometry for college students and built DHMs. The second is the survey of learning results about how the design students apply and learn the DHMs method.

4.1 Anthropometry for College Students and Build DHMs

First, the investigation introduce how to combine anthropometry with a virtual DHM and apply it in the field of product design. Anthropometry is the measurement of human physical attributes. In human factors, anthropometry is used as the basis for setting up a workstation. Three ergonomic design principles based on anthropometry are range design, extreme design, and mean design [24]. Anthropometry and anatomical models will be combined to measure the lengths of hands, arms, and legs for men and women, and performance modeling for better design of products and workstations [2].

This study covers the human-machine interactive, combined with the DHM, and the steps to analyze the user's operation. The investigation considers the human-centered method and uses anthropometry to measure the height of the user. The participants of the experiment are second-grade and third-grade students of Ming Chuan University. The study combines a DHM with product design and tries to apply it to teaching human factors engineering courses in the School of Design.

Participants are eligible for anthropometric data regarding gender, age, skill level, and experience. A total of 63 s-year and third-year students of Ming Chuan University participated in the questionnaire. The participants are 12 males and 51 females, and their ages are between 20 and 22 years old. All participants are students of the Product Design Department of the School of Design. The online questionnaire tool used is Surveycake, and the content of the questionnaire included the sex, hand span, and height of the participants. The questionnaire results found that the average height for females is 160.725 cm and a standard deviation of 5.095 cm. At the same time, the average height for males is 172.25 cm and a standard deviation of 5.048 cm. The data visualization software for statistical values is JMP 14 from SAS. Also, transfer the data into 3D of the DHM and adjust the height, please see Fig. 2.

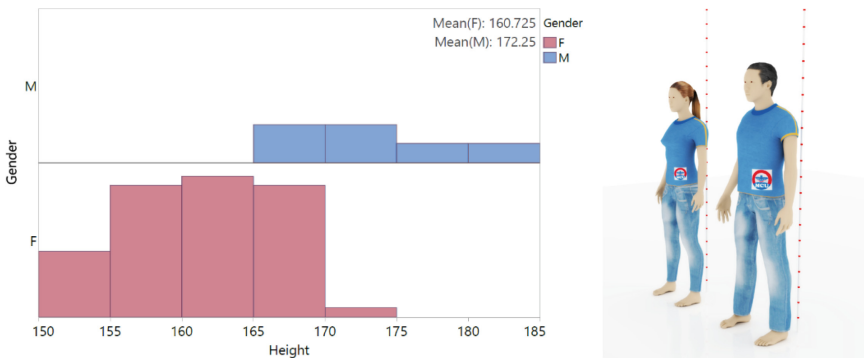


Fig. 2. The participants are 63 students in the second grade and third grade of Ming Chuan University to measure the height. The study combines DHMs and anthropometric measurements to create 3D models with accurate height in software.

4.2 The Survey of Design Students Apply and Learn the DHMs

Next, this study discovers the learning results and performance of DHM applications by design students. Thus, 106 students in Classes A and B are analyzed, and the learning results are divided into three types: better learning results (Type A), normal learning results (Type B), and weak learning results (Type C). Type A students can import 3D models and adjust the gestures or postures of DHMs according to products or furniture size. Type B students can import 3D models but only use existing gestures or postures of DHMs and cannot adjust 3D models. Type C students cannot import 3D DHMs into the 3D scenes. Also, they can only present gestures and pictures of humans by cut-paste pictures. Thus, we can see some scale or symmetry issues in the image.

The learning survey result shows that there are 65 students with type A: better learning result (61.3%), 25 students with type B: normal learning result (23.6%), and 16 students with type C: weak learning result (15.1%). The results are depicted in Fig. 3.



Fig. 3. The learning survey of design students apply a DHM and divided the results into three types include (a) type A: better learning result, (b) type B: normal learning result, and (c) type C: weak learning result.

Therefore, 61.3% can import the 3D of DHM and adjust the gestures and postures of the DHM according to the product or furniture. However, there are still 15.1% of students cannot import the 3D of DHMs and can only present them by cut-paste pictures. In general, students can use this method to be fully creative, low threshold to design, and get good design works.

5 Case Study

The study compares the two methods, applying door handle auxiliary device design as a case using the virtual 3D DHMs and physical product design models method. The article highlights the integration of DHMs and VR in product design, emphasizing the importance of physical prototypes for practical testing and user comfort evaluations. It demonstrates the significance of proper angles in the design of door handle auxiliary devices.

5.1 Apply DHM to Aid Door Handle Auxiliary Device Design

The study combined DHMs and anthropometric measurements to create precise 3D models in software. Thus, we have a relatively accurate height of DHM, which aligns

with the average height of second and third year university students in Taiwan. We also can apply the DHM to design products, including consumer electronics, hand tools, components, medical equipment, etc. On the other hand, we can also apply this model to the human factor design of daily life, including desks and chairs, office furniture, sinks, drinking fountains, and even door handles. Next, this case will combine DHM to study the door handle auxiliary devices on university campuses during the epidemic. The case study is depicted in Fig. 4.



Fig. 4. A case study using a DHM to assist a door handle auxiliary device design.

5.2 Simulations of New Product Usability Applying Virtual Reality

Today's technology could quickly move from 3d modeling scenarios that combine DHMs and product designs to virtual reality (VR). Also, allow experiment participants to simulate the usability of conceptual products. Virtual reality can be roughly divided into two categories: static and interactive. Static VR is limited to exploring the scene, while interactive VR allows users to interact with objects. Some software tools, such as Simlab Composer, could transform 3D modeling into an interactive VR experience in the product design process. Allowing designers to design products in a virtual reality environment is conducive to the integrity of the product. Through the immersive virtual experience, the designers can also obtain critical feedback from participants on the developed products. In this way, virtual reality (VR) is converted from an entertainment tool to an industrial or educational use.

5.3 Apply Rough Prototypes for Design Experiment

In the early stages of the design process, a total of four different types are designed, including (A) an angled door handle auxiliary device, (B) a horizontal door handle auxiliary device, (C) a "U" shaped door handle auxiliary device and (D) an "O" shaped door handle auxiliary device. These four types of designs are all door handle auxiliary devices. You don't need to change the original door handle design, and it doesn't need to be opened by hand. Pressing the device by elbow can avoid direct contact during the epidemic. The four types of rough prototypes are depicted in Fig. 5.

At this time, it is possible to integrate the DHM, and the 3D model of the door handle auxiliary device into the virtual world scene. However, the pressing weight of the

elbow when pressing the door handle, the friction of the material during contact, and the tactile feeling of pressing cannot be simulated. Thus, the user must use a physical model to test and evaluate its effectiveness. However, the DHM can still be used for visual evaluation to find design problems, such as with the “U” and “O” shaped door handle auxiliary devices, chiefly how to configure them (place the device vertically or parallel) to consider which concept is more suitable for use. The goal is to prevent designers from imagining designs out of thin air and choosing a suitable concept design.



Fig. 5. Rough prototypes of the door handle auxiliary device for the design experiment.

5.4 Apply Experiment Prototypes With Angled

Next, this experiment screened the design concepts of the above-mentioned several door handle auxiliary devices and carried out a subjective evaluation of the comfort of these devices. The goal is to understand users' comfort with different inclination angles when using the devices. The difference is used as a reference for the designer to design the door handle auxiliary device. According to our continuous design and modification of the handle auxiliary device, two auxiliary devices are finally designed: 15-degree and 25-degree. The prototypes of two auxiliary devices are depicted in Fig. 6.

Finally, there are nine participants in this design experiment. All of them are right-handed, aged between 18 and 56, with a height between 155 cm and 182 cm. The test scores received after the actual test. Almost everyone prefers a 15-degree door handle auxiliary device to a 25-degree regardless of age and height.

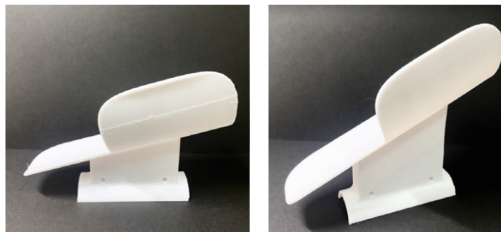


Fig. 6. Two experiment prototypes of different angles: 15-degree and 25-degree.

According to the statistical t-test result of the two devices, the p-value is 0.0002. Due to the p-value being less than 0.05, the null hypothesis is rejected, and the 15-degree

handle device has a statistically significant preference for the experiment participants. Finally, the interviewees found that the 15-degree device is generally more comfortable and smoother. Therefore, according to the subjective evaluation and statistical results after the experiment, it is found that the proper angle between the door handle auxiliary device and the arm had a significant influence. The door handle auxiliary device is better at 15-degree than 25-degree. In general, the same device with two different angles, the 15-degree device, is found to be more in line with the user's needs.

6 Discussion

The applications of DHMs and VR assist in streamlining product design processes, especially in the early design stages. Still, they cannot entirely replace physical prototypes, particularly when assessing a product's tactile and physical aspects. The choice between virtual and physical models depends on the specific design stage and product.

6.1 DHM and Virtual Reality Could Assist in Product Design

After implementing the case study, the main question is how DHMs can assist product design. Whether it is in the early or later stages of the design process. However, in the early stage of design, because designers are exploring problems, the possible answers are relatively divergent. Thus, whether it is a circular closed "O" shaped or a semi-closed "U" shaped door handle auxiliary device may become the final design direction. The DHM can assist in product design virtual simulation and discover problems in different concept designs.

Students of design-related departments have strong practical ability, including 3D modeling, 3D human models construction, scene design, light control, etc. It is not difficult to convert the models of the virtual world into a physical model, and we can see the advantages of design department students to design the objects in the virtual world (metaverse). However, programming is not an advantage for design students. Therefore, if they can complete the details of the virtual world without programming, it could narrow the gap between the virtual world and the real world.

6.2 Compare Virtual 3D Models and Physical Prototypes

Regarding which design stage the DHM is suitable for, in this case, it is more suitable in the early design stage. For example, what is the optimal door handle auxiliary device, whether parallel to the door or perpendicular to it? What kind of suitable situation can simulate the user's hand to avoid hitting the door, instead of imagining it out of thin air? However, in the latter stage of the design, the physical prototype is more valuable than virtual 3D models. For example, the designer determines the different angles of the door handle auxiliary device, which one is the optimal design, 15-degree or 25-degree? Due to some conditions cannot be simulated by digital human body models, such as the weight of pressing, the feel of use, and the difference in materials. In addition to the virtual model, the physical model also plays an important role.

The product design of visualization can be simulated and evaluated, but the feel of the hand cannot be simulated with a virtual model. For example, an Anne CPR training manikin is often used in the education and training of CPR. However, a virtual Annie only can be used to practice the procedure. The students using AR\VR to simulate the feel of Annie's is more challenging. Therefore, the participants who operate on the virtual Anne still need a physical Anne to realize the training practice.

6.3 The limitations of Applying the DHMs and VR in Product Design

Although digital human models (DHMs) and virtual reality (VR) provide many advantages for product design, there are also certain limitations to be aware of. The advantages include reducing product design costs, a speedy product design schedule, and fewer team members to complete the project. The limitations include a lack of accurately simulated sensory feedback and physical feedback. Due to the primary reliance on visual and auditory feedback in DHM and VR simulations, it can be difficult to effectively sense and evaluate physical characteristics including force, texture, weight, and resistance.

If we take the door handle auxiliary device as an example, designers can simulate the user's operation based on the VR environment in the early design stage. However, depending solely on DHMs and virtual reality cannot predict the optimal design. The experimentation of human factors is needed with a physical model to simulate the operation force and determine the optimal design with a suitable angle.

6.4 The Suitable Product Types for Virtual Simulation

Due to different kinds of human factor experiments, there are various product types. In addition to the door handle auxiliary devices design, available products also include mug designs, metro ticket vending machines, glasses, one-handed bottle openers, stylus pens, keyboard cleaning brushes, elevators button layout designs, hair dryers, weighted computer mice, bus handles, tableware for older people, soft material seasoning jars, and juicers. Another question is which product type of virtual 3D model by the DHM is suitable for simulation. After observing the actual design experiment process, the investigation found that large products are more suitable than small products, product usability in human behavior evaluation is more suitable than the operation of hand tools, and visual design is more suitable than tactile design. In addition, the design in the early stage is more suitable than the later stage.

Therefore, this survey takes the design of the door handle auxiliary device as an example. If it is necessary to select the optimal design from two auxiliary devices from different angles in the later stage of the design, the actual test physical model will be a better choice. However, if it is the design concept that needs to be determined in the early stage of design, then virtual methods (AR\VR) can be used for simulation. But this does not mean DHMs can replace rough models, 3D printing, or actual physical prototypes. It can only save the initial production model's cost and achieve the effect of user simulation in the early-stage product design process.

7 Conclusions

This investigation provides a design students' learning result survey and a case study. The case combined DHMs and design education in product design courses to discover how the method can assist the design process. From a design viewpoint, physical prototypes still have irreplaceable value in product design. The virtual DHMs cannot completely replace the user's experience in use. Nevertheless, the operation and simulation of the DHM can help the designer or design students choose the appropriate design direction in the initial design stage. This approach is beneficial for discovering design issues through visualization. A future direction in research will explore how to use DHMs to integrate VR glasses to simulate product operations in design education.

Acknowledgment. The authors are grateful to the students of the Product Design Department of Ming Chuan University.

Funding. This research is funded by MOE, Taiwan, R.O.C., through grant PHA1110214.

References

1. Zhu, W., Fan, X., Zhang, Y.: Applications and research trends of digital human models in the manufacturing industry. *Virt. Reality Intell. Hardware* **1**(6), 558–579 (2019). <https://doi.org/10.1016/j.vrih.2019.09.005>
2. Shah, P., Luximon, Y.: Use of digital human modeling in product design. In *DHM and Posturography*, pp. 591–598. Academic Press (2019). <https://doi.org/10.1016/B978-0-12-816713-7.00044-1>
3. Duffy, V.G.: Human digital modeling in design. *Handb. Hum. Factors Ergon.* 1016–1030 (2012). <https://doi.org/10.1002/9781118131350.ch35>
4. Haffin, D.: Improving digital human modeling for proactive ergonomics in design [Invited plenary paper]. *Ergonomics* **48**, 478–491 (2005). <https://doi.org/10.1080/00140130400029191>
5. Moes, N.C.C.M.: Digital human body modelling to support designing products for physical interaction. In: *DS 36: Proceedings DESIGN 2006, the 9th International Design Conference*, Dubrovnik, Croatia, pp. 757–766 (2006)
6. Colombo, G., Regazzoni, D., Rizzi, C.: Ergonomic design through virtual humans. *Comput.-Aid. Des. Appl.* **10**, 745–755 (2013). <https://doi.org/10.3722/cadaps.2013.745-755>
7. Ahmed, S., Irshad, L., Demirel, H.O., Tumer, I.Y.: A comparison between virtual reality and digital human modeling for proactive ergonomic design. In: Duffy, V.G. (ed.) *HCI 2019*. LNCS, vol. 11581, pp. 3–21. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-22216-1_1
8. Wang, X., Peng, J.: A simplified human body model for assisting electric bicycle design. In: *Digital Human Modeling and Applied Optimization* (2023)
9. Sudarma, M.: Visibility analysis model to traffic light using digital human modeling. *Ergon. Int. J.* (2022)
10. Marshall, R., Brolin, E., Summerskill, S., Högberg, D.: Digital human modelling: inclusive design and the ageing population. In: Scataglini, S., Imbesi, S., Marques, G. (eds.) *Internet of Things for Human-Centered Design*. Studies in Computational Intelligence, vol. 1011, pp. 73–96. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-8488-3_5

11. Schlund, S., Kamusella, C., Knott, V., et al.: Digital ergonomics and digital work planning in university education: experiences from Germany and Austria. *Z. Arb. Wiss.* **76**, 510–524 (2022). <https://doi.org/10.1007/s41449-022-00333-7>
12. Al Meanazel, O.T., Obeidat, M.H., Obeidat, S., Al-Momani, H.A., Hamasha, M.M.: Wrist deviation influence on initial hand force before wheelchair propulsion using digital human modeling. In: Proceedings of the 5th European International Conference on Industrial Engineering and Operations Management, Rome, Italy, July 26–28 (2022)
13. Brink, H., Kilbrink, N., Gericke, N.: Teach to use CAD or through using CAD: an interview study with technology teachers. *Int. J. Technol. Des. Educ.* **33**, 957–979 (2023). <https://doi.org/10.1007/s10798-022-09770-1>
14. Ye, X., Lee, W., Kang, S., Shin, S.Y.: Today's students, tomorrow's engineers: an industrial perspective on CAD education. *Comput. Aided Des.* **36**(14), 1451–1460 (2004)
15. Altay, B.: Multisensory inclusive design education: a 3D experience. *Des. J.* **20**(6), 821–846 (2017)
16. Musta'amal, A.H., Norman, E., Jabor, M.K., Buntat, Y.: Does CAD really encourage creative behaviors among its users: a case study. *Procedia – Soc. Behav. Sci.* **56**, 602–608 (2012)
17. Gelmez, K., Arkan, S.: Aligning a CAD course constructively: telling-to-peer and writing-to-peer activities for efficient use of CAD in design curricula. *Int. J. Technol. Des. Educ.* **32**(3), 1813–1835 (2022)
18. Longhitano, G.A., Nunes, G.B., Candido, G., da Silva, J.V.L.: The role of 3D printing during COVID-19 pandemic: a review. *Progress Additive Manufact.* **6**(1), 19–37 (2021). <https://doi.org/10.1007/s40964-020-00159-x>
19. Tino, R., Moore, R., Antoline, S., Ravi, P., Wake, N., Ionita, C.N., et al.: COVID-19 and the role of 3D printing in medicine. *3D Print Med.* **6**(1), 11 (2020). <https://doi.org/10.1186/s41205-020-00064-7>
20. Chen, K.L., Wang, S.J., Chuang, C., Huang, L.Y., Chiu, F.Y., Wang, F.D., et al.: Novel design for door handle—a potential technology to reduce hand contamination in the covid-19 pandemic. *Am. J. Med.* **133**(11), 1245–1246 (2020). <https://doi.org/10.1016/j.amjmed.2020.05.015>
21. François, P.M., Bonnet, X., Kosior, J., Adam, J., Khonsari, R.H.: 3D-printed contact-free devices designed and dispatched against the COVID-19 pandemic: the 3D COVID initiative. *J. Stomatol. Oral Maxillofac. Surg.* **122**(4), 381–385 (2021). <https://doi.org/10.1016/j.jormas.2020.06.010>
22. Maranhã, V., et al.: Development of a dynamic hands-free door opener to prevent COVID-19 pandemic spreading. *Designs* **5**(3), 56 (2021)
23. Radfar, P., Bazaz, S.R., Mirakhorli, F., Warkiani, M.E.: The role of 3D printing in the fight against COVID-19 outbreak. *J. 3D Print. Med.* **5**(1), 51–60 (2021)
24. Stack, T., Ostrom, L.T., Wilhelmsen, C.A.: Anthropometry. *Occup. Ergon.* 21–76 (2016). <https://doi.org/10.1002/9781118814239.ch3>



A Software for Learning Thai Fingerspelling Sign Language into Thai Consonants with Machine Learning

Supisara Pisuchpen¹, Chansiri Singhtaun²(✉), and Laddawan Suwannachote¹

¹ Kasetsart University Laboratory School Center for Educational Research and Development, Faculty of Education, Kasetsart University, Bangkok, Thailand

² Department of Industrial Engineering, Faculty of Engineering, Kasetsart University, 50 Ngamwongwan Road, Lat Yao, Chatuchak, Bangkok 10900, Thailand
chansiri.s@ku.ac.th

Abstract. Children with hearing impairment or deafness often face difficulties in spelling and reading. This research aims to develop a program that translates Thai fingerspelling sign language into Thai consonants. The objective is to use this program as a teaching tool to help children with hearing impairment understand and learn the Thai language more easily. The program created is a utilization of Artificial Intelligence (AI) techniques, where the program captures real-time hand gesture images in fingerspelling sign language and then translates the results into Thai consonants. It enables them to accurately create hand signs and makes the learning process more engaging and interesting. The program development begins with Hand Detection algorithm, which involves detecting the positions of 21 landmarks on the hand after using image processing techniques. After that, the hand gesture images for all the Thai consonants are captured and saved to be used as ground truth datasets. Next, the ground truth datasets are fed into the machine learning process to create a trained model. This enables the program to match the hand gesture images accurately with the correct consonant flashcards, using Google's Teachable Machine. The results of testing the program revealed that the program has an average accuracy of 79.50%. The program was deployed with 12 users, consisting of 2 teachers and 10 elementary school students from Satsatian School for the Deaf under the Royal Patronage of His Royal Highness Crown Prince Maha Vajiralongkorn. After conducting the user satisfaction survey and applying statistical analysis methods, the results revealed an overall satisfaction rating of "very satisfactory," scoring 83.80% across all aspects. The program developed can be used as a teaching aid for children with hearing impairments to facilitate their learning and academic progress.

Keywords: artificial intelligence · hand detection · machine learning · student engagement · student satisfaction

1 Introduction

Sign language is an essential means of communication for hearing-impaired people. It is a fundamental and effective way of sharing thoughts, feelings, and opinions. However, sign language is not universal, with the diversity of over 7000 sign languages with variability in motion position, hand shape, and position of body parts (Adeyanju et al. 2021). Each country and even regions within countries may have their own sign language with unique vocabulary and grammar. For example, American Sign Language (ASL) is used in the United States, British Sign Language (BSL) is used in the United Kingdom, Auslan is used in Australia, Thai Sign Language (TSL) is used in Thailand, and so on. Any sign language usually has two schemes: sign and fingerspelling schemes (Nakjai and Katanyukul 2019). Signs in sign languages are composed of specific handshapes, movements, locations, and orientations that convey words, phrases, and concepts. Signs are the core components of sign languages and are used for everyday communication. Fingerspelling involves manually spelling out the letters of a word using specific handshapes or finger positions that correspond to individual letters of the written or spoken language. This is especially useful for spelling out words that might not have a corresponding sign or for conveying proper nouns, names of people, places, technical terms, and unfamiliar words that may not have established signs in the language. Additionally, the use of fingerspelling can also aid in enhancing vocabulary and literacy development for individuals with hearing impairments (Alawad and Musyoka 2018), thereby increasing opportunities for inclusive learning alongside their hearing peers.

The academic success of deaf students relies on their ability to read and comprehend the meaning of a text. Thus, the development of literacy skills is considered a critical factor in deaf students' academic success. Difficulty in reading and writing among deaf students adversely affects their learning processes, indicating the importance of literacy skill development for deaf students (William 2012). Sarchet et al. (2014) investigated reading ability and specific language skills of deaf students in third through seventh grade. The results revealed that knowledge of specific ASL structures, including fingerspelling, correlates with reading achievement. In their study, children who scored better on reading tests were competent in associative skills, such as the ability to write down words that were fingerspelled to them as well as the ability to translate initialized signs. Looking specifically at performance on the fingerspelling tasks, fingerspelling ability significantly correlated with reading comprehension.

Setsatian School for the Deaf under the Royal Patronage of His Royal Highness Crown Prince Maha Vajiralongkorn is the first school for the deaf in Thailand. It is a government school under the Department of Empowerment of Persons with Disabilities, Ministry of Education. Currently, there is teaching and learning of fingerspelling integrated with sign language at the elementary school level, which is for children aged between 6 and 10 years old. The current teaching method for fingerspelling involves the teacher demonstrating hand movements for students to observe, and then presenting images of corresponding letters for students to look at. This process is repeated over and over until the students memorize it and can establish the correct mapping between fingerspelling postures and printed letters. Since students need to practice and perform fingerspelling sign language gestures frequently, teachers are required to assess the gestures until they are accurate. This results in limited practice time for each individual

student and consequently hinders the effectiveness of learning. The school aims to promote students' ability to learn and practice accurate fingerspelling, laying a foundation for spelling, reading, and writing at higher grade levels. The purpose of this research is to introduce real-time AI-driven software tailored to engage learners, for use as instructional and training materials in Fingerspelling. The goal is to facilitate accurate learning and enjoyable memorization, aiming to enhance students' interest and willingness to learn. By hypothesizing that integrating software could significantly boost learning enthusiasm, the research aspires to amplify students' engagement in the learning process. The remainder of the paper is organized as follows; Sect. 2 provides the review of related studies. Section 3 presents the methodology to develop the program and program usability assessment. Section 4 the results of the program, program accuracy and user satisfaction scores are presented. The conclusions are presented in Sect. 5.

2 Related Studies

British, American, Russian, Arabic, French, Chinese, and Japanese sign languages rely mainly on still postures (either in a one-hand or two-hand scheme) with only a few exceptions for movement signing. On the contrary, TFSL uses single-hand fingerspelling with an extension using movement and multi-posture signings for the consonants and a two-hand scheme for the vowels and intonation marks. Thai Fingerspelling sign language (TFSL) was developed in 1953 by Lady Kamala Kraireuk, who adapted the ASL fingerspelling system to suit the Thai language's phonological and linguistic characteristics. It is developed to represent 42 consonants, 4 intonation marks and 20 vowels. Noted that there are 44 consonants in Thai language. Two consonants do not have the TFSL because they are not commonly used in everyday life. This research focuses on 42 Thai consonants because it is a significant basic lesson for the children. The TFSL for 42 Thai consonants can be categorized into 3 groups based on the number of postures or the number of hand movements: single-posture, two-posture, and three-posture. The diagram of the postures representing the consonants are shown in Fig. 1.

A one-posture consonant means the consonant can be represented using a single posture or movement. A two-posture consonant and a three-posture consonant mean that these consonants require two and three postures, respectively, to represent them. In Thai consonants, there are 15 one-posture consonants, 24 two-posture consonants, and 3 three-posture consonants. These consonants are produced through single use or by combining 25 different postures (Nakjai and Katanyukul 2019).

Learning and understanding sign language is a complex matter due to the variety of gestures and communication styles, making communication between regular individuals and those with hearing impairments challenging. Hence, there are numerous research works that attempt to develop technologies to facilitate easier communication between regular individuals and those with hearing impairments. Recent advancements in AI have created opportunities for research groups to incorporate AI into sign language recognitions (SLR). There are various ways to apply AI in sign interpreting operations. More recently, further attention has been directed toward intelligent-based SLR systems, as they are now being applied in numerous applications. These applications include

1- posture consonant	2- posture consonant	3- posture consonant
ก:	ข: ค: ฅ:	ง:
พ: ฝ:	ด: ต: ฒ:	ช:
บ: ป: ผ:	ท: ฑ: ฒ:	ฉ:
ด: ฟ: ล:	ษ: ฮ: ป:	
ย: ม: น:	ผ: ฝ: ฮ:	
อ:	ฎ: ฏ: พ:	
	จ: ญ: ฉ:	
	ง: ฑ: ฉ:	

Fig. 1. TFSL for consonants diagram (adapted from Tumsri and Kimpan 2017).

robotics, interpreting services, real-time multi-person recognition systems, games, virtual reality environments, natural language communications, online hand tracking for human communication in desktop environments, and human-computer interaction.

Vision-based SLR techniques here later (VSLR), which is an application of intelligent based SLR systems, are used to develop programs to translate sign language to letters or words. There are research works that apply VSLR techniques to various countries' sign languages especially ASL. The application of these techniques to Thai sign language is more limited. And, using VSLR techniques for TFSL is more complex compared to other languages due to the presence of many multi-posture consonants. For example, research studies by Saengsri et al. (2012), Nakjai and Katanyukul (2019, 2021), and Pariwat and Seresangtakul (2021). The primary applications of VSLR techniques with TFSL are for communication between hearing individuals and those with hearing impairments. There is very little research that uses VSLR to be used as teaching and learning materials.

VSLR techniques can be categorized into five stages: image acquisition, image pre-processing, image segmentation, feature extraction, and classification. Adeyanju et al. (2021) describes stages of VSLR as follows. Image acquisition is the first stage in SLR that can be acquired through self-created or available public datasets. Researchers have used different devices including a camera or webcam, data glove, Kinect, and leap motion controller. Among these devices, a camera or webcam is the most widely used by many researchers because it provides better and natural human-computer interaction without additional devices, unlike data glove based. Data glove has proven to be more accurate in data acquisition but very costly and inconvenient for the users. Kinect is wildly used and effective. It provides both color video and depth video stream simultaneously, however it is expensive. The leap motion controller is a low-cost device with better accuracy than Kinect, but it can operate in a limited range.

The second stage is preprocessing to eliminate unwanted noise and enhance the quality of the image. This can be accomplished by resizing, color conversion, removing

noise, or a combination of techniques from the original image. Image preprocessing techniques can be classified into image enhancements and image restoration.

The third stage is image segmentation. In this stage, the regions of interest in the images are segmented and extracted from the entire image. There are two basic approaches used for segmentation: contextual and non-contextual segmentation. Contextual segmentation employs the relationships between the image features, such as edges, similar intensities, and spatial proximity. A non-contextual segmentation ignores spatial relationships between image features, but group pixels based on global attribute value (Sharma et al. 2021). There are various techniques in segmentation such as Thresholding, Edge detection, Region based, Clustering based, and Artificial neural network (ANN)-based segmentation techniques. ANN-based segmentation techniques are currently famous because they do not require a complex program to work, and they are less prone to noise. However, computational time in training is higher than other techniques. The background of the image also affects the quality of segmentation. In this research area, backgrounds are divided into 2 types: simple backgrounds and complex background. The simple background entails the use of a single color such as green, blue, or white. This can help hand segmentation work more easily, and the system can recognize accuracy at a high level. However, it is not practical due to the complexity of the backgrounds in real-life situations.

The fourth stage is features extraction, which transforms the input image region into feature vectors for recognition. It aims at finding the most distinctive features in the acquired image. The output is the compact feature vector, which is extracted by removing an irrelevant part to increase learning accuracy. The features extracted output supports the classification stage by checking for features that can effectively be distinguished between classes and help achieve high recognition accuracy. The features extraction techniques are used as hand detection algorithms in VSLR. There are several approaches and algorithms used for hand detection: Histogram of Oriented Gradients (HOG), Convolutional Neural Networks (CNNs), YOLO (You Only Look Once), MediaPipe Hands. MediaPipe Hands, developed by Google, is particularly notable for its ability to robustly detect and track human hands in images and videos, making it a valuable tool for applications ranging from gesture recognition and sign language interpretation to augmented reality and human-computer interaction. It employs a deep learning model trained on a large dataset of annotated hand images. This model can detect hands in a variety of poses and orientations, making it adaptable to different scenarios and user movements. One of the standout features of MediaPipe Hands is its suitability for real-time scenarios. It leverages a combination of machine learning and computer vision techniques to achieve impressive performance without sacrificing speed. This is essential for applications where quick and accurate hand tracking is crucial. Furthermore, MediaPipe Hands not only identifies the presence of hands but also provides detailed information about hand landmarks – key points on the hand’s surface – which can be used to infer gestures and movements.

The last stage is classification. After the pre-processing, segmentation, and extraction of features from the images have been completed, it is necessary to use a predictor algorithm to help give valuable meaning to the extracted features. Machines are trained to learn and machine learning improves their performance to match the features of the

new sign image with the stored features in the database for recognition of the given sign language. Machine learning is a subfield of computer science, and it is also classified as an AI method (Voyant et al. 2017). The artificial intelligent techniques used for sign language recognition include supervised or unsupervised. Supervised machine learning took in a set of known training data and used it to infer a function from labelled training data, whereas unsupervised machine learning is used to draw inferences from datasets with input data with no labelled response. Google's Teachable Machine is a web-based tool developed by Google that allows users to easily create and train custom machine learning models without requiring extensive coding or machine learning expertise. It's designed to help individuals understand the basics of machine learning through hands-on experience. Google's Teachable Machine primarily uses supervised machine learning. Users provide labeled examples to train the model, and the model learns to associate patterns in the data with specific classes or labels. With Google's Teachable Machine, users can train models to recognize patterns in different types of data, such as images, sounds, or gestures. The process usually involves providing the tool with labeled examples of data, allowing the model to learn the associations between input data and corresponding output categories.

3 Research Methodology

The research methodology is composed of 6 main steps as follows.

3.1 Program Development Concept

The process of developing the functional part of the program consists of three phases: the training phase, the testing phase, and the prediction phase. The flow diagram of functional part of the developed program is illustrated in Fig. 2.

In the training phase, a trained classifier named Keras.h5 is created. It is a neural network model capable of recognizing hand gesture images associated with Labels. The trained model is then tested with another set of hand gesture images to evaluate the accuracy of model before program deployment. Finally, the program is tested by the actual users, which is the prediction phase to evaluate the performance of the trained neural network model.

3.2 Training Phase of Program Development

The training phase follows 5 stages of VSLR techniques mentioned in the second session. In this research, the TFSL image dataset is a self-created dataset. Thus, the process starts with creating the dataset. After that, this dataset is used to train neural network model to create the trained classifier or trained neural network model. Therefore, there are two parts in this phase as follows.

Data Collection for Sign Language Dataset. This part starts with capturing the hand gesture image of all 42 TFSL using webcam. Three individuals are used to perform the hand gesture of these consonants with different angle and distance referred from

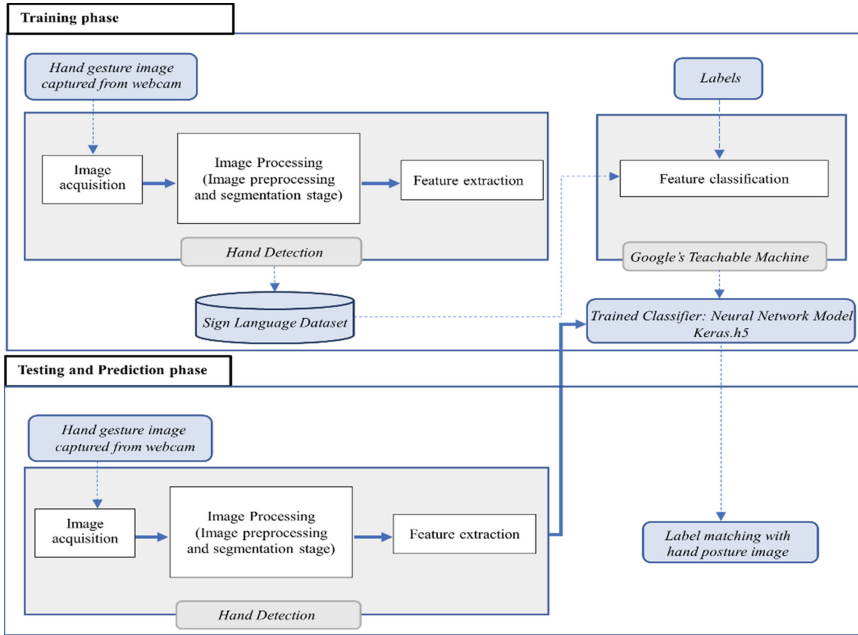


Fig. 2. Program flow diagram.

the webcam. There are 300 images, 500 images, and 800 images created for each one-posture consonant, two-posture consonant, and three-posture consonant respectively. Hand detection process is created using OpenCV and CVZone library of Hassan (2022), which use MediaPipe Hands library. The process starts with using computer vision to detect a hand from an input image and keeps focus on the hand’s movement and orientation. OpenCV is used to perform operations associated with computer vision. After that MediaPipe Hands library is used to perform the actual hand detection and tracking on input images. Using MediaPipe Hands, 21 landmarks according to hand and finger joints as shown in Fig. 3 are extracted. These landmarks are embedded in the hand posture image.

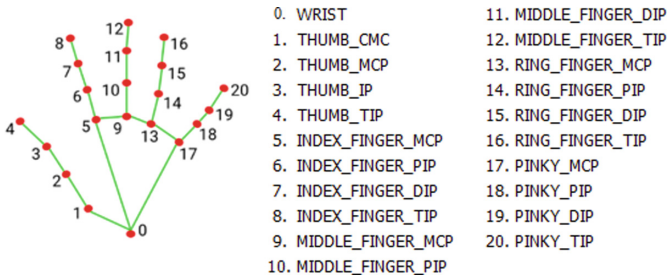


Fig. 3. Hand landmarks from MediaPipe Hands and their indices.

Creation of Trained Model. The dataset of hand posture images and landmarks obtained from the previous section is put into Google's Teachable machine. Google's Teachable Machine primarily uses a technique called transfer learning along with pre-trained neural network architectures to create machine learning models for image and audio classification tasks. Transfer learning involves taking a pre-trained model that has been trained on a large dataset and adapting it for a specific task with a smaller dataset. In this research, it is employed to create machine learning models for image classification from the TFSL dataset. The output obtained from this step is a trained model, which will be used as a trained classifier in the next phase, called Keras.h5.

3.3 Testing Phase of Program Development

The testing phase is to evaluate the trained model on the validation set to assess its performance. In this phase, the trained model named Keras.h5 will take features from another hand posture image dataset performing by the same individuals through the same process as in the data collection steps. After that, it yields predicted labels, which are consonants. The testing will be conducted with every consonant, each repeated 10 times, utilizing various angles and distances from the webcam. Subsequently, instances where the consonant card images are correctly matched to the TFSL gestures are recorded. The average percentage of accuracy for all consonants is then calculated.

3.4 Prediction Phase of Program Development

This phase is also referred to as program deployment. The process is composed of the same steps as mentioned in the testing phase but in this phase, it is done by the actual users. The actual users in this research are the hearing-impaired students who study in elementary grade and teachers from Setsatian School. The actual users have various sizes of hand, TFSL skills and stability of hand gesture, concentrations. These factors may affect the accuracy of the programs.

3.5 Graphic User Interface (GUI) Design

After completing the testing phase of program development, we can ensure that the program functions correctly according to our design. The program will then be used to create a GUI, enabling real users to easily use it, and providing an engaging user experience. One computer, one monitor, and one keyboard are required when using the program. The equipment setup, the hand image (input) and flashcard (output) displayed on the monitor are shown in Fig. 4(a) and Fig. 4(b) respectively.

To use the program, the user performs TFSL in front of webcam. The user captures the image by pressing "s" button on keyboard to send the image into the program and ends the program by pressing "a" button on keyboard. Then, the program shows the flashcard corresponding to the TFSL on monitor. For multi-posture TFSL, the user press "s" button for each posture and press "a" button when finish all of postures.

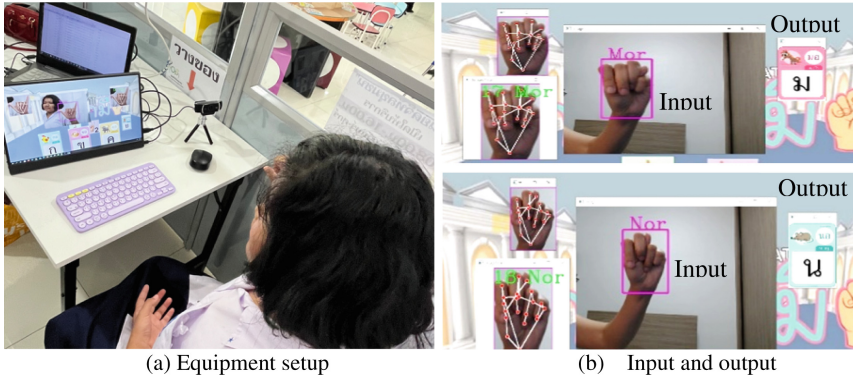


Fig. 4. GUI design (a) Equipment setup (b) Input and output.

3.6 Software Usability Assessment

After program deployment, the developed software usability is assessed. The evaluation is divided into 3 aspects: Aspect 1 represents the program’s functionality, Aspect 2 pertains to the GUI, and Aspect 3 relates to the program’s benefits. These are used to create 8 assessment topics in the questionnaire rating scale 1–5, as shown in Table 1. The rating scores are defined as indicated in Table 2. Criteria are also established to interpret the meaning of the average assessment scores obtained from participants evaluating program usage in each topic. The criteria are in Table 2.

Table 1. Topics of each aspect in the questionnaire.

Aspect 1	Aspect 2		Aspect 3
1. Precision 2. Response time	3. Ease of use 4. The interestingness in usage, featuring novelty and captivating engagement 5. Interestingness of display screen components	6. Text and image size on the screen 7. Displaying the results through flashcards to enhance comprehension	8. Program benefits in enhancing Thai consonant learning

The experiment participants totaled 12 individuals, consisting of male and female schoolchildren, with 10 individuals aged between 6 and 10 years. All participants were of Thai nationality and ethnicity, born with normal hearing, without any other disabilities, and were studying at the elementary level of the case study school. This age and educational level were chosen as they were currently learning Thai consonants according to the school curriculum. The teaching staff involved in the experiment comprised 2 teachers responsible for this class. In the program deployment phase, as shown in Fig. 5, the researcher will explain and demonstrate the program to non-hearing-impaired teachers.

Table 2. Score interpretation.

Score Level	1	2	3	4	5
Range of Average Score	1.00–1.49	1.50–2.49	2.50–3.49	3.50–4.49	4.50–5.00
Meaning	need improvement	somewhat satisfied	moderately satisfied	very satisfied	extremely satisfied

Subsequently, the teachers will convey the usage methods to students using sign language, along with demonstrating program usage. Afterward, the teachers will initiate the software usage trial, followed by the students. Participants in the trial will perform TFSL posture for every consonant twice and observe the flashcards of the letters displayed on the screen by the program.



Fig. 5. Demonstration of program usage and program deployment test.

In the evaluation phase, the researcher explained the objectives and meanings of each topic in the questionnaire to the teachers in detail. The teachers then translated and conveyed this information to the students participating in the software usage experiment using sign language. The translation was done topic by topic. Once the students completed the evaluation for one topic, the next topic was translated and conducted in the same manner until all 8 topics were completed.

4 Results and Discussion

The accuracy of the trained model reported by Google’s Teachable Machine is 100%. However, when tested in the testing phase, it was found that the average accuracy for all consonants decreased to 79.50%. This drop can be attributed to the use of still images for training the model, while real-time motion images were employed for the testing phase. Regarding the satisfaction assessment of the program usage, the results, obtained from 12 real user participants, are presented in Table 3 and visually represented in Fig. 6. From Table 3, overall average score of all assessment topics is 4.19, indicating that users are highly satisfied. When considering satisfaction in each aspect of the program, it was found that the average scores for functionality, GUI, and benefits of the program are

3.75, 4.32, and 4.42 respectively. These scores exceed 3.50 for all aspects, indicating that the users are very satisfied in all aspects. In terms of the functionality of the program, the assessment received the lowest scores, particularly in precision topic. Students gave an average score of 3.70, while teachers gave an average score of 3.50. The precision of the program when displaying the results has the lowest score, which could potentially be influenced by hand stability, hand size, and unfamiliarity with the program. The display will continuously change according to the user’s hand movements until the hand comes to a stop. This causes users, especially students, to take longer time to press the answer submission button and consequently results in low scores in the assessment of response time. In the GUI aspect, it was found that both students and teachers are very and extremely satisfied, indicating that the program is user-friendly, engaging, and captivating for learners. The visual components are appealing, and the presentation of consonants in the form of flashcards makes it easier for students to comprehend. In this regard, there is an additional recommendation that suggests the desire for larger image sizes of the flashcards.

Table 3. Average assessment score.

Aspect	Aspect 1		Aspect 2					Aspect 3
Topic	1	2	3	4	5	6	7	8
Teachers (2 persons)	3.50	4.50	4.50	4.00	4.00	5.00	4.50	4.50
Students (10 persons)	3.70	3.70	5.00	3.70	4.40	3.90	4.50	4.40
All participants	3.67	3.83	4.92	3.75	4.33	4.08	4.50	4.42
	3.75		4.32					4.42

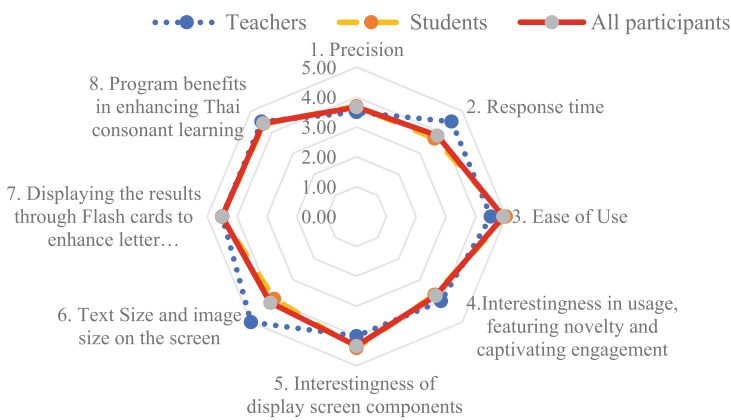


Fig. 6. Radar chart of average assessment score.

The program benefit aspect obtained the highest average score of 4.42, it indicates that overall users are highly satisfied with using the program to derive benefits in learning the Thai consonants. When looking at user groups, the student group provides an average score of 4.40, signifying they are very satisfied to use the program to enhance their TFSL skill. Meanwhile, it can be observed that teachers exhibit extreme satisfaction in utilizing the program for beneficial purposes, scoring up to 4.50. Teachers also provided additional comments that the program seems to be very beneficial in encouraging children to become more interested in learning consonants and can be used as a teaching aid to help children better understand Thai consonants.

5 Conclusion

According to the experimental results, the developed software can be effectively implemented through the designed functions. It can accurately capture students' hand gestures and translate them into corresponding flashcard images. Additionally, when assessing the practical usage with actual users, who are students encompassing all aspects including program functionality, GUI usage, and benefits, it was found that the satisfaction level was very high across all dimensions. The overall average score of satisfaction is 4.19 or 83.8%. The program for translating TFSL into Thai consonants can serve as an educational tool for children with hearing impairments, contributing to improved learning outcomes.

Acknowledgements. We would like to thank the school director of Setsatian School for granting permission for the research team to collect data, as well as the teachers and students of Setsatian School who dedicated their time to participate in the software testing and provided valuable suggestions.

References

- Adeyanju, I.A., Bello, O.O., Adegbeye, M.A.: Machine learning methods for sign language recognition: a critical review and analysis. *Intell. Syst. Appl.* **12**, Article no. 200056, 36 p. (2021)
- Alawad, H., Musyoka, M.: Examining the effectiveness of fingerspelling in improving the vocabulary and literacy skills of deaf students. *Creat. Educ.* **9**, 456–468 (2018)
- Hassan, M.: Easy Hand Sign Detection | American Sign Language ASL | Computer Vision (2022). <https://www.youtube.com/watch?v=wa2ARoUUdU8>
- Nakjai, P., Katanyukul, T.: Hand sign recognition for thai finger spelling: an application of convolution neural network. *J. Sign. Process. Syst.* **91**, 131–146 (2019)
- Nakjai, P., Katanyukul, T.: Automatic Thai finger spelling transcription. *Walailak J. Sci. Technol.* **18**(13), Article no. 11233, 19 p. (2021)
- Pariwat, T., Seresangtakul, P.: Multi-stroke Thai finger-spelling sign language recognition system with deep learning. *Symmetry* **13**(2), 262 (2021)
- Saengsri, S., Niennattrakul, V., Ratanamahatana, C.: TFRS: Thai finger-spelling sign language recognition system. In: 2012 Second International Conference on Digital Information and Communication Technology and it's Applications (DICTAP), pp. 457–462 (2012)

- Sarchet, T., Marschark, M., Borgna, G., Convertino, C., Sapere, P., Dirmyer, R.: Vocabulary knowledge of deaf and hearing postsecondary students. *J. Post Second. Educ. Disabil.* **27**, 161–178 (2014)
- Sharma, A., Badal, T., Gupta, A., Gupta, A., Anand, A.: An intelligent sign communication machine for people impaired with hearing and speaking abilities. *Commun. Comput. Inf. Sci.* **1367**, 75–86 (2021)
- Tumsri, J., Kimpan, W.: Thai sign language translation using leap motion controller. In: *Proceedings of the International Multi-Conference of Engineers and Computer Scientists*, vol. 1, pp. 46–51 (2017)
- Voyant, C., et al.: Machine learning methods for solar radiation forecasting: a review. *Renew. Energy* **105**, 569–582 (2017)
- William, C.: Promoting vocabulary learning in young children who are deaf and hard of hearing: translating research into practice. *Am. Ann. Deaf* **156**, 501–508 (2012)



A Preliminary Study on Developing and Training the Smart Site Safety System (SSSS) for Construction Industry in Hong Kong

Yan-Wai Chan¹, Simon K. S. Cheung², Kwan-Keung Ng³(✉), and Louise Luk²

¹ Hong Kong Institute of Construction, Hong Kong and Department of Science and Environmental Studies, The Education University of Hong Kong,

Tai Po, New Territories, Hong Kong
kevinywchan@hkic.edu.hk

² Hong Kong Metropolitan University, Kowloon, Hong Kong
{kscheung, lluk}@hkmu.edu.hk

³ Ming-Ai (London) Institute, London, UK
s.ng@ming-ai.org.uk

Abstract. In late February 2023, the Hong Kong government committed to fully implementing the Smart Site Safety System (SSSS) on a specific amount of public works for the Construction Industry in Hong Kong. Though the government supports a significant push by the public sector, it is foreseeable that different levels of challenges may be encountered at this initial stage of system development. This study focuses on studying the effect of the implementation of SSSS in the industry and discusses the challenges it will face. The study will include interviews with industry professionals in four fields, including those responsible for creating and selecting systems, training, supervision, implementation and execution. Therefore, interviews with these professionals are conducted to share their views and the doubts raised in the industry and opinions on the solutions and division of labour at various levels to understand better the implementation of a safe and smart construction site system.

Keywords: Smart Site Safety System (SSSS) · Internet of Things (IoT) · Centralised Management Platform (CMP) · Computerised Intelligent System · Construction Industry

1 Introduction

Since 2020, the government has started the development of a computerised intelligent system, the Smart Site Safety System (SSSS), in phases in public works contracts for the Construction Industry in Hong Kong. The Development Bureau of the Government of the Hong Kong Special Administrative Region (DEVB) has carried out a trial plan for the SSSS in 8 public works contracts of different natures - the adjustment and separation effect (Development Bureau, HKSAR 2022; Press Releases HKSAR 2022).

The SSSS aims to strengthen the existing safety management system with the machinery currently equipped with technology, such as Artificial Intelligence (AI) and Virtual Reality (VR). This system will help to change the traditional use of labour to monitor various processes and operations for safety and to reduce the workforce shortage or management.

During the initial implementation of SSSS in construction sites, difficulties may arise during the market development. Construction is characterised as an industry with high mobility of workers, constant changes in the site environment, and involvement of different contractors working on-site simultaneously. These characteristics may increase the risk of accidents. Under the recent direction of DEVB, the Hong Kong Institute of Construction (HKIC) developed a programme to provide training to industry stakeholders on the SSSS. The author was one of the course leaders for the course development, responsible for collecting and consolidating the opinions of contract stakeholders on how to implement the SSSS.

2 Construction Industry in Hong Kong

The construction industry has been notorious for its high accident rate globally. According to the data of the Hong Kong Labor Department (2022), the construction industry in Hong Kong had a total of 4,377 industrial accidents and a total of 51 fatal cases in 2022, accounting for 19.17% of the 266 fatalities in Hong Kong as a whole. It shows that the number of industrial fatalities at construction sites in various places accounts for a relatively high proportion among all industries. Therefore, finding a new method and technology is necessary to effectively reduce fatal industrial accidents on construction sites.

In order to cope with industrial casualties and haze, the Smart Construction Site (SCS) construction site system consists of sensors, displays, and computing elements with embedded intelligence and advanced digital applications. It has attracted more and more attention in the global construction industry. Interest has been aroused among the stakeholders in the industries, and applications have increased, both noted (Jiang et al. 2021). The purpose of promotion includes several industrial elements: a) Smart safety devices for monitoring high-risk construction activities and identifying safety hazards; b) act as a communication network for transmission of data collected from smart devices; and c) provide a centralised management platform as a one-stop hub for data analysis and alerts generation, as well as facilitating follow-up actions with potential hazards and abnormalities identified.

3 Design of the SSSS

The SSSS generally includes recognition monitoring systems, sensor transmission technology and data collection. In the design process of the SSSS, the implementation can be divided into four stages for implementation. Figure 1 shows that they are a) the Planning Stage, b) the Development Stage, c) the Implementation & training Stage, and d) the Operation Stage. All the approaches focus on five management resources aspects: People, Machine, Material, Method, and Environment (Wardani 2017), which must be adopted in all stages.

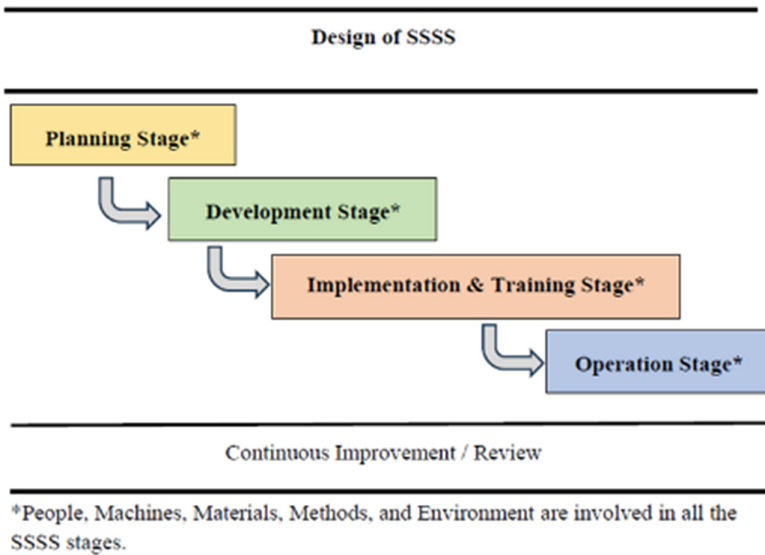


Fig. 1. Four Stages Design of the SSSS

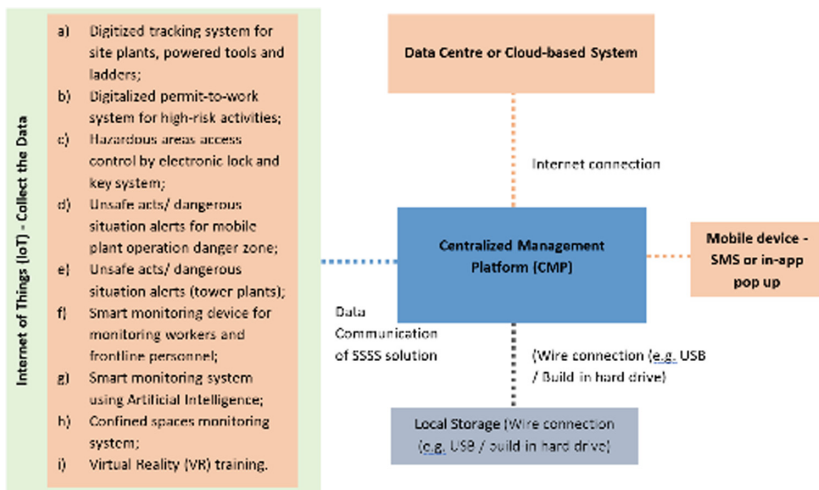


Fig. 2. Essential Elements of Implementing the SSSS

4 Process for Implementing the Smart Site Safety System

In consideration of the management on site, data are collected through sensors (i.e., Internet of Things (IoT), Bluetooth) (Brioso 2023), which includes identifying any gas present in the environment (e.g. Carbon Monoxide (CO), Hydrogen Sulfide (H₂S), Methane (CH₄) and other flammable gases. It also detects information in relation to the human body, e.g. body temperature, heart rate, falling and standstill detection) and

identification of objects/machines. It then converts the collected data into explicit data and application code. These data are centralised and processed in the system. These data will then be transmitted to the central management platform on the construction site through the on-site network (4G, 4.5G or 5G) for synchronisation of visual display and alarm for follow-up actions; IoT involves connecting physical devices and appliances to the internet to collect and exchange data. Also, it can be used to collect the location of on-site personnel data, including temperature, site humidity and fall from height alarm in real-time (Ding et al. 2022).

5 Essential Elements of Implementing the Smart Site Safety System

The focus of the implementation is to identify and mitigate/manage potential risks of on-site damage, strengthen on-site management, and mitigate the occurrence of accidents. Hong Kong Government implemented a compulsory Centralised management platform, and 9 items were given to the construction site for selective procurement to meet the feasibility of the site. It includes Fig. 2 a) Digitised tracking system for site plants, powered tools and ladders; b) Digitalised permit-to-work system for high-risk activities, c) Hazardous areas access control by electronic lock and key system; d) Unsafe acts/dangerous situation alerts for Mobile plant operation danger zone; e) Unsafe acts/dangerous situation alerts (tower plants); f) Smart monitoring device for monitoring workers and frontline personnel; g) Smart monitoring system using Artificial Intelligence; h) Confined spaces monitoring system; i) Virtual Reality training. The system can collect real-time data and transmit it to the one-stop site via, including but not limited to, Wi-Fi/UWB/LoRa/4G/4.5G/5G network/Ethernet/etc.

6 Methodology

The implementation of the new Smart Site Safety System (SSSS) was announced with a notice issued by the Development Bureau, effective from the end of February 2023 (DEVB 2023). In order to better understand their opinions and effectively promote the new system in class, interviews with these stakeholders are to be conducted. This study intends to incorporate the views from the four major areas and to provide feasible suggestions for the industry to solve problems.

The involvement of the professionals in these four fields is crucial in the pre-and post-development stages of the system as these four industry personnel play important roles and divisions of labour in the smart security system throughout. They are involved in the Planning Stage (creating), Development Stage (selecting and building), Implementation and Training Stage, and Operation Stage (supervising and executing). These representatives from the captioned four professional groups were requested to collect their industrial opinions and represent their respective sectors to sit for the sharing session.

The researchers invited a representative from four main fields of the construction industry to attend the sharing session, i.e., (i) Government Engineers, (ii) Supplier Representatives from the Centralised Management Platform, (iii) Site Safety Officers, and

(iv) Safety Trainers. In the foreseeable future, they will also be provided with training courses in the four main fields in the Construction Industry.

Research Questions:

- a. What is the difference between implementing computerised safety smart construction site systems and traditional safety management systems on construction sites in Hong Kong?
- b. What will be the impact on the construction site's working environment and conditions when launching the computerised Safety Smart Site System?
- c. When the construction companies' 4-stage Planning Stage (creating, Development Stage (selecting and building), Implementation and Training Stage, and Operation Stage (supervising and executing) the on-site computerised SSSS, what will you advise the construction sites?
- d. What is your opinion on the future technological development of the Safety Smart Site System in the construction sites in your professional field ((i) Government Engineers, (ii) Supplier Representatives, (iii) Site Safety Officers, and (iv) Safety Trainers)?

7 Interviews Report in the Sharing Session

(i) Government Engineer (GE) Representative

The Government Engineer (GE) commented that the engineering team has smart devices to monitor high-risk work, increasing vigilance and training on-site personnel to pay attention to safety. However, simultaneously, it should be noted that the limitation of equipment transmission distance should be fully understood as to each equipment's coverage and technical performance, including the transmission and selection system used by workers' smart helmets, etc. The system may also cause on-site SSSS due to network problems or main power interruptions during operation. There may be opportunities for direct real-time transmission and technology that only provides visual and auditory data, and the prohibition of workers from entering cannot be completed. However, it is believed that due to technological advancement and government support, the problem can be further resolved in the future. GE has a more positive attitude towards the future development of SSSS on the construction site.

(ii) Supplier Representative from the Centralised Management Platform (SP)

The Supplier Representative from the Centralised Management Platform (SP) believes that developing the SSSS will help accelerate the industry's demand for technological safety. However, at the same time, due to the rapid increase in the use of technology and the high competitiveness of development, the technology currently needs to achieve the actual outcome against the user's requirements and expectations. The deviation of the actual data and the delay of signal transmission may be due to the limitation of network speed, etc. Moreover, the data must be converted to the CMP in cooperation.

(iii) Site Safety Officer (SO) Representative

Site Safety Officer (SO) believes that intelligent devices can monitor practical risks, including knowing the implementation status of confined spaces and using artificial intelligence cameras to feature workers entering dangerous areas, with real-time

transmission and sensing systems to prevent accidents in real-time. However, the current technology still has a buffering period, although the implementation can be seen by the central system - the problem of worker violations. Nevertheless, it will take some time for the final interception.

(iv) Safety Trainer (ST) Representative

Safety Trainer (ST) commented that it is a big challenge for safety trainers as SSSS differs from traditional safety training courses. Traditional safety training uses traditional ideas to change the unsafe behaviours of human beings. It deepens basic safe site behaviours through education and learning about safety knowledge before they can work on-site. At the same time, the SSSS makes use of engineering control measures throughout the system to avoid accidents.

8 Discussion and Analysis

Smart Construction Technologies (SCTs) are the integration of advanced digital and technological solutions into the construction industry to improve productivity, efficiency, safety, and sustainability. These technologies leverage various innovations such as the Internet of Things (IoT), artificial intelligence (AI), data analytics, robotics, augmented reality (AR), and Building Information Modeling (BIM) to transform the way construction projects are planned, executed, and managed (Chen and Han 2021; Getuli et al. 2022; Jiang et al. 2021). SCTs are regarded as all types of technologies used for innovating and improving productivity based on providing information sharing and connectivity for all participants in all stages of construction work, such as in planning, design, construction, maintenance, and demolition (Kim et al. 2022). The responses to the research questions from the representatives in the sharing session are summarised as follows:

- a. What is the difference between implementing computerised safety smart construction site systems and traditional safety management systems on construction sites in Hong Kong?

SP and GE commented that it is different from the traditional ones; it can increase vigilance, train on-site personnel to pay attention to safety, and use technology to help monitor high-risk work.

- b. What will be the impact on the construction site's working environment and conditions when launching the computerised Safety Smart Site System?

ST believes improving the teaching mode regarding safety management training is necessary. In fact, it will take some time for SO to intercept in the end. The interruption of the main power supply may cause the interruption of on-site SSSS operation, and special personnel must be assigned to be responsible for such management and monitoring.

- c. When the construction companies' 4-stage Planning Stage (creating, Development Stage (selecting and building), Implementation and Training Stage, and Operation Stage (supervising and executing) the on-site computerised SSSS, what will you advise the construction sites?

SP thinks that technology is currently unable to create and select systems. It can fully achieve the actual effects and situations imagined by customers. Therefore, it is still necessary to carry out manually for final interception or prompting in preventing unsafe behaviours.

- d. What is your opinion on the future technological development of the Safety Smart Site System in the construction sites in your professional field ((i) Government Engineers, (ii) Supplier Representatives, (iii) Site Safety Officers, and (iv) Safety Trainers.

Both GE and SO believe that the limitation of equipment transmission distance should be fully understood and aware of. The possibility of failure in coverage and technical performance of each piece of equipment and the system due to network problems or interruption of the main power supply may lead to the interruption of on-site SSSS operation.

There are significant positive outcomes in teaching after the interview because the perspectives of various stakeholders were included, and they provided in-depth development directions and recommended measures. These comments help sustaining the implementation of construction safety design and teaching implementation guidelines. From the base on teaching plans and examples to illustrate how to use BIM design to pre-plan installation requirements and related installation areas, the use of 3D drawings to coordinate augmented reality (AR) and virtual reality (VR) training can help in ensuring the project to be consistent and matches with the design intention.

The opinions obtained can lead to more comprehensive training for on-site personnel to pay attention to safety matters and review the critical teaching points. The installation and implementation of this safe, smart site safety system hardware and searches for hazards through software based on risk assessment covers all involved processes. (For example, access to machinery/equipment, moving machinery/equipment to the installation location, lifting and erection methods of accessories, etc.).

It is also helpful to develop emergency support services and backup planning procedures for any events involving the network and a large number of sensors or facilities that may seriously disrupt on-site monitoring activities (such as power outages). Steps such as immediate suspension of all on-site work activities, having a backup workforce to monitor affected work, providing designated escape routes in case of emergency, and reporting the incident immediately, etc., can be included in the teaching materials.

9 Recommendation and Future Work

Under the so-called IR 4.0, the site safety management technology level will have become a new page in the future and will have achieved apparent results in the initial stage. They may have misunderstood and violated more organisational security policies (Gyekye et al. 2009). Therefore, it is still necessary to use manual interception when dealing with unsafe behaviours or when workers face hazardous environments.

After the implementation of the policy, the feasibility and compatibility of selecting appropriate or intelligent safety equipment and transmission equipment can benefit in reducing the number of occupational injuries and death rates under different constraints.

10 Conclusions

Hong Kong Institute of Construction assists in developing, coordinating, and training the Smart Site Safety System (SSSS) for the stakeholders in Hong Kong's Construction Industry. Thus, this paper aims to report a preliminary study on developing and training the SSSS for the Construction Industry in Hong Kong. The Development Bureau of the Government of the Hong Kong Special Administrative Region (DEVB) commissions the Construction Industry Council (CIC) to be the Construction Innovation Technology Fund (CITF) implementation partner.

There may be uncertainty on how the data can be collected in a unified manner as the workers on site are mobile and labour-intensive. The work locations of the workers on site vary on a daily basis. It will also be a challenge in the integration of the system.

References

- Brioso, X., Calderon-Hernandez, C.: Framework for integrating productive, contributory, and noncontributory work with safe and unsafe acts and conditions. *Int. J. Environ. Res. Public Health* **20**(4), 3412 (2023). <https://doi.org/10.3390/ijerph20043412>
- Building Department, HKSAR: Smart Site Safety System. Retrieved from 140723 (2023). https://www.bd.gov.hk/doc/en/resources/codes-and-references/practice-notes-and-circular-letters/circular/CL_SSSS2023e.pdf
- Chen, L., Han, P.: The construction of a smart city energy efficiency management system oriented to the mobile data aggregation of the Internet of Things. *Complexity* **2021**, 1–13 (2021). <https://doi.org/10.1155/2021/9988282>
- Construction Innovation and Technology Fund: About CITF. Retrieved from 100823 (2023). <https://www.citf.cic.hk/>
- Development Bureau, HKSAR: My Blog. New era of site safety - Smart site for the future. Retrieved from 140723 (2022). https://www.devb.gov.hk/tc/home/my_blog/index_id_1514.html
- Development Bureau, HKSAR: Smart Site Safety System. Retrieved from 010723 (2023). <https://www.devb.gov.hk/filemanager/technicalcirculares/en/upload/1393/1/C-2023-03-01.pdf>
- Ding, L., Jiang, W., Zhou, C.: IoT sensor-based BIM system for smart safety barriers of hazardous energy in petrochemical construction. *Front. Eng. Manag.* **9**(1), 1–15 (2022). <https://doi.org/10.1007/s42524-021-0160-6>
- Getuli, V., Capone, P., Bruttini, A., Sorbi, T.: A smart objects library for BIM-based construction site and emergency management to support mobile VR safety training experiences. *Constr. Innov.* **22**(3), 504–530 (2022). <https://doi.org/10.1108/CI-04-2021-0062>
- Gyekye, S.A., Salminen, S.: Educational status and organisational safety climate: does educational attainment influence workers' perceptions of workplace safety? *Saf. Sci.* **47**(1), 20–28 (2009). <https://doi.org/10.1016/j.ssci.2007.12.007>
- Jiang, W., Ding, L., Zhou, C.: Cyber physical system for safety management in smart construction site. *Eng. Constr. Arch. Manag.* **28**(3), 788–808 (2021)
- Kim, Y.-S., Lee, J.Y., Yoon, Y.-G., Oh, T.-K.: Effectiveness analysis for smart construction safety technology (SCST) by test bed operation on small- and medium-sized construction sites. *Int. J. Environ. Res. Public Health* **19**(9), 5203 (2022). <https://doi.org/10.3390/ijerph19095203>
- Liu, H., Song, J., Wang, G.: A scientometric review of a smart construction site in construction engineering and management: analysis and visualisation. *Sustainability* **13**(16), 8860 (2021)

Press Releases, HKSAR: Construction Safety Summit 2022 advocating revamp of site safety culture. Retrieved from 140623 (2022). <https://www.info.gov.hk/gia/general/202211/04/P2022110400785p.htm>

Wardani, R.: The analysis of interest to treatment outpatients back based on management resource approach (man, money, material, machine, method/5m). In: Proceeding Surabaya International Health Conference 2017, vol. 1, no. 1, July 2017



A Review of Smart Education Practices Across Disciplines

Billy T. M. Wong^(✉), Kam Cheong Li, and Mengjin Liu

Institute for Research in Open and Innovative Education, Hong Kong Metropolitan University,
Ho Man Tin, Kowloon, Hong Kong SAR, China
{tamiwong, kcli, mjliu}@hkmu.edu.hk

Abstract. Smart education has become a prevalent way of education delivery to enhance learning effectiveness and efficiency in the digital age. Despite smart education being implemented in a widening range of disciplines, there has been a lack of comprehensive analysis on the disciplinary variation of smart education practices. This paper presents a review of the disciplinary contexts of smart education practices, covering a total of 110 studies collected from Scopus and published from 2010 to 2022. The studies were analysed in terms of the learning modes, techniques, and devices utilised in four disciplinary categories, i.e., hard-pure, soft-pure, hard-applied, and soft-applied, under Biglan's classification. The results show that 15 disciplines were involved in smart education practices, among which language and computer science were the two most frequently reported. In regards to disciplinary categories, smart education has been practiced most frequently in hard-applied disciplines, and least in hard-pure disciplines. Three major modes of learning were identified, with face-to-face learning accounting for the largest proportion, followed by online learning and blended learning. Learning platforms/management systems and mobile devices were predominantly used in the practices. These findings provide a comprehensive overview of smart education practices within various subjects. They contribute to informing the implementation of smart education in relation to the nature of disciplines, as well as directions for future research on this topic.

Keywords: Smart education · smart learning · smart classrooms · disciplines · Biglan's classification

1 Introduction

Smart education is becoming increasingly prevalent alongside the rapid development and widespread use of smart technologies in the past decade. It refers to a form of technology-enhanced learning that emphasises the importance of technological design to enhance learning effectiveness and efficiency (Gros 2016), as well as adaptation to individual learners' needs to provide tailor-made support (Hwang 2014; Wong et al. 2023). Smart education provides various features such as learner-centredness and personalisation (Zhu et al. 2016) and the ability to learn anywhere and anytime (Bajaj and Sharma 2018).

Smart education has been implemented in a wide range of subjects (Li and Wong 2021, 2022a). Disciplinary variation is an essential contextual factor in the development of smart education environments and activities. Given the distinct natures of individual disciplines, disciplinary differences could greatly impact the design and delivery of teaching and learning (Neumann 2001; Trowler 2012).

Despite its importance, there has been a lack of comprehensive investigation on disciplinary variation in the implementation of smart education. Among the related studies, for example, Wong et al. (2023) identified the subjects involved in smart education practices, Li and Wong (2021, 2022a) analysed the trends in the subjects addressed by smart education over the years, and Cheng et al. (2019) focused on student learning outcomes across different disciplines in smart learning environments.

This paper reviews the disciplinary contexts of smart education. It aims to identify the patterns of smart education practices in relation to the learning modes, techniques, and devices across various disciplines.

2 Related Work

There have been a broad range of reviews covering various aspects of smart education. Several of them provide an overview of the development of smart education research and practice. For example, Chen et al. (2021) and Li and Wong (2022b) explored the major research topics and their evolution based on a topic modelling and bibliometric analysis of smart learning publications. Li and Wong (2021) analysed the patterns and trends of smart learning research and practice, covering research issues, research methods, and research participants, as well as various smart devices or tools and smart learning environments.

Technology is another major area of reviews on smart education. For example, Yoo et al. (2015) focused on the use of educational data mining and visualisation dashboards, and proposed an evaluative tool to assess the dashboard functions in a pedagogical and visual manner. Saini and Goel (2019) presented a survey of smart classroom technologies for content preparation, presentation and dissemination, student engagement and interaction, feedback, and assessment. The benefits and challenges brought about by the application of emerging technologies such as the Internet of Things (Kassab and Mazzara 2020) and artificial intelligence (Dimitriadou and Lanitis 2023) for smart education were also covered in similar work.

Related studies also addressed learner issues in smart education. For example, Karimah and Hasegawa (2022) surveyed the latest developments in relation to student engagement in smart education settings. Mbunge et al. (2021) reviewed the hybrid and deep learning models for predicting students' performance in smart learning environments. Li and Wong (2019) studied the effects of deep and surface learning, disorganisation, achievement goals, and self-efficacy for enhancing learners' metacognition for smart learning.

There have only been a few studies which investigated smart education practices within specific disciplinary contexts. Sokolova et al. (2018) examined the roles of English as a foreign language (EFL) teachers in smart learning environments, who serve as facilitators, collaborators/mentors, technicians, course designers, developers, managers,

and evaluators. Diogo et al. (2023) analysed the strengths, weaknesses, opportunities, and threats for engineering education in smart learning environments.

However, related work has not provided a comprehensive overview of smart education practices in various disciplinary contexts. This study addresses the literature gap through a survey of the empirical work on smart education across various subjects.

3 Research Method

This study aims to identify the patterns of smart education practices in higher education in relation to the learning modes, techniques, and devices across subjects. It adopted a systematic review methodology and followed the process suggested by Petticrew and Roberts (2006).

Relevant literature was collected from Scopus. The keywords “smart education”, “smart learning”, “smart teaching”, or “smart classroom” were used to search for relevant journal articles published in the period 2010–2022. An initial search returned 760 results, which were then screened according to a set of inclusion and exclusion criteria. The inclusion criteria were set as (1) focused on smart education, (2) written in English, and (3) available in full text. The exclusion criteria were defined as (1) a lack of detail about the disciplines, (2) a lack of educational-level detail, and (3) not in the context of higher education. Based on these criteria, 110 articles were finally selected for review.

Regarding the smart education practices reported in the articles, the subjects involved were identified and categorised according to *Research Fields, Courses, and Disciplines Classification*—a list of subjects developed by the Australian Bureau of Statistics (1998). As shown in Table 1, the disciplines were categorised as hard-pure, soft-pure, hard-applied, and soft-applied based on Biglan’s classification (Biglan 1973a, 1973b; Stoecker 1993), which has been highly recognised as an organisational system of subjects (Hativa and Marincovich 1995; Simpson 2017). Furthermore, information about the learning modes, techniques, and devices applied in the smart education practices was identified from the articles and categorised based on the coding scheme adapted from Li and Wong (2022a).

Table 1. Classification of subjects in this review

	Hard	Soft
Pure	Mathematics Earth science Physics Biology	Language Arts Political science
Applied	Computer science Engineering Medical science	Education Commerce and management Economics Social studies Law

4 Results

4.1 Overview of Publications

Figure 1 shows the number of publications on smart education from 2010 to 2022. There has been a clear upward trend in the number of publications since 2018, indicating the increasing popularity and application of smart education in recent years.

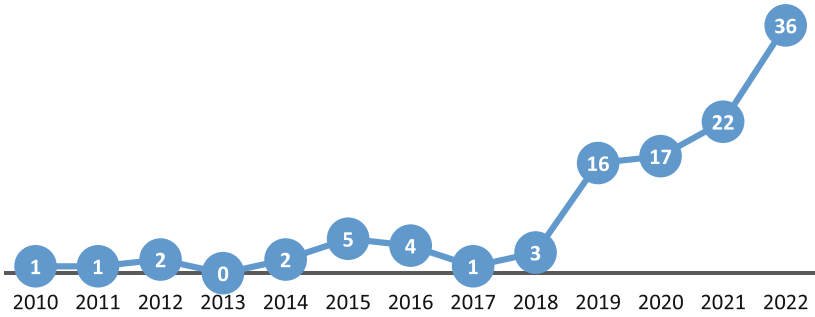


Fig. 1. Number of publications on smart education

4.2 Subjects

Figure 2 presents the subjects involved in the smart education practices. A total of 15 disciplines were covered, among which language and computer science were the two most frequently reported, followed by education, engineering, and mathematics. The results suggest that smart education has been practised across a diverse range of disciplines, and in specific disciplines smart education has been more frequently implemented.

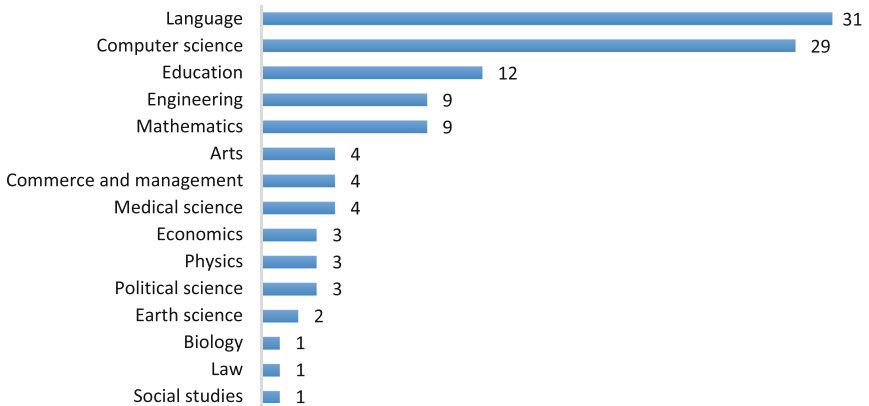


Fig. 2. Subjects reported in the studies

Figure 3 shows the distribution of subjects according to Biglan’s classification. Clearly, smart education has been most frequently adopted in hard-applied and soft-pure disciplines, accounting for 36% and 33% of the articles, respectively. This is due to the large number of smart education practices in computer science and engineering (hard-applied disciplines) and language (a soft-pure discipline). Comparatively, there are only 13% of smart education practices in hard-pure disciplines, including mathematics, physics, earth science, and biology, which were not frequently covered in smart education.

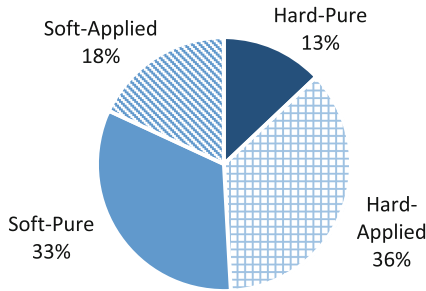


Fig. 3. Disciplinary categories according to Biglan’s classification

4.3 Learning Modes

Figure 4 shows the distribution of learning modes of smart education practices. The face-to-face mode accounts for the largest proportion (48%), while online learning and blended learning are 29% and 23%, respectively. The results suggest that smart education can be implemented in various learning modes.

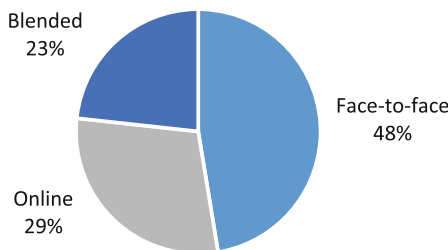


Fig. 4. Learning modes of smart education practices

Figure 5 presents the frequency of learning modes and the distribution of disciplinary categories under Biglan’s classification within each type of learning mode. The percentage of hard-pure disciplines under online learning (6%) is notably smaller than those under the face-to-face (16%) or blended learning mode (19%). On the other hand, there is a larger proportion of soft-applied disciplines under online learning (26%) than under face-to-face (16%) and blended learning (11%).

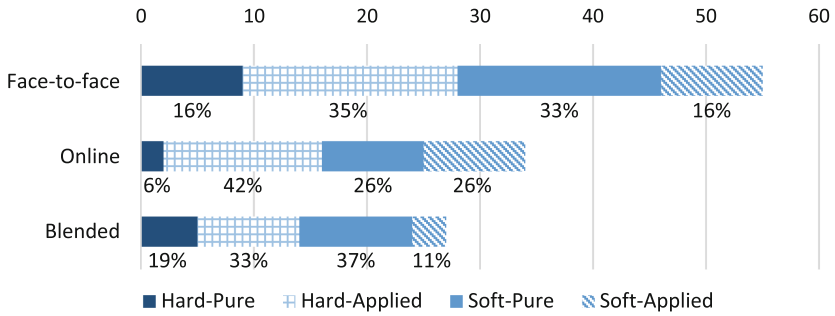


Fig. 5. Learning modes of smart education practices under Biglan’s classification

4.4 Techniques

Figure 6 presents the frequencies of techniques adopted in smart education practices and the distribution of disciplinary categories for each type of technique. Learning platforms/learning management systems is the most frequent type, followed by communication and interaction technologies. This pattern reveals the popularity of using learning management systems, which are available in most higher education institutions for smart education (Singh and Miah 2020). Other frequently used techniques include content display technologies, data mining/learning analytics, and simulation/virtual reality/augmented reality.

More than half of the techniques were applied under all four disciplinary categories. However, the application of computer games/mobile games was limited to applied disciplines, while context awareness technologies were only adopted in pure disciplines. In contrast to cloud computing, which was solely used in hard-applied disciplines, identification and authentication technologies as well as data mining/learning analytics were widely used in all except soft-applied disciplines.

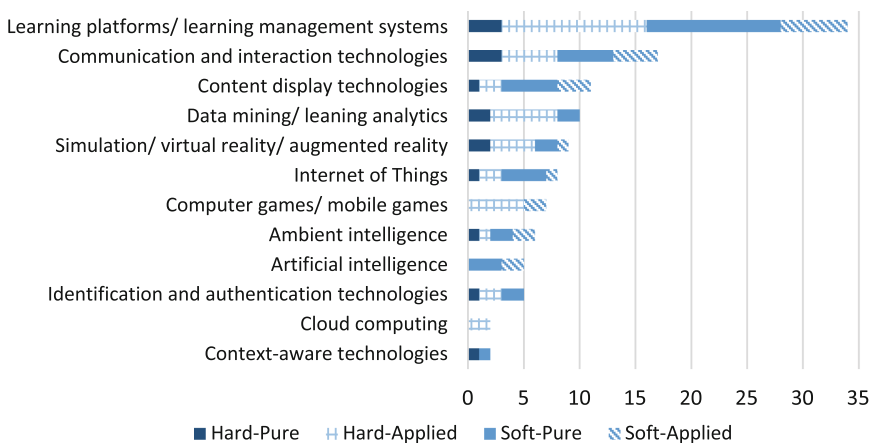


Fig. 6. Frequencies of techniques under Biglan’s classification

4.5 Devices

Figure 7 shows the frequencies of devices employed in smart education practices and the distribution of disciplinary categories for each type of device. It is evident that mobile devices such as tablets and smartphones are most popular. Desktop computers and interactive display devices were also frequently used, but their popularity is much lower than mobile devices.

These three most frequently used devices, together with non-interactive display devices, were adopted under all four disciplinary categories. The use of sensors covers three disciplines. Audio and visual devices were applied in both soft-pure disciplines (e.g., language) and soft-applied disciplines (e.g., education and political science). RFID/NFC readers were only used in a hard-applied discipline (i.e., computer science). Overall, the diversity of smart devices adopted in hard-pure disciplines is the lowest.

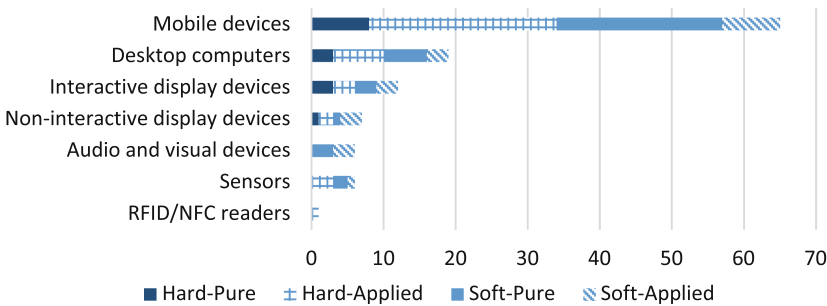


Fig. 7. Frequencies of devices under Biglan’s classification

5 Discussion and Conclusion

The findings of this study show that smart education has been practised across disciplinary contexts. They also reveal the widespread adoption of learning modes, techniques, and devices in different categories of disciplines.

Regarding the disciplinary contexts, language is the most common discipline involved in smart education practices. This is consistent with the results of previous reviews (Li and Wong 2021; Chen et al. 2021). This pattern may be related to the popularity of technological use for language education (Xu et al. 2019). Besides, the preference for smart education in computer science might be due to its convenience for running experiments, as well as the availability of required infrastructure (Tabuenca et al. 2021). Based on Biglan’s classification, hard-applied disciplines have been addressed by the largest proportion of practices. Since hard-applied disciplines are geared towards techniques (Neumann et al. 2002), teachers involved in these disciplines tend to have a high level of ability to integrate technology into their teaching (Starkey et al. 2023). This reflects how the nature of disciplines influences technological integration in smart education practices.

In terms of learning modes, the findings show that online learning was more frequently adopted in soft-applied disciplines than in hard-pure disciplines. This might be related to the various extent of social presence in different disciplines. It has been shown that social presence can positively influence students' participation and motivation in online learning environments (Mazzolini and Maddison 2007), as well as course and instructor satisfaction (Hostetter and Busch 2006). Lim and Richardson (2022) found that students studying in soft-applied disciplines showed a higher level of social presence in online courses than students in hard-pure disciplines. Students' various degree of social presence in relation to their study in disciplines of diverse natures would have an impact on their preference of learning mode in smart education practices.

The results regarding the techniques and devices employed in smart education show that most are not discipline-specific. They suggest that the techniques and devices are, in general, applicable to all disciplines. What would be more important is the pedagogy behind technological use. Being an integral part of smart education, the use of smart technology without an understanding of relevant pedagogical principles may pose a risk of centrifugal effects in the educational process (Daniela 2019). In this regard, future work should place more emphasis on the essential role of pedagogy in smart education, as well as the effective integration of technology and pedagogy in relation to the nature of various disciplines.

Acknowledgments. The work described in this paper was partially supported by grants from Hong Kong Metropolitan University (2021/011 and CP/2022/04).

References

- Australian Bureau of Statistics. Research Fields, Courses and Disciplines Classification (RFCDD) (1998). <https://www.abs.gov.au/ausstats/abs@.nsf/Previousproducts/1297.0Contents11998?opendocument&tabname=Summary&prodno=1297.0&issue=1998&num=&view=>
- Bajaj, R., Sharma, V.: Smart education with artificial intelligence based determination of learning styles. *Procedia Comput. Sci.* **132**, 834–842 (2018)
- Biglan, A.: The characteristics of subject matter in different academic areas. *J. Appl. Psychol.* **57**(3), 195–203 (1973a)
- Biglan, A.: Relationships between subject matter characteristics and the structure and output of university departments. *J. Appl. Psychol.* **57**, 204–213 (1973b)
- Chen, X., Zou, D., Xie, H., Wang, F.L.: Past, present, and future of smart learning: a topic-based bibliometric analysis. *Int. J. Educ. Technol. High. Educ.* **18**, Article no. 2 (2021)
- Cheng, L., Ritzhaupt, A.D., Antonenko, P.: Effects of the flipped classroom instructional strategy on students' learning outcomes: a meta-analysis. *Edu. Technol. Res. Dev.* **67**, 793–824 (2019)
- Daniela, L.: Smart pedagogy for technology-enhanced learning. In: Daniela, L. (ed.) *Didactics of Smart Pedagogy*, pp. 3–21. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-01551-0_1
- Dimitriadou, E., Lanitis, A.: A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. *Smart Learn. Environ.* **10**, Article no. 12 (2023)
- Diogo, R.A., dos Santos, N., Loures, E.F.: Digital transformation of engineering education for smart education: a systematic literature review. In: Ram, M., Xing, L. (eds.) *Reliability Modeling in Industry 4.0*, pp. 407–438. Elsevier (2023)

- Gros, B.: The design of smart educational environments. *Smart Learn. Environ.* **3**, Article no. 15 (2016)
- Hativa, N., Marincovich, M.: *Disciplinary Differences in Teaching and Learning: Implications for Practice*. Jossey-Bass (1995)
- Hostetter, C., Busch, M.: Measuring up online: the relationship between social presence and student learning satisfaction. *J. Scholarsh. Teach. Learn.* **6**(2), 1–12 (2006)
- Hwang, G.J.: Definition, framework and research issues of smart learning environments – a context-aware ubiquitous learning perspective. *Smart Learn. Environ.* **1**, Article no. 4 (2014)
- Karimah, S.N., Hasegawa, S.: Automatic engagement estimation in smart education/learning settings: a systematic review of engagement definitions, datasets, and methods. *Smart Learn. Environ.* **9**, Article no. 31 (2022)
- Kassab, M., Mazzara, M.: Towards designing smart learning environments with IoT. In: Bruel, J.M., Mazzara, M., Meyer, B. (eds.) *DEVOPS 2019*. LNCS, vol. 12055, pp. 152–166. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-39306-9_11
- Li, K.C., Wong, B.T.M.: Enhancing learners' metacognition for smart learning: effects of deep and surface learning, disorganisation, achievement goals and self-efficacy. *Int. J. Smart Technol. Learn.* **1**(3), 203–217 (2019)
- Li, K.C., Wong, B.T.M.: Review of smart learning: patterns and trends in research and practice. *Australas. J. Educ. Technol.* **37**(2), 189–204 (2021)
- Li, K.C., Wong, B.T.M.: How smart learning has been achieved: a review of the literature (2011–2020). *Int. J. Mob. Learn. Organ.* **16**(3), 310–322 (2022)
- Li, K.C., Wong, B.T.M.: Research landscape of smart education: a bibliometric analysis. *Interact. Technol. Smart Educ.* **19**(1), 3–19 (2022)
- Lim, J., Richardson, J.C.: Considering how disciplinary differences matter for successful online learning through the community of inquiry lens. *Comput. Educ.* **187**, 104551 (2022)
- Mazzolini, M., Maddison, S.: When to jump in: the role of the instructor in online discussion forums. *Comput. Educ.* **49**(2), 193–213 (2007)
- Mbunge, E., Fashoto, S., Mafumbate, R., Nxumalo, S.: Diverging hybrid and deep learning models into predicting students' performance in smart learning environments – a review. In: Ngatched, T.M.N., Woungang, I. (eds.) *PAAISS 2021*. LNICST, vol. 405, pp. 182–202. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-93314-2_12
- Neumann, R.: Disciplinary differences and university teaching. *Stud. High. Educ.* **26**(2), 135–146 (2001)
- Neumann, R., Parry, S., Becher, T.: Teaching and learning in their disciplinary contexts: a conceptual analysis. *Stud. High. Educ.* **27**(4), 405–417 (2002)
- Petticrew M., Roberts H.: *Systematic Reviews in the Social Sciences: A Practical Guide*. Blackwell (2006)
- Saini, M.K., Goel, N.: How smart are smart classrooms? A review of smart classroom technologies. *ACM Comput. Surv.* **52**(6), Article no. 130 (2019)
- Simpson, A.: The surprising persistence of Biglan's classification scheme. *Stud. High. Educ.* **42**(8), 1520–1531 (2017)
- Singh, H., Miah, S.J.: Smart education literature: a theoretical analysis. *Educ. Inf. Technol.* **25**(4), 3299–3328 (2020)
- Sokolova, D.G.V.N.L., Shinkaruk, A.R.M.V.M., Kiva-Khamzina, Y.L.: Exploring EFL teacher's role in a smart learning environment–review study. *XLinguae* **11**(2), 265–274 (2018)
- Starkey, L., et al.: Each discipline is different: teacher capabilities for future-focussed digitally infused undergraduate programmes. *Educ. Technol. Res. Dev.* **71**(1), 117–136 (2023)
- Stoecker, J.L.: The Biglan classification revisited. *Res. High. Educ.* **34**, 451–464 (1993)
- Tabuenca, B., et al.: Affordances and core functions of smart learning environments: a systematic literature review. *IEEE Trans. Learn. Technol.* **14**(2), 129–145 (2021)

- Trowler, P.: Disciplines and interdisciplinarity: conceptual groundwork. In: Trowler, et al. (eds.) *Tribes and Territories in the 21st Century: Rethinking the Significance of Disciplines in Higher Education*, pp. 5–29. Routledge, London (2012)
- Wong, B.T.M., Li, K.C., Chan, H.T.: A survey of smart learning practices: contexts, benefits, and challenges. *Int. J. Innov. Learn.* **34**(2), 122–142 (2023)
- Xu, Z., Banerjee, M., Ramirez, G., Zhu, G., Wijekumar, K.: The effectiveness of educational technology applications on adult English language learners' writing quality: a meta-analysis. *Comput. Assist. Lang. Learn.* **32**(1–2), 132–162 (2019)
- Yoo, Y., Lee, H., Jo, IH., Park, Y.: Educational dashboards for smart learning: review of case studies. In: Chen, G., Kumar, V., Kinshuk, ., Huang, R., Kong, S. (eds.) *Emerging Issues in Smart Learning*. LNET, pp. 145–155. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-44188-6_21
- Zhu, Z.T., Yu, M.H., Riezebos, P.: A research framework of smart education. *Smart Learn. Environ.* **3**, Article no. 4 (2016)



Study on the Influencing Factors of Primary School Students' Self-regulated Learning in the Smart Classroom: An Environmental Perspective

Yinghui Shi¹, Kexin Jia¹, Caiqin Cao¹, Yitong Wei², and Harrison Hao Yang^{1,3}(✉)

¹ Central China Normal University, Wuhan 430079, China

harrison.yang@oswego.edu

² Jiangnan University, Wuhan 430056, China

³ State University of New York at Oswego, Oswego 13126, USA

Abstract. Smart classrooms provide a blended and autonomous learning experience for students and greatly promote their learning. In particular, it can support students' self-regulated learning (SRL) process. However, few studies have explored how environmental factors in smart classrooms affect students' SRL. This study uses structural equation modeling techniques to examine the relationships between students' SRL and three environmental influence factors: teacher emotional support (TES), connected classroom climate (CCC), and facilitating conditions (FC). Research data gathered from 386 primary students in China reveal that TES, CCC, and FC are the main determining factors of students' SRL. Additionally, TES and CCC are two mediators between FC and students' SRL. The study proposes a research model of influencing factors in SRL. Furthermore, this study suggests implications for future research and instructional practice.

Keywords: Smart classroom · Self-regulated learning · Teacher emotional support · Connected classroom climate · Facilitating conditions

1 Introduction

The smart classroom, as a technology-enabled and student-centered learning environment, provides opportunities for student learning and participation in formal educational learning experiences (Macleod et al. 2018). The smart classroom enhances students' active learning processes while emphasizing flexibility for different types of instructional delivery (Li et al. 2019; Saini and Goel 2019). The smart classroom typically provides a blended learning experience by employing educational management software (e.g., learning management systems), which enables teachers to monitor, assess, and guide students through immediate feedback. Additionally, the smart classroom can effectively promote the interaction level between teachers and students, enhance students' learning interest, motivation, and engagement (Jena 2013), and thus improve students' learning effect.

Self-regulated learning (SRL) is a comprehensive learning process that requires learners to participate actively in learning in terms of metacognition, motivation, and behavior (Jansen et al. 2020). SRL emphasizes individual autonomy and control, monitoring, directing, and regulating actions to achieve the goals of information acquisition, professional enrichment, and self-improvement (Paris and Paris 2003). Students with high SRL abilities can plan, manage, and control their learning process, and they can thus learn faster and obtain better learning achievement (Kizilcec et al. 2017). Previous studies have also shown that using self-regulation in the learning process enables students to become “intelligent learners” (Zimmerman 2002). Additionally, research has shown that students with strong SRL skills are more likely to succeed (Broadbent 2017).

Many studies have shown that SRL has a great influence on student learning achievement (Wang et al. 2013). Therefore, what factors will affect students' SRL has been an important research hotspot. Based on Bandura's social learning theory, Zimmerman (1989) proposed a three-dimensional model of SRL, which revealed that students' SRL was determined by the interaction of three elements: internal, environmental, and behavioral influencing factors. Existing studies mostly focus on traditional learning environments and online learning environments, which have investigated the impact of students' internal and behavioral factors on SRL (e.g., Pekrun et al. 2002; Schunk 2023). In research related to smart classrooms and SRL, researchers have found that emerging technologies integrated into smart classrooms can effectively support students' SRL process (Singh and Miah 2020). However, few studies have systemically investigated the key factors that influence students' SRL from the environmental aspect. As Garcia et al. (2018) pointed out the current category of SRL which lacks research is environmental structuring. To address this research gap, this study attempts to investigate the relationship between primary students' SRL and key environmental influencing factors from psychological, emotional, and physical perspectives, which are connected classroom climate (CCC), teacher emotional support (TES), and facilitating conditions (FC) in the smart classroom environment. Specially, the research purposes of this study are:

1. Propose a structural research model of students' SRL for understanding students' SRL in smart classrooms.
2. Explore the relationships between factors of the proposed structural research model to promote students to achieve better achievement through SRL in smart classrooms.

2 Theoretical Framework

There are various environmental factors that play a crucial role in influencing student learning. A commonly accepted framework categorizes these factors into three perspectives: psychological, emotional, and physical (Western Governors University 2021). Physical factors encompass the elements within the learning environment, such as the furniture used to fill the room as well as the equipment utilized by the teacher and students to enhance the learning experience, which directly impact the learning experience. Psychological factors involve those that support student engagement with the curriculum and their peers within the learning environment. Emotional factors encompass the nurturing of emotional expression and the freedom to communicate feelings within the learning environment. In this section, we examine the key factors affecting

students' SRL from these three perspectives and subsequently propose a research model and hypotheses aligned with our findings.

2.1 Psychological Perspective: Connected Classroom Climate

According to Dwyer et al. (2004), CCC refers to students' perception of a supportive and collaborative communication environment in the classroom (Dwyer et al. 2004). This definition emphasizes the construct of connectedness among students. Previous studies have also shown that a positive classroom climate is beneficial for students. For example, CCC supports students with more engagement, stronger self-determined motivation, and more enjoyment of learning, and satisfies their psychological needs in traditional classrooms (Macleod et al. 2018). The smart classroom increases student-to-student relationships, which contributes to creating a positive classroom climate. Researchers have determined that CCC has effects on students' SRL. It is suggested that a positive classroom climate contributes to students' SRL (Macleod et al. 2018). In addition, Paris and Newman (1990) suggested that CCC influenced students' self-regulated strategies.

2.2 Emotional Perspective: Teacher Emotional Support

Teacher emotional support (TES) means that the teacher likes and cares about the student as an individual (Patrick et al. 2007). Tvedt et al. (2021) defines it as a factor that reflects the interpersonal bonding between teacher and student and is assumed to support students' need for relatedness. When students perceive that their teachers appreciate them, acknowledge their perspectives, and provide feedback, they are more likely to reach positive academic and psychological outcomes. Therefore, TES is an important environmental factor that affects individual development (Li et al. 2019), since teachers can create opportunities for students to support one another (Whipp and Chiarelli 2004; Perry et al. 2002). Teachers constantly interact with students in the classroom, such as teaching, questioning, praising, and criticizing. All those interactions influence students' self-regulation (Burić 2015). For example, teachers provide more autonomy support to students with a higher level of SRL, while they may more like those students who are more motivated and put more effort into their learning, and give them more support (Schuitema et al. 2016). In addition, there is a possible link between CCC and TES, because teachers are inclined to encourage student-to-student connectedness to facilitate students' involvement in class (Sidelinger and Booth-Butterfield 2010).

2.3 Physical Perspective: Facilitating Conditions

From the physical perspective, facilitating conditions refer to the degree to which an individual believes that an effective organizational and technical infrastructure exists to support the use of the system (Venkatesh et al. 2003). In a technology-enabled learning environment, these conditions encompass access to technology, technical support, and organizational support. Previous studies have found that access to technology equips students with digital tools, enhancing their digital literacy and self-directed learning capabilities (US Department of Education 2017). Technical support, provided by teachers

and IT staff, ensures students' confidence and competence in utilizing technology, a fundamental aspect of self-regulated learning (Songkram et al. 2023).

Meanwhile, organizational support refers to the policies and procedures that support the use of technology in the classroom. This can include guidelines for the appropriate use of technology, as well as training and professional development opportunities for teachers (US Department of Education 2017). When teachers are well trained and supported in their use of technology, they are better able to create a supportive and engaging learning environment that promotes self-regulated learning (Songkram et al. 2023).

2.4 Research Model and Mediating Variables

Based on our literature review, the research model is established. As shown in Fig. 1, the following hypotheses are proposed:

Hypothesis 1 (H1): TES has a positive impact on students' SRL in the smart classroom.

Hypothesis 2 (H2): CCC has a positive impact on students' SRL in the smart classroom.

Hypothesis 3 (H3): TES has a positive impact on students' CCC in the smart classroom.

Hypothesis 4 (H4): FC has a positive impact on students' TES in the smart classroom.

Hypothesis 5 (H5): FC has a positive impact on students' SRL in the smart classroom.

In addition, this research model indicates that CCC and TES could also act as mediating variables. A mediating variable is a variable that transmits influence between an independent and dependent variable. In other words, mediating variables help explain relationships between independent and dependent variables. Thus, in this study, it is expected that CCC and TES mediate the relationship between FC and students' SRL.

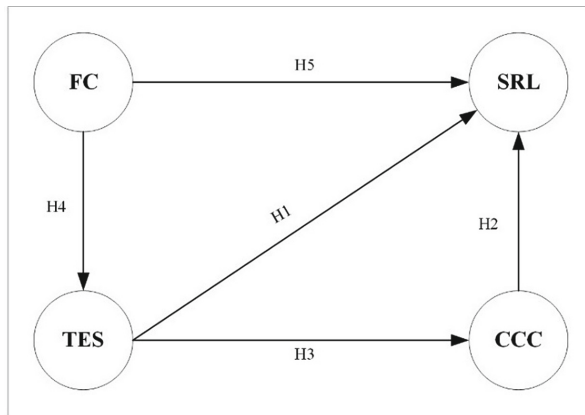


Fig. 1. Influencing factor model of students' SRL.

3 Research Methods and Processes

3.1 Sample and the Smart Classroom Learning Environment

The study sample is students from three grades (Grades 3 to 5) in a public primary school in central China. This school had smart classrooms and carried out regular teaching applications in grades three to five. All students had previous experience learning in the smart classroom. Students participated in an online survey. A total of 382 questionnaires were collected.

The smart classroom involved in this study is a tablet-based learning environment, including an interactive touch-controlled all-in-one computer, which is mainly used for presenting digital resources and instructional content. The teachers can quickly analyze student learning, send students learning materials, and provide real-time feedback and interaction with students. Each student is equipped with a tablet computer that can interact with the all-in-one computer, including receiving resources from teachers and submitting answers to in-class tests. In addition, the smart classroom is also equipped with relevant teaching systems, such as a cloud-based learning platform, intelligent interactive recording, and a broadcasting system.

3.2 Instruments

By referring to the relevant measurement standards and scales that were used in previous studies on SRL (Hung et al. 2010), TES (Tvedt et al. 2021), CCC (Dwyer et al. 2004), and FC, as parts of the UTAUT scale (Venkatesh et al. 2003), this study developed a questionnaire by integrating these scales. The questionnaire is divided into two parts, part one is the basic information about students (3 items), such as name, and gender. Part two includes four subscales (18 items in total), namely SRL (5 items), FC (5 items), TES (4 items), and CCC (4 items), and each item is designed with a 5-level Likert scale style, from one (strongly disagree) to five (strongly agree).

3.3 Data Collection and Analysis

Data collection for this study was conducted via an online survey. All the students were asked to complete the questionnaire within 30 min. They were told that the survey was voluntary and anonymous, and the information they provided would only be used for educational research. After completing the questionnaire, all the answers were downloaded and then imported into SPSS and Amos for data analysis. The structural equation modeling (SEM) method was employed to investigate students' SRL in smart classrooms. The Partial Least Squares (PLS) method was used to verify the research model presented in this study and it is appropriate for the sample size of this study (Chin 1998). The procedure of PLS-SEM analyses includes (1) evaluating the measurement of the research model (e.g., validity, reliability, and goodness-of-fit), then (2) assessing the path coefficients of direct and indirect effects in the structural model. The analyses were conducted with SPSS 26 and AMOS 24.

4 Data Statistics and Analysis

4.1 Overview of the Survey

Table 1 provides an overview of the participants' mean scores and standard deviations on this survey. The results showed responses that range between 4.33 and 4.48. FC ($M = 4.33$, $SD = 0.81$) was shown to have the lowest average response value examined. Each factor was shown to have a high average response value in the survey.

Table 1. Descriptive statistics of the survey

Scale	M	SD
SRL	4.48	0.58
CCC	4.42	0.78
TES	4.41	0.73
FC	4.33	0.81

4.2 Validity and Reliability Test

Validity was assessed in terms of the convergent and discriminant validity. As shown in Table 2, AVE values showed a range of 0.59 to 0.70. These results were all greater than 0.50, which validated the convergent validity of the construct and indicated that the items of each scale were appropriately correlated with each other. As shown in Table 2, almost the square root of each AVE was greater than the respective correlation coefficient, which suggested acceptable discriminant validity for all measures.

As shown in Table 2, coefficient alpha values from the participants' responses showed a range of 0.75–0.89 and CR coefficients showed a range of 0.87–0.90. These values were all above 0.70, confirming adequate reliability. Accordingly, the consistency of participants' scores collected in the present study was considered to have good reliability and relatively little error.

Additionally, to test the measurement model a confirmatory factor analysis (CFA) was conducted. Different fit indices were used and the results obtained were as shown in Table 3. The values of all fit indices were acceptable. Therefore, the goodness-of-fit for the proposed model was fairly good.

4.3 SEM Analysis

To verify the research hypotheses, a SEM analysis was conducted. In addition to the standardized regression coefficients (β), Table 4 shows the associated T statistics and the significance levels (p-value). The results showed that the five hypotheses were all supported by the proposed model. As shown in Fig. 2, the values of R² showed a range of 0.44–0.58.

Table 2. Validity and reliability analysis of the measurement model

	Reliability		Convergent validity	Discriminant validity			
	CR	α	AVE	SRL	CCC	TES	FC
SRL	0.88	0.88	0.59	0.77			
CCC	0.90	0.89	0.68	0.63	0.83		
TES	0.90	0.89	0.64	0.67	0.74	0.80	
FC	0.87	0.75	0.70	0.57	0.65	0.63	0.84

Boldface numbers represent the square roots of the AVE. AVE = average variance extracted; CR = composite reliability.

Table 3. Goodness-of-Fit analysis of the measurement model

	Fit index	Actual value	Recommended value	Judgment
Absolute fit indices	χ^2/df	2.820	≤ 3.00	Yes
	RMSEA	0.069	≤ 0.08	Yes
Incremental fit measurement	CFI	0.947	≥ 0.90	Yes
	TLI	0.937	≥ 0.90	Yes
	NFI	0.920	≥ 0.90	Yes

χ^2 = Chi-square; df = Degree of freedom; RMSEA = Root mean square error of approximation; CFI = Comparative fit index; TLI = Tucker-Lewis index; NFI = Normed fit index.

Table 4. Test of Hypotheses

Hypotheses	Independent variable	Dependent variable	β	T	Result of hypotheses
H1	TES	SRL	0.490	5.915***	Supported
H2	CCC	SRL	0.257	3.315***	Supported
H3	TES	CCC	0.762	12.575***	Supported
H4	FC	TES	0.673	11.09***	Supported
H5	FC	SRL	0.182	2.736**	Supported

** $p < 0.01$; *** $p < 0.001$. β = beta correlation.

4.4 Mediation Analysis

To examine CCC and TES’s mediating effect, the statistical significance of the indirect effect was examined by adopting bootstrapping. As shown in Table 5, two significant mediated paths were found, indicating that TES and CCC mediated the relationships between FC and student’s SRL.

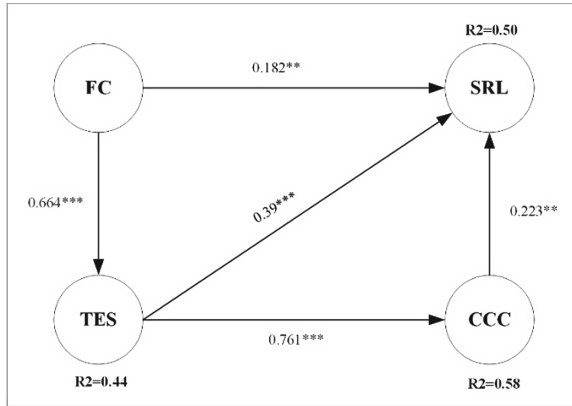


Fig. 2. Model hypothesis verification and path coefficient.

Table 5. Results of Mediation Analysis

Model Path	β	T	95% CI		Result of hypotheses
			LL	UL	
FC → TES → SRL	0.259	2.88**	0.10	0.463	Supported
FC → TES → CCC → SRL	0.112	1.84*	0.003	0.245	Supported

Estimates represent 5000 bootstrapping testing. CI = confidence interval, LL = lower limit; UL = upper limit. * $p < 0.05$, ** $p < 0.01$.

5 Discussion and Conclusion

This study analyzes the factors affecting primary school students' SRL in the smart classroom environment through an online questionnaire survey. The influencing factors of students' SRL were examined from the environmental perspectives, which are categorized into three aspects: psychological, emotional, and physical perspectives. The structural equation model is used to verify and analyze the relationship between FC, TES, CCC, and students' SRL in the smart classroom environment. The results show that TES, FC, and CCC have a significant and directly positive impact on students' SRL. Upon conducting further analysis, it becomes evident that FC indeed plays a significant and positive role in enhancing students' SRL within the smart classroom environment. In other words, FC not only has a direct and significant impact on students' SRL but also exerts an indirect but significant impact on students' SRL as mediated by two key factors: TES and CCC. To this end, teachers should make full use of the advantages of the smart classroom, to improve students' SRL by providing timely feedback and support to students and constructing a collaborative communication environment with information and communication technology.

Despite some previous SRL research focusing on other learning environments such as traditional classrooms, this study contributed the initial evidence of students' SRL

research examining smart classrooms in primary education. The major findings of this study indicated that FC, TES, and CCC directly impacted students' SRL in smart classrooms, and they collectively accounted for 50% of the R². These findings supported H1 (TES), H2 (CCC), and H5 (FC). FC positively influenced TES and collectively accounted for 44% of the R². These findings confirmed H4 (FC). The TES positively influenced CCC and collectively accounted for 58% of the R². These findings confirmed H3 (TES). Additionally, TES and CCC mediated the relationships between FC and students' SRL.

Firstly, the teacher's emotional support (TES), as well as students' perception of a supportive and collaborative communication environment in the classroom (i. e. CCC) were the main deciding factors of students' SRL examined in this study. Additionally, TES had a positive effect on CCC. This finding is consistent with previous studies (e.g., Schuitema et al. 2016; Sidlinger and Booth-Butterfield 2010). Teachers should encourage student-to-student connectedness to facilitate students' learning involvement through educational software. Therefore, this study suggests that TES is the most important determinant of students' SRL in smart classrooms. TES not only directly influences students' SRL but also influences CCC. If students' SRL is to be most effectively improved, TES needs to be provided that support, assist, and encourage students' efforts for active learning. Previous research suggested that teachers could provide students with academic support by offering appropriate support in the form of feedback and guidance (Law 2011) and providing emotional support through behaviors such as caring, respect and concern for students, listening, and encouragement (Sakiz et al. 2012). Given CCC is a very important factor in improving students' SRL, it is essential to encourage classroom connectedness. Teachers should consider students' peer interaction and preferences toward the smart classroom when analyzing, designing, developing, implementing, and evaluating learning activities in such an environment to maximize the potential of smart classrooms (Lu et al. 2021). For example, with a hybrid instructional model that encourages interaction and incorporates student-centered constructivist learning activities, students' perception of CCC is higher than other models (Xu et al. 2018).

Secondly, FC significantly affects TES, and FC influences students' SRL indirectly through TES and CCC. The finding suggests that when educational institutions select a smart classroom service provider, they must consider whether the instructional and technical support is readily available to respond to student needs (Hanham et al. 2021). Hence, the schools should provide adequate infrastructural facilities and resources, such as learning resources, tablet equipment, and smooth network links, to the students. It will undouble promote the use of smart classrooms by students and teachers (Li and Zhao 2021).

It should be noted that the present study has certain limitations. Firstly, participants in this study were limited to one primary school in central China, so the applicability of the findings needs to be further improved. Therefore, more students from different levels of education and different schools are encouraged to participate in future studies. Secondly, this study only examines three factors that affect students' SRL in the smart classroom. More influencing factors, such as factors from parents' perspective can be considered in future studies.

References

- Broadbent, J.: Comparing online and blended learner's self-regulated learning strategies and academic performance. *Internet High. Educ.* **33**, 24–32 (2017). <https://doi.org/10.1016/j.iheduc.2017.01.004>
- Burić, I.: The role of social factors in shaping students' test emotions. A mediation analysis of cognitive appraisals. *Soc. Psychol. Educ.* **18**(4), 785–809 (2015)
- Chin, W.W.: The partial least squares approach to structural equation modeling. *Mod. Methods Bus. Res.* **295**(2), 295–336 (1998)
- Dwyer, K.K., Bingham, S.G., Carlson, R.E., Prisbell, M., Cruz, A.M., Fus, D.A.: Communication and connectedness in the classroom: development of the connected classroom climate inventory. *Commun. Res. Rep.* **21**(3), 264–272 (2004)
- Garcia, R., Falkner, K., Vivian, R.: Systematic literature review: self-regulated learning strategies using e-learning tools for computer science. *Comput. Educ.* **123**, 150–163 (2018)
- Hanham, J., Lee, C.B., Teo, T.: The influence of technology acceptance, academic self-efficacy, and gender on academic achievement through online tutoring. *Comput. Educ.* **172**, 104252 (2021)
- Hung, M.L., Chou, C., Chen, C.H., Own, Z.Y.: Learner readiness for online learning: scale development and student perceptions. *Comput. Educ.* **55**(3), 1080–1090 (2010)
- Jansen, R.S., van Leeuwen, A., Janssen, J., Conijn, R., Kester, L.: Supporting learners' self-regulated learning in massive open online courses. *Comput. Educ.* **146**, 103771 (2020)
- Jena, P.C.: Effect of smart classroom learning environment on academic achievement of rural high achievers and low achievers in science. *Int. Lett. Soc. Humanist. Sci.* **3**(3), 1–9 (2013)
- Kizilcec, R.F., Pérez-Sanagustín, M., Maldonado, J.J.: Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Comput. Educ.* **104**, 18–33 (2017)
- Law, Y.K.: The role of teachers' cognitive support in motivating young Hong Kong Chinese children to read and enhancing reading comprehension. *Teach. Teach. Educ.* **27**(1), 73–84 (2011)
- Li, Y., Yang, H.H., MacLeod, J.: Preferences toward the constructivist smart classroom learning environment: examining pre-service teachers' connectedness. *Interact. Learn. Environ.* **27**(3), 349–362 (2019)
- Li, Y., Zhao, M.: A study on the influencing factors of continued intention to use MOOCs: UTAUT model and CCC moderating effect. *Front. Psychol.* **12**, 528259 (2021)
- Lu, K., Yang, H.H., Shi, Y., Wang, X.: Examining the key influencing factors on college students' higher-order thinking skills in the smart classroom environment. *Int. J. Educ. Technol. High. Educ.* **18**(1), 1–13 (2021)
- Macleod, J., Yang, H.H., Zhu, S., Li, Y.: Understanding students' preferences toward the smart classroom learning environment: development and validation of an instrument. *Comput. Educ.* **122**, 80–91 (2018)
- Paris, S.G., Newman, R.S.: Development aspects of self-regulated learning. *Educ. Psychol.* **25**(1), 87–102 (1990)
- Paris, S.G., Paris, A.H.: Classroom applications of research on self-regulated learning. In: *Educational Psychologist*, pp. 89–101. Routledge (2003)
- Patrick, H., Ryan, A.M., Kaplan, A.: Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *J. Educ. Psychol.* **99**(1), 83 (2007)
- Pekrun, R., Goetz, T., Titz, W., Perry, R.P.: Academic emotions in students' self-regulated learning and achievement: a program of qualitative and quantitative research. *Educ. Psychol.* **37**(2), 91–105 (2002)

- Perry, N.E., VandeKamp, K.O., Mercer, L.K., Nordby, C.J.: Investigating teacher-student interactions that foster self-regulated learning. *Educ. Psychol.* **37**(1), 5–15 (2002)
- Saini, M.K., Goel, N.: How smart are smart classrooms? A review of smart classroom technologies. *ACM Comput. Surv. (CSUR)* **52**(6), 1–28 (2019)
- Sakiz, G., Pape, S.J., Hoy, A.W.: Does perceived teacher affective support matter for middle school students in mathematics classrooms? *J. Sch. Psychol.* **50**(2), 235–255 (2012)
- Schuitema, J., Peetsma, T., van der Veen, I.: Longitudinal relations between perceived autonomy and social support from teachers and students' self-regulated learning and achievement. *Learn. Individ. Differ.* **49**, 32–45 (2016)
- Schunk, D.H.: Self-regulation of self-efficacy and attributions in academic settings. In: *Self-regulation of Learning and Performance*, pp. 75–99. Routledge (2023)
- Sidelinger, R.J., Booth-Butterfield, M.: Co-constructing student involvement: an examination of teacher confirmation and student-to-student connectedness in the college classroom. *Commun. Educ.* **59**(2), 165–184 (2010)
- Singh, H., Miah, S.J.: Smart education literature: a theoretical analysis. *Educ. Inf. Technol.* **25**(4), 3299–3328 (2020)
- Songkram, N., Chootongchai, S., Osuwan, H., Chuppunnarat, Y., Songkram, N.: Students' adoption towards behavioral intention of digital learning platform. *Educ. Inf. Technol.* **28**, 11655–11677 (2023)
- Tvedt, M.S., Bru, E., Idsoe, T.: Perceived teacher support and intentions to quit upper secondary school: direct, and indirect associations via emotional engagement and boredom. *Scand. J. Educ. Res.* **65**(1), 101–122 (2021)
- US Department of Education: Reimagining the role of technology in education 2017 national education technology plan update. National Education Technology Plan Update (2017)
- Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: toward a unified view. *MIS Q.* **27**(3), 425–478 (2003)
- Wang, C.-H., Shannon, D.M., Ross, M.E.: Students' characteristics, self-regulated learning, technology self-efficacy, and course outcomes in online learning. *Distance Educ.* **34**(3), 302–323 (2013). <https://doi.org/10.1080/01587919.2013.835779>
- Western Governors University: 3 Types of learning environments (2021). <https://www.wgu.edu/blog/3-types-learning-environments2111.html>
- Whipp, J.L., Chiarelli, S.: Self-regulation in a web-based course: a case study. *Educ. Technol. Res. Dev.* **52**(4), 5–21 (2004)
- Xu, J., Yang, H.H., MacLeod, J.: STEP on connected classroom climate in a hybrid learning environment. *Int. J. Innov. Learn.* **23**(4), 430–443 (2018)
- Zimmerman, B.J.: Models of self-regulated learning and academic achievement. In: Zimmerman, B.J., Schunk, D.H. (eds.) *Self-Regulated Learning and Academic Achievement*. SSCD, pp. 1–25. Springer, New York (1989). https://doi.org/10.1007/978-1-4612-3618-4_1
- Zimmerman, B.J.: Becoming a self-regulated learner: an overview. *Theory Pract.* **41**(2), 64–70 (2002)

Artificial Intelligence in Education



Integration of AI Learning into Higher Education: A Case of Using Microsoft Learn for Educators

Shao-Fu Li¹, Kwan-Keung Ng², and Lap-Kei Lee³(✉)

¹ Chung Hua University, Hsinchu, Taiwan
shaofu@g.chu.edu.tw

² Ming-Ai (London) Institute, London, UK
s.ng@ming-ai.org.uk

³ Hong Kong Metropolitan University, Kowloon, Hong Kong, China
lkleee@hkmu.edu.hk

Abstract. The integration of AI Learning into modern higher education has become a prevailing trend. Maintaining a student-centered approach to AI teaching is vital, considering students' interests and needs to enhance engagement. Continuous updating of AI course materials is necessary to keep pace with the ever-changing AI field. Microsoft Learn for Educators (MSLE) offers valuable AI courseware, enabling non-IT teachers to incorporate AI content into classes. This paper presents a case of how MSLE can be used for integrating AI Learning into an architectural computer graphics course at a Taiwan University, which details a revised implementation process and experience spanning from the academic year 2019–2020 to the present, and proposes a course plan for the upcoming academic term. The case sharing in this paper is expected to facilitate education practitioners without an IT background to integrate AI content into their curricula, which helps equip students with the necessary AI knowledge and skills for the AI-driven future.

Keywords: teaching and learning · Microsoft Learn for Educators (MSLE) · artificial intelligence (AI)

1 Introduction

Artificial Intelligence (AI) training and programming have ceased to be exclusive domains limited to information science students in modern higher education (see, e.g., (Li et al. 2021), and the survey of Ng et al. (2023)), and even in K-12 and early childhood education (see the surveys of Wang and Lester (2023) and Su et al. (2023), respectively). With the advent of the Internet, their significance has become closely tied to a country's economic and technological advancement. Consequently, nearly all universities now offer comprehensive Information Technology (IT) courses that are essential knowledge for both educators and students. This drive for self-improvement is bound to impact existing curricula across various departments.

Simultaneously, the pursuit of innovative and creative teaching approaches has become a prevailing trend, urging academic staff to adapt their instructional methods (see (Li et al. 2022) for an example). An evident illustration of this can be seen in Taiwan's promotion of university social responsibility (USR) programs (Executive Yuan, R.O.C. 2019). The primary goal of this initiative is to enable students to gain knowledge beyond the confines of traditional classroom learning through engagement in social outreach services. These programs aim to enhance students' social participation skills.

Taking into account the teaching demands arising from the above two developments, this paper presents a revised implementation process and experience spanning from the academic year 2019–2020 to the present. Additionally, it proposes a course plan for the upcoming academic term, focusing on the mandatory computer graphics course offered during the first year of architectural studies at a university in Taiwan.

2 Relevant Studies

IT Industry's Investment in School Teaching. Microsoft is considered a tech company that is deeply invested in education partnerships. The report by Fullan et al. (2020) combined Microsoft's three major phases that must be considered for future learning in response to the pandemic: (1) disruption, (2) transition, and (3) reimagining. This report specifically mentions that this is some sort of hybrid model where digital technologies must be employed to expand, accelerate, and connect learners to learning with a conscious focus on global competencies and academic standards. These new skills and attributes, combined with learning processes that combine the best of distance and on-campus learning with digital engagement, drive deep learning and bring equity and learning for all to the fore. No one knows what post-pandemic education will look like, but technology use is expected to continue due to its increased use during the COVID-19 pandemic (Escola 2021; Ng et al. 2021; Lee et al. 2022). Aires et al. (2021) highlight the fact that technology is a complement, not a substitute, as it is not a substitute for schools from the perspective of socialization and civic behavioral learning. They mentioned that Microsoft 365 can be a bridge between the technological culture of students and teachers, as it offers a wide variety of tools and possibilities, while also fostering conditions for better communication with the educational community, especially with the parents of students. Meyer and Billionniere (2021) even directly compared the three teaching resources of Amazon Academy, Microsoft Learn for Educators (Azure Fundamentals), and IBM Skills Academy in an in-depth comparison. Microsoft Azure Training organizes solid course material into distinct units. Each unit allows for lectures and exercises on self-contained topics. Instructors found the navigation and progress indicators to be confusing at times. However, the advantage of the Microsoft Azure textbook is that it solidly covers basic cloud computing and business systems. Focusing on the business side of Azure is important for IT professionals to understand business considerations. Relatively speaking, Amazon and IBM are less involved in the learning system of the education sector.

Intervention in Social Responsibility Issues. Data-driven artificial intelligence provides important basic information processing functions, which in educational applications, based on the combination of two approaches, represents a natural path forward.

In contrast, education usually focuses on the gradual development of theoretical and conceptual structures in a particular field. Holmes and Tuomi (2022) made the above points, so in the context of education, the development of artificial intelligence may be more constructively seen as the joint development of human and artificial cognition. In this context, the question arises: what is the relationship between the content of human development and the university? Rahbarianyazd (2017) mentioned that universities face the responsibility to address global challenges. They must engage with communities, promote diversity, and develop ethical research frameworks. As catalysts for change, they collaborate with stakeholders to find sustainable solutions. Dignum (2021) also emphasized that the digital age is a time for reinvention and creativity. Capacity building must embrace these skills alongside technological expertise. This shows that the traditional separation between humanities, arts, and social sciences is not suitable for the needs of the digital age. Today, the world's complexities demand collective efforts. Universities hold the key to a brighter future by equipping students to create a better world.

3 Background of This Study

Course Requirements for a Department with Non-IT Backgrounds. This study was performed in a private university near Taiwan's Hsinchu Science Park, where experiential learning about AI is one of the major characteristics of teaching that all faculties need to work on (Chung Hua University, n.d.). As a pilot study, computer graphics in the architecture and urban planning programme was chosen to be the department's introductory course for AI experiential learning. An architectural computer graphics course typically focuses on the application of computer graphics techniques and tools specifically tailored to the needs of architects and designers. The content of such a course may vary depending on the institution offering it, but some common and expected topics include 3D modelling, CAD software, BIM (Building Information Modelling), animation, digital fabrication, virtual reality (VR), and augmented reality (AR), rendering engines and plugins, presentation techniques, architectural photography, and post-processing. Throughout the course, students may be required to work on individual and group projects, creating 3D models, visualizations, and presentations to demonstrate their understanding and skills in architectural computer graphics. These courses aim to equip aspiring architects and designers with the necessary tools and knowledge to leverage computer graphics effectively in their profession.

AI-SCE Courses for Teachers to Apply for Offering. To effectively promote AI courses for non-IT professionals, the school's AI+ Experience Centre has designed a set of progressive courses called the AI-SCE (Starter, Candidate, and Expert) programmes for interested teachers to apply for on their own (Chung Hua University 2023). Teachers who apply can add AI elements to any of their existing curricula. Qualified and executed curricula can be included in the scoring of the teaching assessment. There are 3 levels of specifications: (1) Interdisciplinary AI Professional Course, (2) Microsoft MSLE Courseware, and (3) AI Module Application Courseware.

An Interdisciplinary AI Professional Course is an initial level of introductory material that can be taught in only 10 to 15 h of an existing course. The course content shall include courses on smart manufacturing, smart building, smart marketing, smart retail, smart finance, robotics applications, and/or any AI algorithm. One of the 10 h of the course is at an Entry Level, and completion of the course report and student feedback is sufficient. The Advanced Level requires a minimum of 15 h. In addition to the entry-level course content, it will be required to design and execute an enterprise AI application project that focuses on solving practical problems and building cross-domain AI models using non-simulated data. The number of these courses awarded and funded per semester is based on 30 courses for Entry Level and 10 courses for Advanced Level. The maximum number of courses that can be awarded per semester is two per teacher.

Microsoft MSLE Courseware is unfunded but can be included in a teacher's evaluation of teaching. The applicant must be a teacher who has previously passed a school-run Seed Teacher Training Camp. Teachers have access to the latest Microsoft AI-related materials, course structure examples, and teaching guides to incorporate into their courses. During the course, students can use Microsoft desktop tools to practice application problems in their professional fields and obtain certified skills and certificates to add value to their resumes. Microsoft provides six types of teaching materials for schools: (1) Azure AI Fundamentals (AI-900), (2) Azure Fundamentals (AZ-900), (3) Power Platform Fundamentals (PL-900), (4) Designing and Implementing a Microsoft Azure AI Solution (AI-102), (5) Microsoft Azure Administrator (AZ-104), (6) Analyzing Data with Microsoft Power BI (DA-100).

AI Module Application Courseware is applied by teachers with programming experience, so the number of courses offered in each semester is not fixed. In particular, the AI+ Experience Centre will provide adequate assistance to meet the practical needs of students. In particular, it is necessary to pay for the materials required by the students during their practical work. The Centre has prepared 11 types of established teaching materials to assist teachers in conducting classes, namely, (1) Intelligent Voice Experience, (2) Intelligent Building Inspection, (3) Customized Visual Recognition Models, (4) Azure IoT Services, (5) Intelligent Sales Forecast, (6) Smart Chat Booking, (7) Wisdom Recognition of Instant Images, (8) Intelligent monitoring of mechanical equipment, (9) Wisdom Detection of CNC abnormality, (10) Intelligent Prediction for Vending Machines, and (11) Wisdom Analysis of Learning Classification Model.

Table 1 below summarizes the three types of AI-SCE courses described above.

Table 1. Comparison summary of the three types of AI-SCE Courses.

	AI Module Application Courseware	Microsoft MSLE Courseware	Interdisciplinary AI Professional Course
Hourly Requirements	A minimum of 3-h practical course	No specific requirements	10 to 15 h of AI course content
Curriculum	Teaching materials provided by the AI+ Experience Centre	Ready-to-use PowerPoints, videos	Teachers prepare their materials
Assessment	Teacher Customization	Module quizzes and practice certification exam	Teacher Customization
Project-Based Learning	Yes	Teacher Customization	No
Certification Exam	Azure Data Scientist	Azure Fundamentals	No
Prerequisite(s)	Programming experience is required	No	No
Accreditation	Yes	Yes	No

4 Implementation into Teaching and Learning

Based on the strong push for AI teaching and learning in the school as an administrative measure, the dedicated teachers in each department approached it with a mindset of experimentation. Most teachers from non-IT backgrounds start with Interdisciplinary AI Professional Courses. The authors of this paper have responded to the school's request to select Computer Graphics as the only AI course in the Department of Architecture and Urban Planning since the academic year 2019.

4.1 Computer Graphics Courses for Employment Needs

The first year of computer graphics is devoted to teaching the most basic applications. The department focuses on internationally recognized Autodesk products, with other lightweight 3D applications such as SketchUp (SU) as a secondary focus. In the first semester, students are required to complete AutoCAD (ACAD) architectural drawings and basic SketchUp 3D modeling exercises. The exposure to BIM applications (Revit, RVT) is immediate in the second semester, which takes up at least two-thirds of the time. The remaining time will be used for additional illustrative rendering or animation. Table 2 shows a 36-week curriculum plan for a full academic year.

The above learning progression has been implemented for more than 10 years. With this experience, students will be able to cope with future architectural design requirements. Intensive quizzes will be arranged in the first semester to stimulate students' learning rhythm. The number of tests will be significantly reduced in the second semester. The compact time is mainly reserved for RVT case practice. In a two-hour class every

Table 2. Arrangement of Year-1's computer graphics course before AI learning.

Week	1st Semester	2nd Semester
1	Introduction and general settings	SU Plugin - architectural rendering
2	ACAD general drafting	SU Plugin - spatial animation
3	ACAD general drafting	SU Plugin - case study
4	ACAD architectural plans	Quiz
5	Quiz	RVT Model - general concepts
6	ACAD architectural drafting	RVT Model - building structures
7	ACAD architectural drafting	RVT Model - rendering and animation
8	ACAD site planning	RVT Model - site formation
9	Quiz	Quiz
10	ACAD 3D concepts	RVT Families - geometric forms
11	ACAD 3D concepts	RVT Families - furniture design
12	Quiz	RVT Families - furniture design
13	SU modeling - geometric forms	RVT Families - review
14	SU modeling - building forms	RVT Families - structure openings
15	Quiz	RVT Families - structure openings
16	SU modeling - building forms	RVT Families - building envelope
17	SU modeling - contour map	RVT Families - case study
18	Quiz	Overall review

week, the intermittent lectures take up about half an hour, and the remaining one-and-a-half hours are spent assisting the students with their exercises. Frankly speaking, both teachers and students think that there is not enough time.

4.2 Interdisciplinary AI Professional Course

As mentioned earlier, the Entry Level of AI instruction requires a 10-h allocation, and this is in conjunction with the student's practice assignments. The AI handouts were created by the teachers themselves, with the materials being rearranged almost exclusively from public information. The AI+ Experience Centre only provides general one, the content of which is not much different from that on the Internet. Therefore, teachers mainly look for materials that are close to their specializations.

Figure 1 shows the changes made by the authors to the courses over the academic years. The first academic year (2019–2020) took quite a long time to prepare handouts and look for online AI services, especially free ones. This course is about drawing, so graphic manipulation is a priority. The plan for the year was to set aside about half an hour for each lesson. The contents were planned as follows: (1.1) Machine learning; (1.2) Big data analytics; (2.1) AI in graphic design; (2.2) Quick Draw, AutoDraw, and Scrying Pen; (2.3) PaintsChainer; (2.4) Deep Art; (2.5) Pix2Pix; (2.6) SketchAR; (2.7)

Academic Year	Semester	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
before AI learning	1st	intro	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	quiz	SU	SU	quiz	SU	SU	quiz
	2nd	SU	SU	SU	quiz	RVT	RVT	RVT	RVT	quiz	RVT	RVT	RVT	RVT	review	RVT	RVT	RVT	RVT	review
2019-2020	1st	intro	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	quiz	SU	SU	quiz	SU	SU	quiz	
	2nd	SU	SU	SU	quiz	RVT	RVT	RVT	RVT	quiz	RVT	RVT	RVT	RVT	review	RVT	RVT	RVT	RVT	review
2020-2021	1st	intro	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	quiz	SU	SU	quiz	SU	SU	quiz	
	2nd	SU	SU	SU	quiz	RVT	RVT	RVT	RVT	quiz	RVT	SS 2	RVT	RVT	review	RVT	RVT	RVT	RVT	review
2021-2022	1st	intro	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	ACAD	quiz	via Teams	via Teams	via Teams	via Teams	via Teams	via Teams
	2nd	SU	SU	SU	quiz	SU	SU	SU	quiz	RVT	RVT	RVT	RVT	quiz	RVT	RVT	RVT	RVT	RVT	review
2022-2023	1st	intro	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	ACAD	quiz	Azure	Azure	Azure	Azure	Azure	SU	SU	SU	SU	quiz
	2nd	SU	SU	SU	quiz	MSLE	MSLE	MSLE	MSLE	MSLE	SS 1	SS 2	SS 3	SS 4	RVT	RVT	RVT	RVT	RVT	quiz
2023-2024	1st	intro	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	ACAD	quiz	SU	SU	SU	quiz	Azure	Azure	Azure	Azure	Azure	
	2nd	SU	SU	SU	quiz	RVT	RVT	RVT	RVT	quiz	RVT	RVT	RVT	RVT	RVT	RVT	RVT	RVT	RVT	quiz

Fig. 1. Arrangements and changes in AI content intervention in computer graphics courses by academic years. The Interdisciplinary AI Professional Course ran for 5 semesters.

Application review; (3.1) Smart buildings; (3.2) Smart homes; (3.3) Green design; (4.1) Natural ventilation; (4.2) Daylighting; (4.3) Thermal comfort; (5.1/SS1) Space Syntax’s depthmapX; (5.2/SS2) Visibility Graphic Analysis; (5.3/SS3) Axial Map Analysis; (5.4/SS4) Agent Analysis; (5.5/SS5) Gate Counts; and (5.6/SS6) overall review.

Nos. 1.1 to 1.2 is an introductory note with information from the literature and infographics of major IT companies. Nos. 2.1 to 2.6 are the most popular generative drawing tools from the Google search system. AI issues related to the built environment (Nos. 3.1 to 3.3) are introduced before the end of the semester.

The second semester continues with an introduction to physical environmental factors in architecture (Nos. 4.1 to 4.3). Nos. 5.1 to 5.6 introduce analysis software that uses algorithms to calculate the visual range in space. This course uses Space Syntax’s depthmapX developed by the School of Architecture, UCL.

Students must submit another study report before the end of the semester, the content of which is mainly based on AI-generated drawings, supplemented by a summary of each person’s experience. This open-ended personal point of view has also become the basis for why AI elements are added to the course.

The AI contents of the first two academic years were more or less the same, except for the second semester of the following year, which was affected by the epidemic and was replaced by the MSLE courseware. It was only because of the experience of borrowing from Microsoft lectures that we found them to be full of content. The program was discontinued after one more semester, making a total of five semesters.

4.3 Microsoft MSLE Courseware

According to the careful comparison of Meyer and Billionaires (2021), Microsoft’s service benefits from solid AI learning material. Teachers only need to report to the AI+ Experience Center to participate in teaching, and then they can get a Microsoft education account. The latest teaching materials of the six categories mentioned above can be downloaded in full. The Chinese and English bilingual versions of this course are downloaded and distributed on the student’s learning platform.

During the academic year 2022–2023 (Fig. 1), except for teaching Space Syntax, the rest of the time will be based on Microsoft’s Azure services. And of the six categories of material, only the most basic Azure AI Fundamentals (AI-900) was selected to be introduced; after all, teachers’ AI knowledge is also not yet mature. There are four main modules in the AI-900: (1) Introduction to AI; (2) Machine Learning; (3) Computer Vision; and (4) Natural Language Processing. The USR project, in which the authors are involved, has been specifically addressed in the teaching of AI this academic year. To effectively fulfill the work requirements of the program, only interested students are solicited to participate. In the end, six students participated in the whole process. The work of the sub-project was to visit the townships to collect pedestrian activity flows, which were then modeled and evaluated using Space Syntax. The AI+ Experience Centre does not specify how much Azure AI practice is to be done, and in terms of volume, it is completely indulgent in letting the teachers dictate what is to be learned. Microsoft only requires students to take a pre- and post-course questionnaire, but this part still counts the company’s internal data. Figure 2 illustrates the course registration on MSLE Portal.

The screenshot shows the 'Courses' page on the MSLE Portal. The page title is 'Courses | 2 courses found'. A search bar is located at the top right. On the left side, there is a navigation menu with options: Home, Profile, Course Intake Form, Courses (selected), Voucher, Practice Key, Training events, Program Documents, Microsoft Learn, and Learning Download Center. The main content area displays a table with the following data:

Course Name	Course Start Date	Course End Date	Certipoint Exam Group Number	Course Number	Pre-course Survey	Post-course Survey	Student Pre-Survey Completion %	Student Post-Survey Completion %	Status
[Link]	09/08/2022	01/05/2023		80510 1A	Link to survey	Link to survey			Completed
[Link]	02/16/2023	06/15/2023		80510 2A	Link to survey	Link to survey			Completed

Fig. 2. General MSLE Portal pages for course registration.

4.4 Adaptation of Teaching

For teachers with non-IT backgrounds, applying for two types of AI courses should be the limit. Although the school never interferes with the teachers’ operation of the courses, the acceptance of AI by the students often affects the whole teaching process. This is something that teachers care about. As far as architecture is concerned, the most important subject is architectural design. Therefore, the task of the computer graphics course is to help students fulfill their ability to express themselves in the design process of building projects. If the intervention of AI content is going to seriously slow down the students’ practice of their professional skills, the authors/teachers are going to give up on it.

Introduction to AI Tools. After running the Interdisciplinary AI Professional Course for five consecutive semesters, the authors decided to terminate the application for the next session (2nd semester of the academic year 2021–2022). In addition to bringing students’ drawing training back to a tabulated progression, teachers themselves can also

rethink. It is important to note that students are not averse to AI or related learning. Students are also not sure of the progress that should be made in architectural drawing and are willing to accept new learning as long as it is new. This is not catchy whether it's for the construction industry, algorithms, or AI-generated graphics. The students are kind of happy with it. However, that's the problem. Once you have a satisfactory tool, what's the next step? Traditional hand-drawn drawings, or today's digital computer graphics, were designed to satisfy the expressive approach to architectural design. The objectives of their behavior are both clear and specific. What the last five semesters have brought to the students is just to get to know a lot of new things, and then there is nothing more. The authors can understand the problem of such intuition from the changes in the fineness of the graphical representations in the students' final AI Reports. To put it bluntly, this end-of-term report on AI insights has no idea what it's all about.

Introduction to USR Outreach Activities. After a semester hiatus, the authors have applied again for recognition as an AI course. After all, that's what AI seed teachers are for! For the reboot academic year, the courses introduced a small selection of work items from the USR program as topics for AI learning. In the classroom instead, using real-life examples from our department or other schools to illustrate what USR is trying to do. And again, Azure AI is interspersed throughout the storytelling. The authors' simple idea is to borrow living examples in the hope that students will take enough interest to understand the issues around them. If we want to solve the problems in these cases, can the tools of AI help? This past academic year, we switched to using MSLE materials. With Microsoft's online learning resources, students do not necessarily have to do exercises during class and can learn more about social issues with the time saved.

Planning for the Next Period. From the distribution of the AI lecture time in Fig. 1, it can be noticed that it was changed from the beginning of a scattered introduction to intensive learning. The advantage of this is that students can concentrate on a new subject, while the disadvantage is obvious: the classroom practice of computer graphics is drastically compressed. Thanks to the current development of online self-learning videos, such as the YouTube channel with a variety of computer graphics instructional videos. Diligent students can learn a lot of skills from them. Therefore, in the past academic year, the authors have been looking for suitable short films from the Internet to make up for the lack of time. Of course, we will also give a brief explanation in class first, and then ask students to take the time to practice after class.

Returning to the timetable, the main AI content will be distributed at the end of the first semester and the beginning of the second semester. Considering that the end of the semester is the busiest time of the year for architectural design classes, students need to learn how to use ACAD in the first semester. At the beginning of the second semester, the Chinese New Year had just been celebrated, and spring break was soon to follow. Therefore, we plan to use the winter holidays as a buffer for AI learning and think about the true meaning of USR in the meantime. A hybrid of USR and AI can also be done at this time. Slower students can also use the whole semester to complete their USR work. The last two weeks of the second semester will be used to summarize and review the learning process of the whole semester. Figure 3 shows the planned schedule of time slots.

Academic Year	Semester	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
2022-2023	1st	intro	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	ACAD	quiz										
	MSLE										Azure	Azure	Azure	Azure	Azure		SU	SU	SU	quiz
	2nd	SU	SU	SU	quiz							RVT	RVT	RVT	quiz	RVT	RVT	RVT	quiz	
2023-2024	1st	intro	ACAD	ACAD	ACAD	quiz	ACAD	ACAD	ACAD	quiz	SU	SU	SU	quiz						
	MSLE						Azure	Azure	SS 1	SS 2	SS 3	SS 4				Azure	Azure	Azure	Azure	
	2nd					SU	SU	SU	quiz	RVT	RVT	RVT	quiz	RVT	RVT	RVT	quiz			
	MSLE	SS 1	SS 2	SS 3	SS 4														Azure	Azure

Fig. 3. Difference in the timing of MSLE classes between two academic years.

5 Conclusion

This study presents a case of how Microsoft Learn for Educators (MSLE) can be used for integrating AI learning into an architectural computer graphics course at a Taiwan University, which details a revised implementation process and experience spanning from the academic year 2019–2020 to the present, and proposes a course plan for the upcoming academic term. This study highlights the importance of providing students with AI learning experiences and knowledge to prepare them for the demands of the rapidly evolving technological landscape. By introducing AI content into the curriculum, students gain exposure to the latest tools, technologies, and applications that are increasingly relevant across various fields, including architecture. It also emphasized the significance of maintaining a student-centered approach to AI teaching and learning, and the importance of continuously updating and improving AI course materials to keep pace with the rapidly evolving field of AI.

Future Work. There is still a lack of relevant discussions on the evaluation of MSLE courseware and courses. In the future, we will interview teachers and students to gain insights into the advantages and limitations of integrating AI learning into higher education using MSLE. It would also be meaningful to generalize the study with courses in other universities and departments.

References

Aires, A.P., Escola, J., Lopes, N.: Microsoft 365: a teaching and learning resource during the pandemic. In: INTED2021 Proceedings, pp. 3647–3652. IATED (2021)

Chung Hua University (n.d.): AI+ 體驗與資訊教育[AI+ experience and information education]. <https://www.chu.edu.tw/p/412-1000-490.php?Lang=zh-tw>. Accessed 30 June 2023

Chung Hua University: AI+ 華大AI-SCE 課程申請[AI+ Chung Hua University AI-SCE programme application] (2023). <https://ai.chu.edu.tw/p/412-1072-2408.php?Lang=zh-tw>. Accessed 30 June 2023

Dignum, V.: The role and challenges of education for responsible AI. *Lond. Rev. Educ.* **19**(1), 1–11 (2021)

Escola, J.J.J.: Comunicação educativa: perspectivas e desafios com a COVID-19. *Educação Real.* **45** (2021)

Executive Yuan, R.O.C.: University social responsibility (2019). <https://english.ey.gov.tw/News3/9E5540D592A5FECD/0bf4aaf5-7332-44d7-bc6c-55d9a4e2a590>. Accessed 30 June 2023

- Fullan, M., Quinn, J., Drummy, M., Gardner, M.: Education reimaged: the future of learning. A collaborative position paper between new pedagogies of deep learning and Microsoft education (2020)
- Holmes, W., Tuomi, I.: State of the art and practice in AI in education. *Eur. J. Educ.* **57**(4), 542–570 (2022)
- Lee, L.K., Chui, K.T., Fung, Y.C.: The study of the dilemma on the control of COVID-19 spread and face-to-face learning and its trade-off solutions. In: *Digital Innovation for Healthcare in COVID-19 Pandemic*, pp. 137–151. Academic Press (2022)
- Li, S.F., Ng, K.K., Lee, L.K.: A study on the application of AI experiential learning in the architecture and design courses of a Taiwan university. In: Li, R., Cheung, S.K.S., Iwasaki, C., Kwok, L.F., Kageto, M. (eds.) *ICBL 2021. LNCS*, vol. 12830, pp. 103–115. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-80504-3_9
- Li, S.F., Ng, K.K., Lee, L.K.: Progress tracking and online learning assisted by social media tools: a case study of teaching architectural design. In: *2022 International Symposium on Educational Technology (ISET)*, pp. 113–117. IEEE (2022)
- Meyer Jr., L., Billionniere, E.: AWS academy vs Microsoft learn for educators vs IBM skills academy: the educators choice. In: *Society for Information Technology & Teacher Education International Conference*, pp. 528–534 (2021). Association for the Advancement of Computing in Education (AACE)
- Ng, R.Y.K., Ng, K.K., Lam, R.Y.S., Lee, L.K.: A study of vocational and professional education and training (VPET) students' online learning experience during the outbreak of pandemic. In: *2021 International Symposium on Educational Technology (ISET)*, pp. 89–95. IEEE (2021)
- Ng, D.T.K., Lee, M., Tan, R.J.Y., Hu, X., Downie, J.S., Chu, S.K.W.: A review of AI teaching and learning from 2000 to 2020. *Educ. Inf. Technol.* **28**(7), 8445–8501 (2023)
- Rahbarianyazd, R.: Sustainability in historic urban environments: effect of gentrification in the process of sustainable urban revitalization. *J. Contemp. Urban Aff.* **1**(1), 1–9 (2017)
- Su, J., Ng, D.T.K., Chu, S.K.W.: Artificial intelligence (AI) literacy in early childhood education: The challenges and opportunities. *Comput. Educ. Artif. Intell.* **4**, 100124 (2023)
- Wang, N., Lester, J.: AI education for K-12: a survey. In: Wang, N., Rebolledo-Mendez, G., Dimitrova, V., Matsuda, N., Santos, O.C. (eds.) *AIED 2023. CCIS*, vol. 1831, pp. 44–49. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-36336-8_6



A Review of Data Augmentation and Data Generation Using Artificial Intelligence in Education

Kwok Tai Chui¹(✉), Lap-Kei Lee¹, Fu Lee Wang¹, Simon K. S. Cheung¹, and Leung Pun Wong²

¹ Hong Kong Metropolitan University, Hong Kong, China
{jktchui, lklee, pwang, kscheung}@hkmu.edu.hk

² Tung Wah College, Hong Kong, China
lpwong@twc.edu.hk

Abstract. Technological advancement of artificial intelligence enhances the quality of education. Sufficient and high-quality data are important elements in building accurate artificial intelligence models. The real-world scenarios are often small-scale, attributable to various reasons, such as the high cost of data collection, low occurrence of events, protected data, privacy issues, and ethical issues. Data augmentation and generation provide additional training data to facilitate machine learning model construction. This paper reviewed 202 publications to analyze the methodologies, results, and applications of data synthesis approaches for educational research in 2010–2023 (up to August 2023). Basic characteristics were studied, including the number of annual publications, subject areas of publications, top ten publishers, and a word cloud of keywords. It was followed by an in-depth discussion of the top ten most cited publications. Several open challenges and suggestions for potential future research directions were considered.

Keywords: artificial intelligence · data augmentation · data generation · small-scale dataset · smart education

1 Introduction

In 2015, member states of the United Nations (UN) signed the 2030 Agenda for Sustainable Development (UN General Assembly 2015), where the goals and targets related to quality education and lifelong learning opportunities were specified in the SDG4-Education 2030 Sustainable Development Agenda (UNESCO 2015). Different tools were studied to find the best path to achieve the agenda. Artificial intelligence is being recognized as the most promising tool, provided that regulatory frameworks are implemented to address policy debates (UNESCO 2021).

Artificial intelligence is increasingly being adopted in educational research, such as assessment, grading, personalized teaching and learning, educational administration, and performance prediction (Chiu et al. 2023; Su et al. 2023). The COVID-19 pandemic in late 2019 and ChatGPT in late 2022 have driven such development, where

people witnessed the benefits of artificial intelligence via online learning and generative artificial intelligence. Attention is drawn to the first step of artificial intelligence model construction in education, which requires sufficient, high-quality data. Whether a dataset is sufficient for model training depends on various factors, such as (i) Complexity of model formulations: typical formulations include clustering, classification, anomaly detection, and prediction of future events. Clustering requires grouping of unlabeled data into multiple clusters, classification manages multi-class detection, anomaly detection tackles one-class classification, and prediction of future events overcomes a series of unseen data; (ii) Type of machine learning algorithms: Deep learning algorithms are usually working with large-scale datasets, whereas traditional machine learning algorithms handle small-scale datasets; and (iii) Data domain knowledge: The understanding of the data domain affects the difficulty and complexity of feature extraction. It helps to define relevant features to characterize the data. Without good domain knowledge, deep learning algorithms are promising solutions for automatic feature extraction.

In both situations of insufficient data and sufficient data, data augmentation (Castiglioni et al. 2021) and data generation (Jabbar et al. 2021) provide synthetic data as extra training data, to enhance the performance of machine learning models (the former situation is more significant). This paper aims to survey artificial intelligence approaches for data augmentation and data generation in educational research. To our understanding, this research is the first of its kind.

The organization of the paper is as follows. Section 2 presents the search process for relevant articles. The following section analyzes the articles. In Sect. 4, research challenges and future research directions are illustrated. At last, a conclusion is drawn.

2 Search Process

We used Scopus Database - advanced document search (with TITLE-ABS-KEY search strategy) to search for data augmentation and data generation using artificial intelligence in education. The query string consisted: (“data augmentation” OR “data generation” OR “generative adversarial network” OR “data synthesis”) AND (“education” OR “educational research” OR “teaching” OR “teacher” OR “student” OR “blended learning” OR “online learning” OR “e-learning” OR “hybrid learning” OR “innovative learning”). The initial search returned 3369 articles. The following inclusion and exclusion criteria were used to shortlist relevant articles. In total, 202 publications (140 articles, 59 conference papers, and three book chapters) were shortlisted. Ninety of them (44.6%) are open-access publications. Figure 1 summarizes the shortlisting process.

- Including only English articles. One hundred and thirty-seven articles were excluded for other languages including, Chinese, Portuguese, Spanish, French, Turkish, German, Persian, Japanese, and Italian.
- Including document types of articles, conference papers, and book chapters. One thousand one hundred and sixty-three articles were excluded, including reviews, conference reviews, books, short surveys, notes, editorials, letters, and erratum.
- Including source types of journals, conference proceedings, book series, and trade journals. Twenty books were excluded.

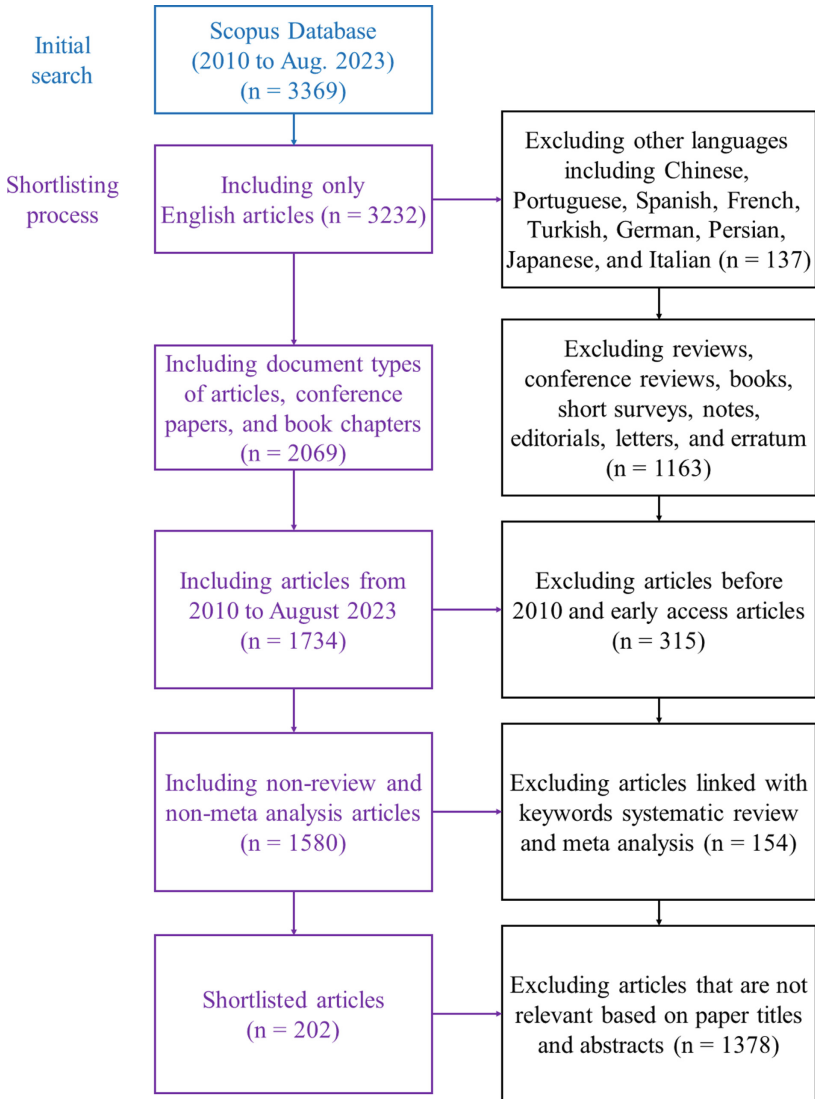


Fig. 1. Shortlisting process of relevant articles of data augmentation and data generation using artificial intelligence in education.

- Including articles from 2010 to August 2023. Three hundred and fifteen articles were excluded.
- Excluding keywords that were related to reviews: systematic review and meta-analysis. One hundred and fifty-four articles were excluded.
- Carefully read the paper titles and abstracts to confirm the publications' relevance. One major reason for the significant exclusion of existing works was that the keywords related to teachers and students were often used in the research of generative models.

3 Analysis of Related Works

This section analyzes the 202 shortlisted articles. It starts with the presentation of characteristics, followed by the discussions of algorithms and applications in the top ten most cited publications.

3.1 Characteristics

Figure 2 shows the number of publications between 2010 and August 2023. Considering several increments and decrements in the number of publications in these periods and small numbers, the research on data synthesis using artificial intelligence in education is receiving fluctuating attention. Beginning in 2020 ($n = 29$), the average annual growth rate was 79.5% towards 2022 ($n = 55$). The projected number of publications in 2023 will be similar to that in 2022. One of the reasons for the recent trend was the pandemic: the social acceptability of artificial intelligence in education increased. The authors look forward to studying the research area in the post-pandemic era.

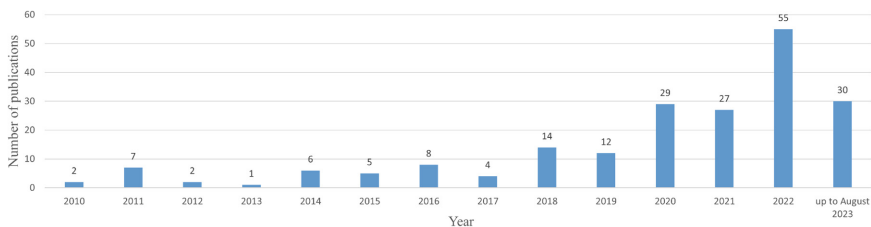


Fig. 2. Number of publications of related works between 2010 and August 2023.

To investigate the subject areas of the related works, Fig. 3 summarizes the number of publications in 24 subject areas. The top five subject areas are computer science (25.3%), social sciences (22.9%), engineering (12%), mathematics (10.1%), and medicine (4.53%), contributing to 74.9% of the population. Considering the first, third, and fourth leading subject areas, these suggested existing works focused more on the algorithms than educational perspectives (second subject area).

The top ten source titles of journals and conferences are summarized in Table 1. These journals covered 22.9% ($n = 32$), and the conferences covered 42.3% ($n = 25$) of the population. Yet, some journals and conferences published identical numbers of publications in education; the list in Table 1 was based on the priorities in the Scopus Database. Among twenty source titles, reflecting by the name of journals/conferences, five journals and three conferences were the main themes of education.

To study the main themes of related works, Fig. 4 presents the word cloud of the top 160 keywords (limit of Scopus Database). In total, there are 973 counts. Among all keywords, 28.8% ($n = 280$) is artificial intelligence-related, 20.0% ($n = 195$) is specific targets-related, 17.6% ($n = 171$) is education-related, and 9.56% ($n = 93$) is data-related. It is worth noting that some researchers may not repeat the terms utilized in the paper title and abstract as keywords. However, research studies revealed the benefits of keyword

repetition in the paper title (Bornmann et al. 2018) and abstract (Sohrabi and Iraj 2017), such as increasing the number of citations, download rates, and search result priorities. This is important to promote the adoption of artificial intelligence in education.

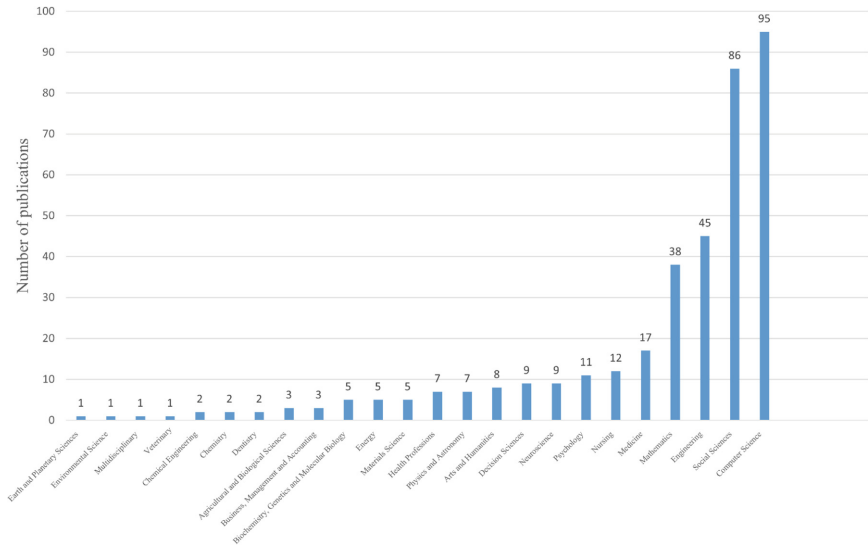


Fig. 3. Number of publications of related works in 24 subject areas.

3.2 Top Ten Most Cited Articles

To investigate the algorithms and applications, we carry out in-depth discussions on the top ten most cited articles in 2019–2023 (reflecting the latest development). Table 2 summarizes the number of citations (based on Scopus and Google Scholar), applications, datasets, algorithms (related to data augmentation and data generation), and results. Various applications were considered, such as student performance prediction, academic emotion inference, individualized educational game creation, and teaching the concepts of generative models. Some existing works performed preliminary analysis so that ablation studies were omitted to investigate the effectiveness of the data augmentation or data generation algorithms.

Table 1. Top ten source titles of 114 journals and 44 conferences.

Journals	Number of publications	Conferences	Number of publications
Computational Intelligence and Neuroscience	7	Lecture Notes in Computer Science Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics	13
South African Journal of Education	5	Communications in Computer and Information Science	2
IEEE Access	3	Lecture Notes in Electrical Engineering	2
International Journal of Learning Teaching and Educational Research	3	Proceedings of The International Conference on E Learning Icel	2
International Journal of Nursing Terminologies and Classifications	3	2014 8th Malaysian Software Engineering Conference MYSEC 2014	1
Nurse Education Today	3	2020 International Conference on Multimedia Analysis and Pattern Recognition MAPR 2020	1
Computers and Electrical Engineering	2	2021 6th International Conference on Image Vision and Computing ICIVC 2021	1
Education Sciences	2	2023 11th International Conference on Information and Education Technology ICIET 2023	1

(continued)

Table 2. Top ten most cited articles in related works.

Work	Number of citations	Applications	Datasets	Algorithms	Results
Chui et al. (2020)	59/76	Predicting university students' performance	Secondary school students' performance (Portugal)	Integrating cGAN, infoGAN, and acGAN	Sensitivity enhancements by 2.75–5.66%
Bian et al. (2019)	30/41	Academic emotion inference	Video clips and images of university students	An increment of images by 800% using image scaling, rotation, translation, dropout, noise insertion, elementwise multiplication, perspective transformation, and cropping and padding	Accuracy enhancement by 55.6%
Park et al. (2019)	24/36	Creation of adaptive educational games for personalized learning	The ENGAGE game-based learning environment supported by Unity	Creation of 5000 levels using deep convolutional generative adversarial networks	4993 out of 5000 solvable levels
Guo et al. (2020)	19/26	Bias-based optimization for prediction of graduate students' employment	Sixty-four majors in a Chinese university (graduated in 2017)	Generating samples using generative adversarial networks to balance classes	Accuracy enhancement by 0.812%
Dev & Patnaik (2020)	18/39	A face recognition-based student attendance system	A collection of images from individual students	Denosing images via generative adversarial networks	An ablation study was not carried out

(continued)

Table 2. (continued)

Work	Number of citations	Applications	Datasets	Algorithms	Results
Waheed et al. (2021)	14/19	Predicting at-risk university students	Open University (UK) learning analytics dataset	Balancing classes using sequential conditional generative adversarial networks	Precision enhancement by 4.40–6.17%
Ali et al. (2021a)	14/28	Teaching students about generative models	Generative stories (text)	Four generative tools: AI Duet, Sketch RNN, AI News Anchor, and This Person Does Not Exist	Generator correctness of 78.6–100% and discriminator correctness of 42.9–78.3%
Häusler et al. (2020)	11/13	Education and training via content-compliant training data	Customer relationship management dataset	Generating 500,000 data	An ablation study was not carried out
Ali et al. (2021b)	7/9	Teaching secondary school students the basic theories and applications of generative adversarial networks	Generative images	Four generative tools: AI Duet, Sketch RNN, AI News Anchor, and This Person Does Not Exist	Generator correctness of 88% and discriminator correctness of 60%
Baruah and Baruah (2021)	6/6	Predicting student performance	Student performance dataset	Bootstrap method	An ablation study was not carried out

4 Open Challenges

According to the United Nations Educational, Scientific and Cultural Organization (UNESCO), in May 2023, less than 10% of the respondents (schools and universities) were ready (developed guidelines and policies) to utilize generative artificial intelligence, e.g., chatbots (UNESCO 2023). Many open challenges remain unsolved before promoting artificial intelligence (e.g., data augmentation and data generation) in education:

- **Lack of technical knowledge:** Artificial intelligence is one of the most difficult educational technologies (Alam et al. 2021), particularly for teachers and students not equipped with a solid foundation in artificial intelligence. More efforts are required to deliver the content to the audience, e.g., two of the top ten most cited publications in Table 2 (Ali et al. 2021a, b). In addition, advanced artificial intelligence algorithms may not be feasible to the public.
- **Datasets availability:** The pre-requisite of data augmentation and data generation is the availability of datasets, either open access or granted upon reasonable requests. The concept of open data has been promoted for many years; however, the results are not satisfactory. UNdata, Data.gov, Data.world, Open Data China, and European Data Portal are examples of organizations that contribute open data repositories to this initiative. Many concerns are raised, such as data privacy, data security, ethics, transparency, non-profit usage, artificial intelligence regulations, and resources for data repositories (Kras et al. 2020).
- **Artificial intelligence investment:** Initial investment is required to set the basic architecture, platform, and services of artificial intelligence-based educational technologies (Nemorin et al. 2023). This brings up a challenge to increase the expenditure of facilities and equipment. It may require support from the government if the school/university is government-funded. For self-financing school/university, it may lead to an increase in tuition fees.

5 Conclusion

This paper reviewed 202 publications in the research area of data augmentation and data generation using artificial intelligence in education. The study of basic characteristics revealed (i) the trends of the number of annual publications, where increased during the pandemic; (ii) the distribution of subject areas of publications, with the top five areas, i.e., computer science, social sciences, engineering, mathematics, and medicine, covering 75% of areas; (iii) top ten journal and conference publication sources; and (iv) a word cloud study of top 160 keywords, mainly artificial intelligence-related, specific targets-related, education-related, and data-related. Attention was also drawn to the top ten most cited publications, discussing applications, datasets, algorithms, and results. At last, three key open challenges were presented.

Funding. The work described in this paper was fully supported by a grant from Hong Kong Metropolitan University (RIF/2021/05).

References

- Alam, A.: Possibilities and apprehensions in the landscape of artificial intelligence in education. In: 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA), pp. 1–8. IEEE, Nagpur, India (2021)
- Ali, S., DiPaola, D., Lee, I., Hong, J., Breazeal, C.: Exploring generative models with middle school students. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1–13. ACM, Yokohama, Japan (2021)

- Ali, S., DiPaola, D., Breazeal, C.: What are GANs?: introducing generative adversarial networks to middle school students. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 15472–15479. AAAI, Online (2021)
- Baruah, A.J., Baruah, S.: Data augmentation and deep neuro-fuzzy network for student performance prediction with MapReduce framework. *Int. J. Autom. Comput.* **18**, 981–992 (2021)
- Bian, C., Zhang, Y., Yang, F., Bi, W., Lu, W.: Spontaneous facial expression database for academic emotion inference in online learning. *IET Comput. Vision* **13**(3), 329–337 (2019)
- Bormmann, L., Haunschild, R., Hug, S.E.: Visualizing the context of citations referencing papers published by Eugene Garfield: a new type of keyword co-occurrence analysis. *Scientometrics* **114**, 427–437 (2018)
- Castiglioni, I., et al.: AI applications to medical images: from machine learning to deep learning. *Physica Med.* **83**, 9–24 (2021)
- Chiu, T.K., Xia, Q.C., Zhou, X., Chai, C.S., Cheng, M.: Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Comput. Educ. Artific. Intell.* **4**, 100118 (2023)
- Chui, K.T., Liu, R.W., Zhao, M., De Pablos, P.O.: Predicting students' performance with school and family tutoring using generative adversarial network-based deep support vector machine. *IEEE Access* **8**, 86745–86752 (2020)
- Dev, S., Patnaik, T.: Student attendance system using face recognition. In: 2020 International Conference on Smart Electronics and Communication (ICOSEC), pp. 90–96. IEEE, Trichy, India (2020)
- Guo, T., et al.: Graduate employment prediction with bias. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 670–677. AAAI, New York, USA (2020)
- Häusler, R., Staegemann, D., Volk, M., Bosse, S., Bekel, C., Turowski, K.: Generating content-compliant training data in big data education. In: Proceedings of the 12th International Conference on Computer Supported Education, pp. 104–110. Springer, Prague, Czech Republic (2020). <https://doi.org/10.5220/0009513801040110>
- Jabbar, A., Li, X., Omar, B.: A survey on generative adversarial networks: variants, applications, and training. *ACM Comput. Surv.* **54**(8), 1–49 (2021)
- Kras, A., Celi, L.A., Miller, J.B.: Accelerating ophthalmic AI research: the role of an open access data repository. *Curr. Opin. Ophthalmol.* **31**(5), 337 (2020)
- Nemorin, S., Vlachidis, A., Ayerakwa, H.M., Andriotis, P.: AI hyped? A horizon scan of discourse on artificial intelligence in education (AIED) and development. *Learn. Media Technol.* **48**(1), 38–51 (2023)
- Park, K., Mott, B.W., Min, W., Boyer, K.E., Wiebe, E.N., Lester, J.C.: Generating educational game levels with multistep deep convolutional generative adversarial networks. In: 2019 IEEE Conference on Games (CoG), pp. 1–8. IEEE, London, UK (2019)
- Sohrabi, B., Iraj, H.: The effect of keyword repetition in abstract and keyword frequency per journal in predicting citation counts. *Scientometrics* **110**, 243–251 (2017)
- Su, J., Ng, D.T.K., Chu, S.K.W.: Artificial intelligence (AI) literacy in early childhood education: the challenges and opportunities. *Comput. Educ. Artific. Intell.* **4**, 100124 (2023)
- UN General Assembly: Transforming our world: the 2030 Agenda for Sustainable Development, A/RES/70/1 (2015). <https://documents-dds-ny.un.org/doc/UNDOC/GEN/N15/291/89/PDF/N1529189.pdf?OpenElement>
- UNESCO. SDG4-Education 2030, Incheon Declaration (ID) and Framework for Action. For the Implementation of Sustainable Development Goal 4, Ensure Inclusive and Equitable Quality Education and Promote Lifelong Learning Opportunities for All, ED-2016/WS/28 (2015). <https://unesdoc.unesco.org/ark:/48223/pf0000245656>
- UNESCO. AI and education: guidance for policy-makers (2021). <https://doi.org/10.54675/PCS/P7350>. <https://unesdoc.unesco.org/ark:/48223/pf0000376709>

- UNESCO. Generative Artificial Intelligence in education: what are the opportunities and challenges? (2023). <https://www.unesco.org/en/articles/generative-artificial-intelligence-education-what-are-opportunities-and-challenges>
- Waheed, H., et al.: Balancing sequential data to predict students at-risk using adversarial networks. *Comput. Electr. Eng.* **93**, 107274 (2021)



The Role of ChatGPT in Higher Education: Some Reflections from Public Administration Students

Aleksander Aristovnik^(✉), Lan Umek, Nejc Brezovar, Damijana Keržič,
and Dejan Ravšelj

Faculty of Public Administration, University of Ljubljana, Ljubljana, Slovenia
{aleksander.aristovnik, lan.umek, nejc.brezovar, damijana.kerzic,
dejan.ravselj}@fu.uni-lj.si

Abstract. Launched in November 2022, ChatGPT, an artificial intelligence-based chatbot, can generate coherent, informative, human-like responses and have natural language conversations with users. In the higher education sector, ChatGPT has been met with both excitement and scepticism. Therefore, the aim of the paper is to investigate the perception of ChatGPT by students of public administration. Findings show that ChatGPT seems to be an attractive platform for public administration students. Students have generally positive perceptions of ChatGPT in relation to different study issues, suggesting the further impact of using ChatGPT in concrete public administration study issues is recommended to enhance learning from students' perspective. However, students feel less effective in developing and improving critical thinking, creativity and personal communication skills. Accordingly, the findings of this paper add to the existing scientific knowledge and facilitate evidence-based policy-making, which is crucial for higher education in the future.

Keywords: ChatGPT · higher education · public administration · students' perceptions · empirical analysis

1 Introduction

An artificial intelligence-based chatbot, ChatGPT, was launched in November 2022 and can generate cohesive and informative human-like responses and engage in natural language conversations with users. In higher education, ChatGPT has sparked both excitement and scepticism. On the one hand, it is believed that ChatGPT will inevitably become an integral part of higher education, and therefore some education practitioners recommend involving students with such tools to enrich the learning experience rather than prohibition (Farrokhnia et al. 2023; Rudolph et al. 2023). In this context, it is believed that Chat GPT can help students develop different skills, such as reading, writing, information analysis, critical thinking and problem-solving, generating practice problems, etc., being relevant in public administration practice and research (AlAfnan et al. 2023; Rahman & Watanobe 2023; Rudolph et al. 2023). On the other hand, it is

believed that ChatGPT will bring multiple challenges, such as copyright issues, bias, fairness, excessive reliance by students, data privacy and security, etc. (Farrokhnia et al. 2023; Lo 2023; Rahman & Watanobe 2023). Although some existing studies provide initial insights into the potentials and challenges of Chat GPT, they do this from a teacher rather than a student perspective.

Early on, after ChatGPT was offered to the public, Wagholikar et al. (2023) found very important results in their online survey study that, due to its ease of use, awareness of ChatGPT did not differ between genders or between fields of study. While the ChatGPT is seen as a disruptive element in education, innovative teachers will know how to use it to their advantage and the benefit of their students (Rudolph et al. 2023), as it can be used in education to enhance teaching and learning for both teachers and students. For students, ChatGPT could act as a tutor and support their learning, for example, by answering questions and summarising information (Lo 2023; Flores Limo et al. 2023; Rahman & Watanobe 2023). ChatGPT tutoring is a form of one-to-one tutoring where advanced technology allows instant interaction, so students do not have to wait for a tutor or teacher to answer their questions. The results of the survey reported by Flores Limo et al. (2023) showed that almost 70% of the participants used ChatGPT as a tutor for their learning. The results also showed that satisfaction after the use of ChatGPT, self-confidence and the motivation to learn were significantly related to the length of the individual tutoring session. Even more interesting is that learners can argue with ChatGPT about the explanations, solutions and other suggestions given, and other suggestions (Rahman & Watanobe 2023), developing competence in argumentation and critical thinking.

ChatGPT can also facilitate collaboration by generating different scenarios for students to solve problems together and provide personalized feedback and guidelines (Lo 2023; Rudolph et al. 2023). ChatGPT can also help with writing skills by providing grammar corrections and suggestions (Lo 2023) or even assist the student by providing them with the structure of the essay (AlAfnan et al. 2023). Similarly, the strengths of ChatGPT, such as the generation of plausible answers and the provision of personalised and real-time answers, were identified by Farrokhnia et al. (2023). They saw opportunities in the increased accessibility of information and the facilitation of personalised and complex learning.

The study conducted by Ali et al. (2023) between English language learners investigated motivation during learning. The research shows that the involvement of ChatGPT in language learning was generally a motivating factor, especially in reading and writing, while it wasn't perceived in listening and speaking. The authors conclude that the new AI general-purpose conversation chatbot is recommended to be a part of English language learning, as it motivates learners to learn autonomously and should not be avoided in the learning process. Similar is reported by Sotelo Muñoz et al. (2023), who have found that ChatGPT is supposed to increase students' motivation and engagement in the learning process. With motivations to read and write more than other learning processes in English language learning.

Shoufan (2023) presented a survey conducted among senior computer science students to assess the potential of ChatGPT and its impact on teaching and learning. In the first phase, they evaluated ChatGPT in an open-ended question after having used

it in one of the study activities. In the second phase, they had a longer period of using ChatGPT for prepared activities, and after that, they answered a questionnaire. Students rated ChatGPT as easy to use, motivating them to work. The students were impressed by ChatGPT's capabilities and emphasised the human-like interface with clear and structured answers, and expressed positive attitudes. However, they also pointed out that the answers are sometimes not precise, concerned about manipulations and malicious use and exposed that ChatGPT could not substitute human intelligence.

However, to fully understand the impact of ChatGPT on higher education, it is crucial to understand students' experiences and perceptions of it. Namely, students' perceptions are highly relevant for educators as they can significantly impact their motivation, engagement and academic achievement. Therefore, the paper aims to examine public administration students' perception of ChatGPT. Specifically, it is focused on students' perceptions of ChatGPT's capabilities, study issues within its context, and its ability to facilitate skills development and enhance outcomes. The rest of the paper is structured as follows. The next section presents the data and research methods used in the study. The final section provides a conclusion, summarising the main findings and implications.

2 Data and Research Methods

The data were a part of the ChatGPT Student Survey entitled "Students' Perception of ChatGPT". The online questionnaire was designed in English and consisted of closed-ended questions covering different aspects relevant to the opportunities and challenges of ChatGPT for students. The developed online questionnaire consisted of selected socio-demographic characteristics of the survey participants and several important aspects/sections relevant to the opportunities and challenges of ChatGPT for public administration students: 1) students' awareness, focusing on how they are familiar with ChatGPT; 2) students' knowledge, concentrating on the technical aspects, recent advancements, potential implications and limitations of ChatGPT; 3) ChatGPT capabilities, such as its ability to understand and respond in human language, improve various industries, and perform certain tasks better than humans; 4) students' concerns, focusing on the potential negative consequences of ChatGPT, such as job loss, misuse, the spread of misinformation, and invasion of privacy; 5) students' satisfaction, focusing on their experience interacting with ChatGPT; study issues from a student and teacher perspective; 6) implications for the labour market; 7) regulatory and ethical considerations; and 8) usage potential of ChatGPT in improving learning and working. Individual aspects were measured on a 5-point Likert rating scale ranging from 1 (e.g., strongly disagree) to 5 (e.g., strongly agree).

The survey was intended for all higher education students enrolled in public administration study programmes in Slovenia. The data collection process took place between May and July 2023 and was facilitated through various communication systems and social media to obtain a representative sample of public administration students. The empirical analysis relies on descriptive statistics, which unveil the mean values of perceptions among the students, encompassing various aspects of ChatGPT. These aspects encompass students' perceptions of ChatGPT's capabilities, study issues within its context, and its ability to facilitate skills development and enhance outcomes. The mean

values are separately presented for each individual aspect of ChatGPT through bar plots. The arrangement of these bars in a plot is such that they are organized in descending order of the calculated means.

3 Data and Research Methods

By the end of July 2023, 157 public administration students (students from the Faculty of Public Administration, University of Ljubljana, Slovenia) had participated in the survey. Since the students were not obliged to complete the questionnaire in full, the number of respondents across questions may vary due to the missing values. The socio-demographic characteristics of the survey participants are presented in Table 1. First, most of the students (66.5%) were female and enrolled in the first level of study (81.5%). Moreover, most of the students had a general understanding of what ChatGPT was and what it could do (36.9%) and even used ChatGPT before and had a good understanding of its capabilities (35.7%).

Table 1. Socio-demographic characteristics of the survey participants.

Socio-demographic characteristics	Number (#)	Share (%)
Gender		
Male	52	33.5
Female	103	66.5
Level of study		
First	128	81.5
Second	21	13.4
Third	8	5.1
Familiarity with ChatGPT		
Never heard of ChatGPT before	3	1.9
Heard of ChatGPT, but I am not sure what it is used for	9	5.7
General understanding of what ChatGPT is and what it can do	58	36.9
Used ChatGPT before and have a good understanding of its capabilities	56	35.7
Very familiar with ChatGPT and have already used it extensively	31	19.5

Source: Own calculations based on ChatGPT Student Survey

Regarding public administration students' perceptions of the capabilities of ChatGPT, the results show that they – as shown in Fig. 1 – on average tend to agree the most with the statement that ChatGPT has a wide range of capabilities ($M = 4.1$). This is not surprising if we consider that ChatGPT can not only answer simple answers but can also be used for more complex tasks such as writing music, planning holidays, writing

summaries, job applications etc. As more than half (55.2%) of the responding students claim to be very familiar with ChatGPT and have used it extensively (19.5%) or at least have used ChatGPT before and have a good understanding of its capabilities, it is also not surprising that they have tested different ways of using ChatGPT and have come to the conclusion that it has a wide range of capabilities. Students, on average, also agree that ChatGPT can summarize and paraphrase lengthy articles ($M = 4.04$), which can be attributed to the fact that students often must summarize lengthy articles or textbook chapters to solve different tasks or to prepare themselves for certain exams. Since attending lectures is not obligatory at the faculty, the students can find such a capability very useful also in situations when they did not attend lectures and therefore did not prepare notes. ChatGPT can assist them with that. On the other hand, students least tend to agree that ChatGPT can write poems in different languages and styles ($M = 3.42$) or can write music in any genre ($M = 3.31$). One possibility for such a result is that the students did not test the beforementioned capabilities, but further research should be done before drawing any concrete conclusions.

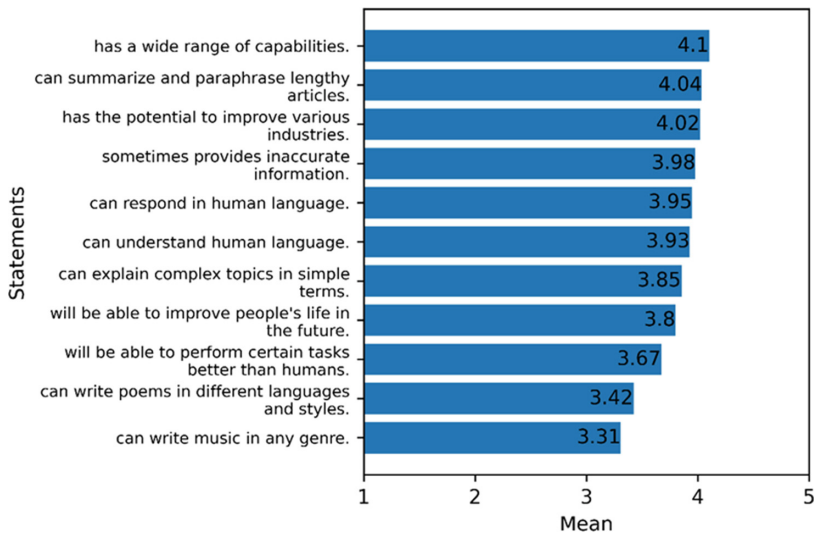


Fig. 1. Students' perceptions of the capabilities of ChatGPT. Source: Own calculations based on ChatGPT Student Survey.

Regarding students' perceptions of study issues in the context of ChatGPT, the results show that they – as shown in Fig. 2 – on average tend to agree the most with the statements that ChatGPT can help with their studies ($M = 4.15$) and can save them time for assessment ($M = 4.15$). Both perceptions are in line with the students' perceptions regarding the capabilities of ChatGPT, which include the fact that students agree that ChatGPT can summarize and paraphrase lengthy articles. Students can therefore save time for assessment and use ChatGPT as a help tool with their studies. This is not surprising, taking into account some of the significant applications of ChatGPT in education

which can be used by students, such as allowing students to ask better questions, giving straightforward responses, providing text analysis, crafting essays, boosting exam preparation, providing exact information, summarising large documents (Javaid et al. 2023). On the other hand, students out of the given possible statements, on average ($M = 3.45$), at the least perceive ChatGPT as being able to or providing personalized education. Since possibilities exist for students, such as receiving personalised feedback using ChatGPT based on their writing style to help understand and concentrate on the areas they need to improve, being provided with on-demand, live tutoring or for everything to be explained to them multiple times (Flores Limo et al. 2023; Javaid et al. 2023; Lo 2023; Rudolph et al. 2023) it is probable that the students did not use ChatGPT in such a way or that they did and were not satisfied with the level of personalization.

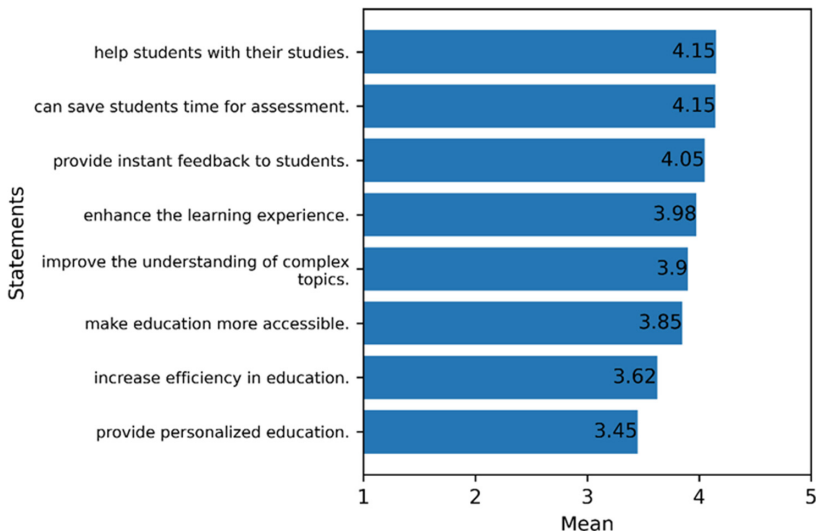


Fig. 2. Students' perceptions of study issues in the context of ChatGPT. Source: Own calculations based on ChatGPT Student Survey.

Regarding students' perceptions of the ability of ChatGPT to facilitate skills development, the results show that they – as shown in Fig. 3 – on average tend to agree the most with the statement that ChatGPT has the ability to improve students' digital communication skills ($M = 3.57$). Some studies observe that tools such as ChatGPT can potentially reduce some of the skills that students should master, including critical, creative and collaborative thinking skills and creative writing skills (Shidiq 2023, p. 354). Why students chose to perceive improving digital communication skills as a top ability of ChatGPT regarding facilitating skills development can be attributed to the fact that digital communication with another person electronically through tools such as text messaging, messaging applications, e-mail, social media sites (such as Instagram, Facebook, Twitter, Snapchat etc.) is very close to them and they use them often on a day-to-day or better minute-to-minute basis. It is only natural that students have therefore tested or used ChatGPT as a tool to help with their digital communication skills, which, as they

perceived it, improved through the use of ChatGPT. On a similar note of “communication skills”, a completely different result emerged regarding students’ perception of ChatGPT’s ability to facilitate and improve students’ face-to-face communication skills ($M = 2.19$). Some authors see ChatGPT as a tool eroding face-to-face communication, empathy, and emotional intelligence, which could potentially fuel a fragmented society where AI language models supplant the human touch in our daily interactions (Ray & Das 2023, p. 1). Their views could become a reality if ChatGPT will not offer satisfying results in using or improving skills in face-to-face communications. In one podcast, an interview was done with ChatGPT where the algorithm was asked if we (humans) still need to learn how to communicate better now that ChatGPT exists, to which ChatGPT replied that while ChatGPT can provide information and generate responses, it cannot replace the human interaction and emotional intelligence that comes with face-to-face communication (Abrahams 2023). As it would seem for now, considering the presented students’ perceptions, there is little correlation between ChatGPT and the perception of it as a tool to facilitate and improve students’ face-to-face communication skills.

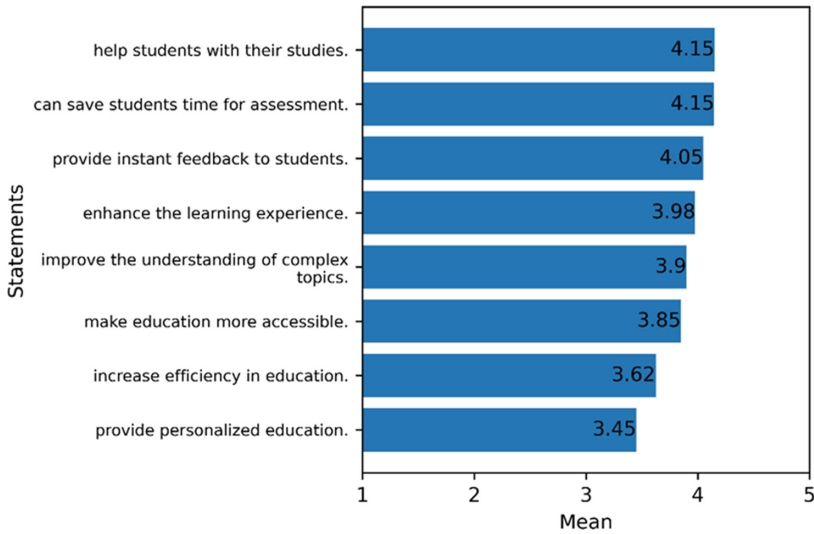


Fig. 3. Students’ perceptions of the ability of ChatGPT to facilitate skills development. Source: Own calculations based on ChatGPT Student Survey.

Students were also asked about their perceptions of ChatGPT to improve outcomes. Results (presented in Fig. 4) show that they, on average, tend to agree the most with the following statements regarding ChatGPT 1. improves students’ satisfaction ($M = 3.62$), 2. increases the chance of completing the study ($M = 3.61$), and 3. increases the chance of completing co-curricular activities ($M = 3.61$). Students, therefore, perceive ChatGPT as a positive and helpful tool to be used in their studies, preferably because ChatGPT can save them time, which would otherwise be used for certain simple tasks such as summarizing and paragraphing lengthy articles (Fig. 1) or certain assessments (Fig. 2). By saving time for simpler tasks, students can focus on more demanding aspects

of their studies and still have time to complete co-curricular activities and therefore have a better chance of completing their studies (in time). Some other studies implied similar improvement outcomes of ChatGPT, such as improving the performance of students on exams, which is especially true for low-performing students and those who suffer under time constraints (Choi et al. 2023, p. 12). The improvement of outcomes in certain areas is perceived by students as lacking. Such areas of possible improvements of outcomes are especially 1. reducing under-skilling ($M = 2.97$), 2. resolving skills obsolescence ($M = 2.96$) and 3. improving students' employability ($M = 2.93$). Students, therefore, do not perceive ChatGPT as a tool for improving their skills or something that would help with their employability. Why students, on average, disagree with ChatGPT as a tool to be used is a topic for further research. As for now, we can conclude that ChatGPT can be used to improve certain student skills, such as language and writing skills, to boost their exam preparation and prepare them for interviews by improving their communication skills (Javaid et al. 2023, p. 8 and 10).

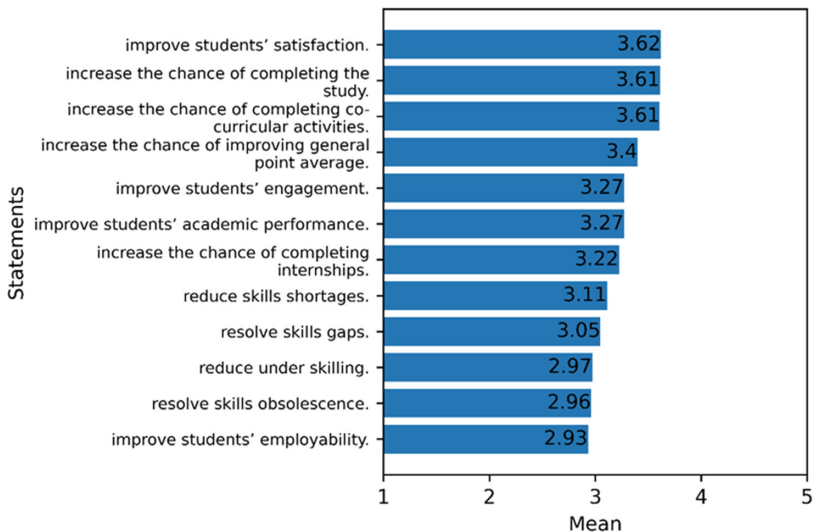


Fig. 4. Students' perceptions of ChatGPT to improve outcomes. Source: Own calculations based on ChatGPT Student Survey.

4 Conclusion

ChatGPT is perceived by public administration students as a tool that can enhance learning experiences and not as a disruptive tool. Most (92.1%) of the students have already used ChatGPT or are at least familiar with what it is and how it works. Students perceive ChatGPT as having a wide range of capabilities, such as summarizing and paragraphing lengthy articles, while they do not perceive ChatGPT as much as a tool for writing poems or music. This is understandable if ChatGPT was used for study purposes or is,

in their perception, connected to study issues, while public administration studies are rarely related to writing poems or music. Students also perceive ChatGPT as a tool that can help them with their studies and save them time. Students generally have positive perceptions of ChatGPT in relation to different study issues, which implies that further impact of the use of ChatGPT in concrete public administration study issues is recommended to improve the learning process from the students' perspective. Students also see ChatGPT as a tool that can facilitate the development and improvement of certain skills, such as digital communication, foreign language skills and information literacy. At the same time, students perceive that it is not able to facilitate to the same extent the development and improvement of skills such as critical thinking, creativity and face-to-face communication skills, which it seems can be attributed to ChatGPT's lack of human interaction and emotional intelligence. Students, in general, agree that ChatGPT is able to improve certain outcomes, such as student satisfaction, increase chances of completing the study and co-curriculum activities, compared to reducing under skilling, resolving skill obsolescence or improving students' employability where students perceive ChatGPT as having a lesser impact. On average, public administration students also perceive ChatGPT as having a wide range of capabilities and, in the context of certain study issues, wider impact (Figs. 1 and 2) compared to ChatGPT abilities to facilitate skill development or improve outcomes, where students on average perceive these abilities as not so wide (Figs. 3 and 4). These findings have several practical implications within the higher education landscape. First, given that students perceive ChatGPT as a tool that can enhance learning experiences and help with their studies, higher education institutions should consider integrating ChatGPT into curricular and co-curricular activities. Second, while ChatGPT is seen as a valuable tool for enhancing digital communication, foreign language skills, and information literacy, higher education institutions should also recognize its limitations in fostering skills such as critical thinking, creativity, and face-to-face communication. Finally, since ChatGPT is also perceived as a tool to improve student satisfaction, higher education institutions should leverage ChatGPT to create a more engaging and efficient learning environment. These practical implications suggest that a balanced approach, combining artificial intelligence tools like ChatGPT with traditional teaching methods, is crucial for the holistic development of higher education. Nevertheless, further research is possible in the field of the reasoning behind the public administration students' perceptions of ChatGPT capabilities and impact.

Acknowledgements. The authors acknowledge the financial support from the Slovenian Research Agency (research core funding no. P5-0093, project no. Z5-4569 and project no. BI-US/22-24-083). The preliminary version of the paper has been presented at the EGPA 2023 Conference "Steering the European Union through poly-crises storms: The Role of Public Administration".

References

- Abrahams, M.: Quick thinks: AI has entered the Chat—"Conversation" with ChatGPT [podcast]. Insights by Stanford Business (2023). <https://www.gsb.stanford.edu/insights/quick-thinks-ai-has-entered-chat-conversation-chatgpt>

- AlAfnan, M.A., Dishari, S., Jovic, M., Lomidze, K.: ChatGPT as an educational tool: opportunities, challenges, and recommendations for communication, business writing, and composition courses. *J. Artific. Intell. Technol.* **3**, 60–68 (2023). <https://doi.org/10.37965/jait.2023.0184>
- Ali, J., Shamsan, M.A.A., Hezam, T., Mohammed, A.A.Q.: Impact of ChatGPT on learning motivation: teachers and students' voices. *J. English Stud. Arabia Felix* **2**(1), 41–49 (2023). <https://doi.org/10.56540/jesaf.v2i1.51>
- Choi, J.H., Hickman, K.E., Monahan, A., Schwarcz, D.B.: ChatGPT goes to law school. *J. Legal Educ.* (Forthcoming). <https://ssrn.com/abstract=4335905>
- Farrokhnia, M., Banihashem, S.K., Noroozi, O., Wals, A.: A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innov. Educ. Teach. Int.* (2023). <https://doi.org/10.1080/14703297.2023.2195846>
- Flores Limo, F.A., Hurtado Tiza, D.R., Arias Gonzales, J.L.: Personalized tutoring: ChatGPT as a virtual tutor for personalized learning experiences. *Social Space J.* **23**(1), 292–312 (2023). <https://socialspacejournal.eu/menu-script/index.php/ssj/article/view/176/81>
- Javaid, M., Haleem, A., Singh, R.P., Khan, S., Khan, I.H.: Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil Trans. Benchmarks Stand. Eval.* **3**(2), 100115 (2023). <https://doi.org/10.1016/j.tbench.2023.100115>
- Lo, C.K.: What is the impact of ChatGPT on education? a rapid review of the literature. *Educ. Sci.* **13**, 410 (2023). <https://doi.org/10.3390/educsci13040410>
- Rahman, M.M., Watanobe, Y.: ChatGPT for education and research: opportunities, threats, and strategies. *Appl. Sci.* **13**, 5783 (2023). <https://doi.org/10.3390/app13095783>
- Ray, P.P., Das, P.K.: ChatGPT and societal dynamics: navigating the crossroads of AI and human interaction. *AI Soc.* (2023). <https://doi.org/10.1007/s00146-023-01713-1>
- Rudolph, J., Tan, S., Tan, S.: ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *J. Appl. Learn. Teach.* **6**(1), 342–263 (2023). <https://doi.org/10.37074/jalt.2023.6.1.9>
- Shoufan, A.: Exploring students' perceptions of ChatGPT: thematic analysis and follow-up survey. *IEEE Access* **11**, 38805–38818 (2023). <https://doi.org/10.1109/ACCESS.2023.3268224>
- Shidiq, M.: 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. *Proc. Int. Conf. Educ. Soc. Hum.* **1**(1), 353–357 (2023). <https://ejournal.unuja.ac.id/index.php/icesh/article/view/5614>
- Sotelo Muñoz, S.A., Gutiérrez Gayoso, G., Arias Gonzales, J.L.: Examining the impacts of ChatGPT on student motivation and engagement. *Soc. Space J.* **23**(1), 1–27 (2023). <https://socialspacejournal.eu/menu-script/index.php/ssj/article/view/156>
- Waghlikar, S., Chandani, A., Atiq, R., Pathak, M., Waghlikar, O.: ChatGPT -Boon or Bane: a study from students perspective. In: International Conference on Advancement in Computation & Computer Technologies (InCACCT), Gharuan, India, 2023, pp. 207–212. *IEEE Explore* (2023)



Automated Computational Intelligence Based Course Timetabling Tool

Marisa Kuntasup^{1,2}, Pupong Pongcharoen², and Thatchai Thepphakorn³(✉)

- ¹ Faculty of Business Administration and Accountancy, Phitsanulok University, Phitsanulok, Thailand
- ² Centre of Operations Research and Industrial Applications (CORIA), Department of Industrial Engineering, Faculty of Engineering, Naresuan University, Phitsanulok, Thailand
- ³ Faculty of Industrial Technology, Pibulsongkram Rajabhat University, Phitsanulok, Thailand
thatchai.t@psru.ac.th

Abstract. Some universities or institutions still have consequences from the unprecedented pandemic. An effective course timetable has directly affected resource utilisation and travelling cost. Solving university course timetabling problem manually without automated program is extremely difficult, time-consuming and may require a group of experts to work for several days. A novel Automated Course Scheduling (ACOS) tool was proposed to seek feasible timetables and to optimise resource utilisation and cost. The development of the ACOS program was based on two sequential steps. The first step was to design a conceptual framework and analyse the related timetabling processes and activities. The second step was to develop the proposed graphic user interface tool that embedded computational intelligence algorithm. The performance of the ACOS program were tested using real-world problems. The statistical analysis on the experimental results suggested that the ACOS tool produced feasible timetables on all problem sizes. The proposed tool also generated the best timetables with the lowest travelling cost as well as convergence speed for all cases. In addition, the reduction on the number of used classrooms associated with the best timetables created by the proposed tool was up to 29% lower than the manual approach.

Keywords: course scheduling · resource allocation · automated timetabling · software development · metaheuristics

1 Introduction

The wide spread of COVID-19 pandemic has a globally severe impact on all business sectors including higher education (Kaliappen & Ghani 2023; Schleicher 2020). For some educational institutes or universities, there still has been an effect on the continuity on the number of the international enrolled students. Educational management becomes more important to ensure the ongoing academic activities especially teaching and learning to be carried out with limited resources and financial efficiency. The utilisation of resources (e.g., classrooms) is one of the crucial resources for educational activity (Thepphakorn et al. 2020). The effective utilisation of classrooms could benefit to the operating and maintenance costs as well as travelling or moving time.

In a large university campus, moving around campus especially for students with busy course timetable can sometimes be a hassle. It commonly happens in each semester when the students' timetables are generated without avoiding unnecessary movements between buildings in the campus. Various forms of transportation in large university campus can be either mass (e.g., minibus service) or personal (e.g., walking, bike, scooter, car). Mass transportation may not be the best choice if the bus schedule is not well designed to coordinate with the timetabling periods. When the movement time between periods is too tight and the academic buildings are scattered away, many students therefore rely on scooter or motorbike to avoid late class attending. Potential accidents, on-campus transportation service, and green campus are still the challenge issues for university management. A well-designed course timetabling could play a crucial role for dealing with those issues.

The university course timetabling problem (UCTP) can be solved either manually by academic staffs or automatically by using course timetabling software. Solving very large UCTP without any course timetabling software is extremely difficult requiring lots of efforts and data (Thepphakorn & Pongcharoen 2023). It may require days or weeks whilst the generated timetables may not be satisfied regarding hard constraints (Iqbal et al. 2021). The UCTP is known to be a Non-deterministic Polynomial (NP) hard problem, meaning that the computational time required to find the optimal solutions increases exponentially with problem size (Pongcharoen et al. 2008). To overcome these difficulties especially for solving very large-size UCTP, the preferable design task is to develop an automated course timetabling tool embedding some computational intelligence (CI) methods to shorten the time to find the optimal timetable. Many research works have been carried out with several CI algorithms to automate course timetabling program, so that the finding process can be conducted even more better and faster (Muklason et al. 2022).

Computational intelligent based automated timetabling tool have become more popular especially for large-scale real-world course scheduling (Thepphakorn et al. 2021), so that the optimal or near optimal timetables can be generated within the acceptable computational time (Lewis 2008). According to the systematic literature review, many algorithms have been applied in the automated timetabling software. The differences on objective functions considered in the literature were also reviewed. However, the automated course scheduling tool that optimises student's travelling time was rarely found.

The main objectives of this paper were to: (i) present the development of a novel automated course scheduling (ACOS) tool for solving real-world UCTP based on travelling cost; (ii) compare the performance of the proposed tool with manual approach on the feasible timetable (no hard constraints violation); and (iii) demonstrate the performances of the ACOS in terms of the solution quality (travelling cost), classroom usage, computational time, and convergence speed.

The next section of this paper briefly explains the educational course timetabling. Section 3 describes the ACOS tool including its conceptual framework and sequential diagram. Section 4 presents the experimental design and the analysis of computational results. Conclusions and future works are drawn in the last section.

2 Educational Course Timetabling

Timetabling is one of the operational problems that has been found in various domains (Siddiqui & Arshad Raza 2021) such as healthcare (e.g., nurse, physician) (Legrain et al. 2020), maintenance (Chansombat et al. 2019), manufacturing (Sooncharoen et al. 2020), sports (Januario & Urrutia 2016), transport (Liu & Ceder 2018), and education (Thepphakorn & Pongcharoen 2023). Educational timetabling is one of ongoing challenge problems due to the global constraints and local preferences. Common timetabling faced by academic institutions (e.g., high school, college, universities) are courses and exam scheduling (Abdipoor et al. 2023; Kakkar et al. 2021). The design task is to allocate a set of courses or events into the appropriate timeslots and classrooms according to a set of predefined constraints for each academic semester (Abdipoor et al. 2023). Moreover, academic timetabling can also be classified into three main problems including the high-school timetabling problem, the university course timetabling problem (UCTP), and the university examination timetabling problem (Ceschia et al. 2023). However, the UCTP has received a great worldwide interest from researchers from various academic areas (Abdipoor et al. 2023).

Hard and Soft Constraints

A course timetabling problem usually involves multiple constraints, such as lecturers and students' preferences, the variety of lectures or courses, educational institutions policies, inadequacy of equipment, facilities, and classrooms (Chen et al. 2023; Mokhtari et al. 2021). Finding the practical timetables for lecturers, students, and classrooms without any hard constraint violation must be operated whereas all users' preferences as well as educational resources' utilisation should be optimised (Thepphakorn et al. 2020).

The university course timetabling problem (UCTP) can be classified as an NP-hard problem because of the solution space and various of constraints involved (Abdipoor et al. 2023). Generally, the UCTP constraints can be classified into two groups including hard and soft constraints (Ceschia et al. 2023). Hard constraints (HC) are the most important constraints whilst soft constraints (SC) are more relaxed (Thepphakorn & Pongcharoen 2023). All hard constraints must always be satisfied to achieve feasible or practical timetables. The violation on hard constraints means that the timetable is crashed. Soft constraints are not necessary (Rezaeipanah et al. 2021) but contribute to be the objective function (Ceschia et al. 2023). The number of soft constraint violations should therefore be optimised.

3 Automated Course Scheduling (ACOS) Tool

The development of the proposed ACOS program was divided into two phases. The first phase was to design a conceptual framework and identify all relating timetabling activities using the sequential diagram. The second phase was to develop the ACOS program with embedded computational intelligence (CI) by using the TCL/TK programming language. A conceptual framework for this research can be illustrated in Fig. 1.

The ACOS tool was designed to construct the best practical course timetables, in which all hard constraints must be satisfied whilst the violation of soft constraints should

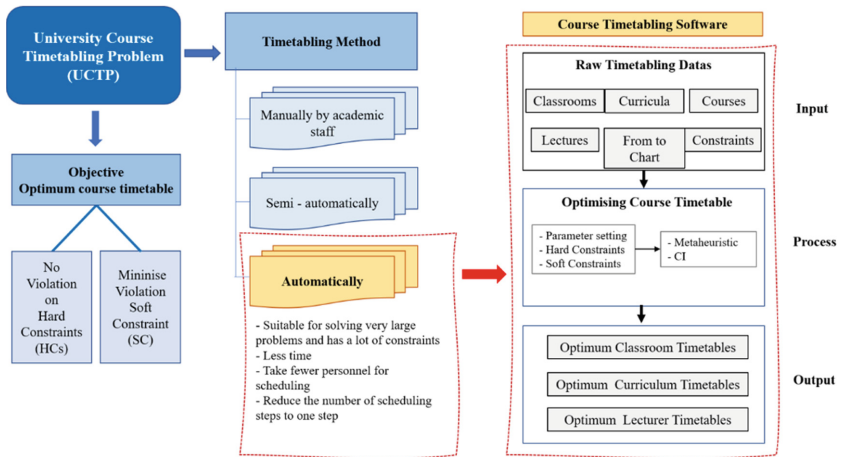


Fig. 1. The conceptual framework for the ACOS design.

be minimised. The ACOS program consists of three phases: input, process, and output. The sequential diagram for all three phases can be illustrated in Fig. 2. In the first phase, all timetabling data (e.g., curricula, courses, lecturers, classrooms, students and so on) was uploaded as selected data file into the ACOS program. Then, all parameters related with the CI and problem constraints were assigned via graphic user interface (GUI) as shown in Fig. 3.

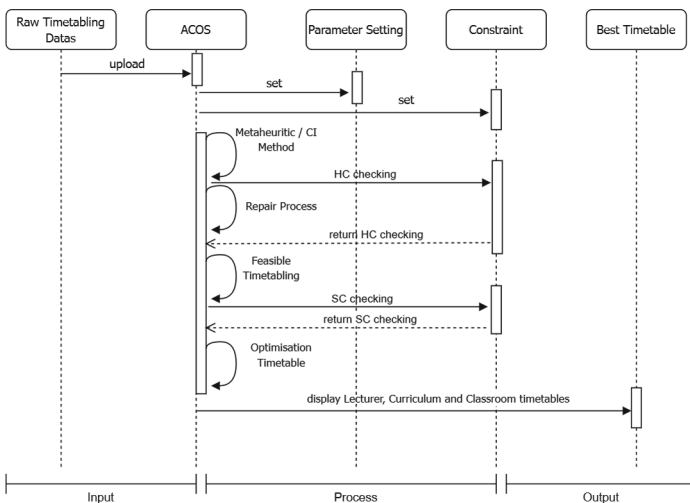


Fig. 2. The sequential diagram for the ACOS program.

Complex optimisation problems have been successfully solved by various CI algorithms, such as Genetic Algorithm (Pongcharoen et al. 2008), Artificial Bee Colony

(Pansuwan et al. 2010), Ant Colony Optimisation (Thepphakorn et al. 2014), Artificial Immune System (Thepphakorn et al. 2015), Backtracking Search Algorithm (Vitayasak et al. 2017), Cuckoo Search (Thepphakorn & Pongcharoen 2020), Teaching Learning Based Optimisation (Vitayasak & Pongcharoen 2020), Local Search (Rezaeipannah et al. 2021), Particle Swarm Optimisation (Thepphakorn et al. 2021), Simulated Annealing (Siddiqui & Arshad Raza 2021), Biogeography-based Optimisation (Sooncharoen et al. 2022), Hyper-heuristic (Muklason et al. 2022), Firefly Algorithm (Thepphakorn & Pongcharoen 2023), and Moth Flame Optimisation (Sahoo et al. 2023). MFO is one of the nature-inspired optimisation algorithms inspired by the navigating mechanism of moths in nature (Mirjalili 2015). Unfortunately, there is no report on applying MFO to solve the real-world university course timetabling problems.

During the process phase, each candidate timetable generated by using CI algorithm was checked the violation of hard constraints (HC). In this work, six hard constraints (HCs) were regularly considered: (HC₁) all lectures/laboratories (elements) required for each course must be scheduled and assigned to distinct periods; (HC₂) students and lecturers can only attend one lecture at a time; (HC₃) only one lecture can take place in a room at a given time; (HC₄) lecturers and students must be available for a lecture to be scheduled; (HC₅) all courses must be assigned into proper rooms according to their given requirements, including building location, room facilities, and room types; (HC₆) all lectures within a course requiring consecutive periods must be obeyed.

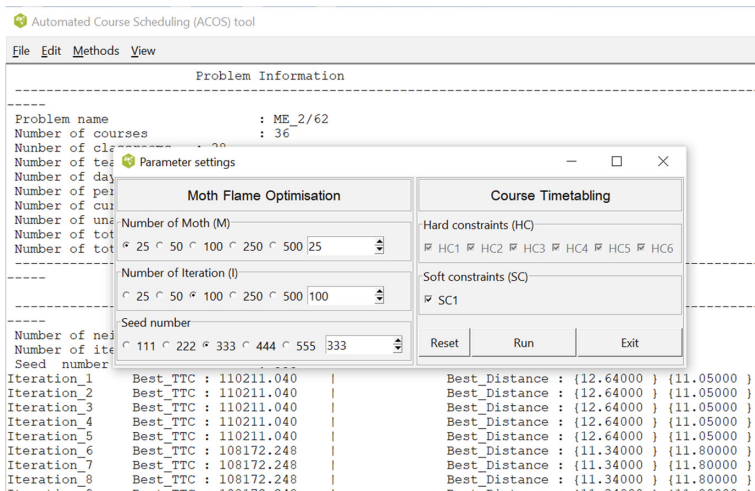


Fig. 3. An example of graphic user interface (GUI) provided in the ACOS program.

The repair process was applied to rectify the infeasible timetables, in which HCs violations were found. After rectification, the feasible timetables are then determined regarding soft constraint violations. In this work, the soft constraint (SC) was the minimisation of logistics fuel cost associated with the students travelling building to building for attending classes regarding their timetables. The minimisation of the total travelling

cost (TTC) was considered as the objective function $f(x)$ as formulated in Eq. (1).

$$\text{Minimise } f(x) = TD \times ST \times WE \times FC \times OP \quad (1)$$

TD is a total travelling distance for each student moving between buildings (kilometre per week per person). ST is the number of enrolled students (person). WE is the number of weeks per semester (week). FC is fuel consumption rate (litre/kilometre). Finally, OP is oil price (currency unit per litre).

After determining soft constraint violations, the candidate timetable with minimum travelling cost was recorded as the best timetable (solution). In the last phase, the ACOS tool displayed the timetables for all teachers, students, and classrooms regarding the best so far solution obtained. Figure 4 depicts the examples of timetables for: (i) students who enrolled in curriculum no. 1; (ii) classroom number IE509; and teacher with ID G03005.

4 Experimental Results and Analysis

A series of computational experiments were designed to: (i) compare the performance of the ACOS tool with manual approach based on the number of hard constraints violations; and (ii) demonstrate the tool performance in terms of the travelling cost, classroom usage, computational time, and convergence speed. A set of personal computers with Core i7 3.50 GHz of CPU and 6 GB of RAM was delegated to conduct all experimental runs and to determine the executional time required. Three real-case course timetabling problems obtained from Naresuan University Engineering Faculty are described in Table 1.

Table 1. Characteristics of the university course timetabling problems.

Problem Instances	No. of Courses	No. of Events	No. of Rooms	No. of Days/week	No. of Periods/day	No. of Lecturers	No. of Curricula	No. of Students	From to chart
Chemical	35	99	26	5	12	24	5	3,180	14
Mechanical	36	92	28	5	12	29	5	3,942	14
Materials	35	96	25	6	11	24	4	4,181	14

4.1 Hard Constraint Violations: ACOS Tool Versus Manual Approach

This experiment was aimed to compare the performances between the proposed tool and the manual approach in term of the number of HCs violations. The timetable with non-violation of hard constraints is feasible or practical timetable. The computational runs for each instance were repeated ten times by using different numbers of random seeds. The computational results are shown in Table 2. It can be seen that the total number of HC_1 – HC_6 violations associated with the timetables generated by the ACOS program for all problem instances were zero. The ACOS tool produced practical timetables whilst the manual approach found few violations associated with HC_2 and HC_5 .

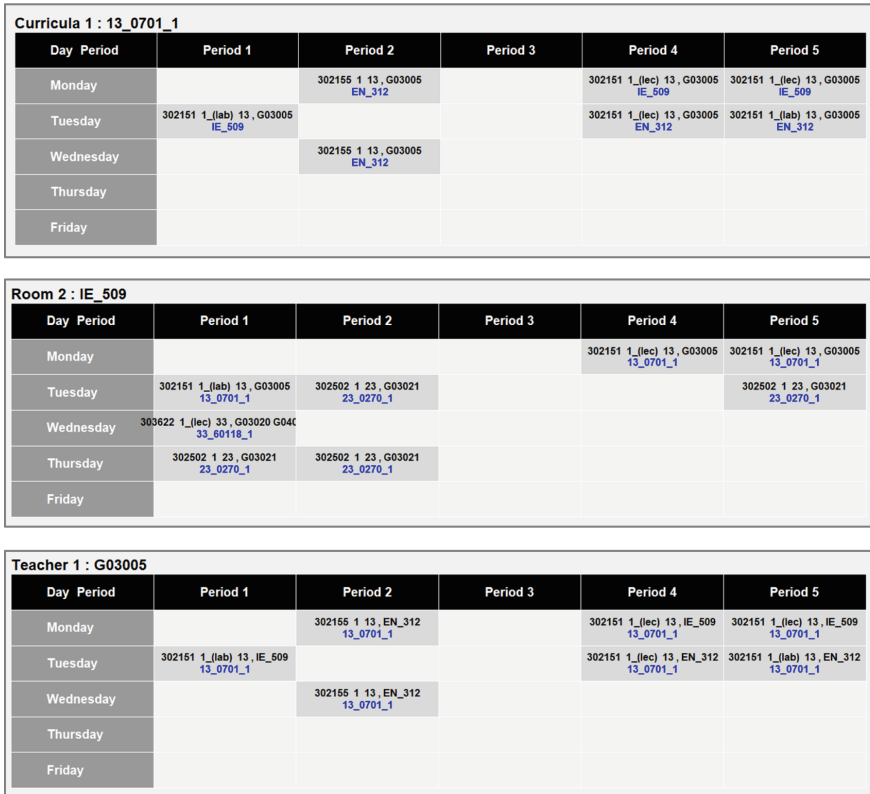


Fig. 4. Examples of teacher, student, and classroom timetables generated by ACOS tool.

Table 2. Number of violations on hard constraints (HCs).

Problem Instances	Methods	Numbers of HC violations						Total number of HCs violations
		HC ₁	HC ₂	HC ₃	HC ₄	HC ₅	HC ₆	
Chemical	ACOS	0	0	0	0	0	0	0
	Manual	0	3	0	0	2	0	5
Mechanical	ACOS	0	0	0	0	0	0	0
	Manual	0	1	0	0	4	0	5
Materials	ACOS	0	0	0	0	0	0	0
	Manual	0	2	0	0	4	0	6

4.2 The Performances of the ACOS Tool

After the created feasible timetables were guaranteed for all cases, the ACOS performance based on total travelling cost (TTC), computational time, convergence speed,

and classroom utilisation was therefore investigated. The computational runs for each instance were repeated ten times by using different numbers of random seed. The travelling costs associated with the best timetables and computational times are statistically reported in Table 3. It can be seen that the average values of TTC associated with the timetables generated for all three engineering curriculums were significantly different. However, the computational times taken to find the best so far timetables were marginally different. The convergence graphs on the average TTC values of the best so far solutions found on each iteration for each problem instance are shown in Fig. 5, 6 and 7. It was found that the convergence speeds for chemical and materials engineering were better than the mechanical engineering case.

Table 3. Computational results obtained from ACOS tool.

Problem Instances	Total travelling cost: TTC (Currency unit)				Time (Minutes)
	Minimum	Maximum	Average	Std. Dev	
Chemical	42,661.17	47,005.01	44,484.17	1,562.39	5,802.90
Mechanical	27,364.84	34,101.45	30,481.66	2,246.79	4,150.60
Materials	75,542.90	100,353.26	91,861.43	10,952.32	3,852.50

In addition, the classroom utilisations associated with the best so far timetables created by the ACOS tool for each problem instance are reported in Table 4. The classroom utilisation was based on the minimisation on the number of the used classrooms. It can be seen that the ACOS tool produced feasible timetables with high utilisation of classroom by having free rooms up to 28.93%. This means that there will be more free rooms available for other academic activities.

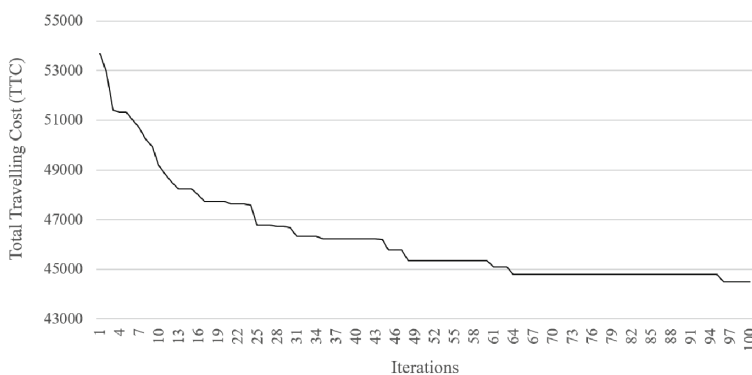


Fig. 5. Convergence speed of the best so far solutions for chemical engineering.

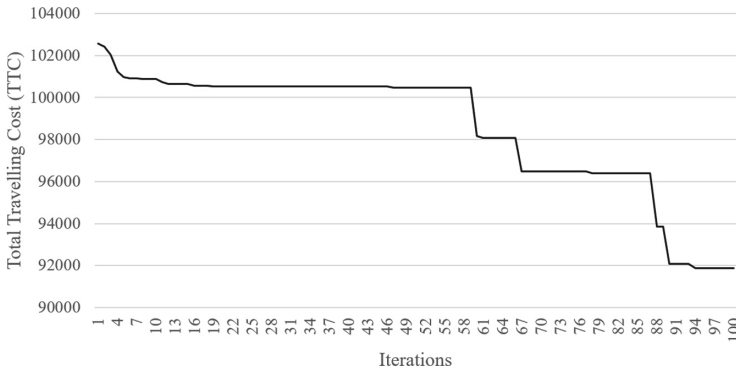


Fig. 6. Convergence speed of the best so far solutions for mechanical engineering.

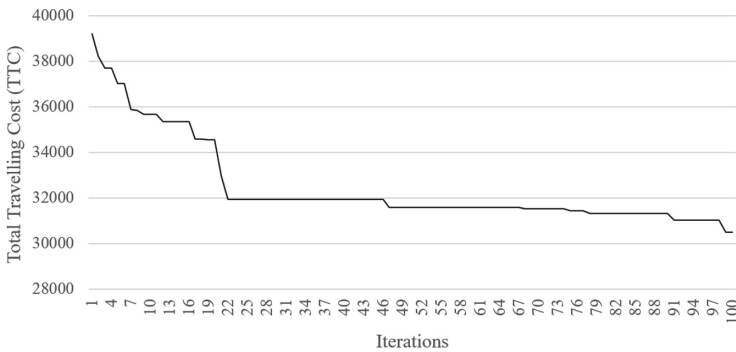


Fig. 7. Convergence speed of the best so far solutions for materials engineering.

Table 4. Classroom utilisation associated with the timetables created by the ACOS tool.

Problem Instances	No. of available Classrooms	No. of classroom used (10 replications)				Percent Room free
		Minimum	Maximum	Average	Std. Dev	
Chemical	26	20	23	21.00	1.054	19.23
Mechanical	28	18	23	19.90	1.595	28.93
Materials	25	17	22	19.90	1.595	20.40

5 Conclusions

Solving large timetabling problems manually is extremely difficult and time-consuming. Many commercial programs for course timetabling are based on a semi-automatic approach, in which constructing the optimal timetables, especially for high numbers of lecturers, students, classrooms, and given constraints cannot be achieved within the acceptable time. This paper presents the conceptual design and analysis framework and

the development of a novel Automated Course Scheduling (ACOS) tool, which embedded a computational intelligent (CI) algorithm aiming to seek the best feasible timetables with non-violation on hard constraints as well as minimise the violations on soft constraint. The conceptual framework was provided to describe the related timetabling activities before the sequential diagram presenting the main processes of the proposed program. Three real-world problem instances obtained from the collaborating university were applied for demonstrating the performances of the ACOS tool. The statistical analysis suggested that best so far timetables produced by the ACOS tool were all feasible timetables, which means that all hard constraints were satisfied. The practical timetables produced by the proposed program can yield low travelling cost and increase room utilisation for all cases. Further works may focus on the alternative CI algorithms and its performances to seek the best feasible timetables especially for very large problem sizes. Alternative soft constraints can also be considered depending on the faculty or institution policy and regulations.

Acknowledgements. This work was partially supported by the Ministry of Higher Education, Science, Research and Innovation under the grant number RGNS 63-204 and the National Research Council of Thailand and Naresuan University under the grant number N42A650329.

References

- Abdipoor, S., Yaakob, R., Goh, S.L., Abdullah, S.: Meta-heuristic approaches for the University Course Timetabling Problem. *Intell. Syst. Appl.* **19**, 200253 (2023)
- Ceschia, S., Di Gaspero, L., Schaerf, A.: Educational timetabling: problems, benchmarks, and state-of-the-art results. *Eur. J. Oper. Res.* **308**(1), 1–18 (2023)
- Chansombat, S., Musikapun, P., Pongcharoen, P., Hicks, C.: A Hybrid Discrete Bat Algorithm with Krill Herd-based advanced planning and scheduling tool for the capital goods industry. *Int. J. Prod. Res.* **57**(21), 6705–6726 (2019)
- Chen, M., Werner, F., Shokouhifar, M.: Mathematical modeling and exact optimizing of university course scheduling considering preferences of professors. *Axioms* **12**(5), 498 (2023)
- Iqbal, Z., Ilyas, R., Chan, H.Y., Ahmed, N.: Effective solution of university course timetabling using particle swarm optimizer based hyper heuristic approach. *Baghdad Sci. J.* **18**(4), 1465–1475 (2021)
- Januario, T., Urrutia, S.: A new neighborhood structure for round robin scheduling problems. *Comput. Oper. Res.* **70**, 127–139 (2016)
- Kakkar, M.K., Singla, J., Garg, N., Gupta, G., Srivastava, P., Kumar, A.: Class schedule generation using evolutionary algorithms. *J. Phys: Conf. Ser.* **1950**, 012067 (2021)
- Kaliappen, N., Ghani, A.B.A.: The impacts of covid-19 on higher education and the rise of e-learning. *Global Higher Education and the COVID-19 Pandemic: Perspectives, Challenges, and New Opportunities*, pp. 105–117 (2023)
- Legrain, A., Omer, J., Rosat, S.: An online stochastic algorithm for a dynamic nurse scheduling problem. *Eur. J. Oper. Res.* **285**(1), 196–210 (2020)
- Lewis, R.: A survey of metaheuristic-based techniques for University Timetabling problems. *OR Spectrum* **30**(1), 167–190 (2008)
- Liu, T., Ceder, A.: Integrated public transport timetable synchronization and vehicle scheduling with demand assignment: a bi-objective bi-level model using deficit function approach. *Transport. Res. Part B: Methodol.* **117**, 935–955 (2018)

- Mirjalili, S.: Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowl. Based Syst.* **89**, 228–249 (2015)
- Mokhtari, M., Vaziri Sarashk, M., Asadpour, M., Saeidi, N., Boyer, O.: Developing a model for the university course timetabling problem: a case study. *Complexity* **2021**, 9940866 (2021)
- Muklason, A., et al.: Flexible automated course timetabling system with lecturer preferences using hyper-heuristic algorithm. In: *ACM International Conference Proceeding Series*, pp. 258–262 (2022)
- Pansuwan, P., Rukwong, N., Pongcharoen, P.: Identifying optimum Artificial Bee Colony (ABC) algorithm's parameters for scheduling the manufacture and assembly of complex products. In: *Proceedings of the 2nd International Conference on Computer and Network Technology, ICCNT 2010*, pp. 339–343 (2010)
- Pongcharoen, P., Promtet, W., Yenradee, P., Hicks, C.: Stochastic Optimisation Timetabling Tool for university course scheduling. *Int. J. Prod. Econ.* **112**(2), 903–918 (2008)
- Rezaeipanah, A., Matoori, S.S., Ahmadi, G.: A hybrid algorithm for the university course timetabling problem using the improved parallel genetic algorithm and local search. *Appl. Intell.* **51**(1), 467–492 (2021)
- Sahoo, S.K., et al.: Moth flame optimization: theory, modifications, hybridizations, and applications. *Arch. Comput. Meth. Eng.* **30**(1), 391–426 (2023)
- Schleicher, A.: *The Impact of COVID-19 on Education: Insights from "Education at a Glance 2020"*. OECD Publishing (2020)
- Siddiqui, A.W., Arshad Raza, S.: A general ontological timetabling-model driven metaheuristics approach based on elite solutions. *Expert Syst. Appl.* **170**, 114268 (2021)
- Sooncharoen, S., Pongcharoen, P., Hicks, C.: Grey Wolf production scheduling for the capital goods industry. *Appl. Soft Comput.* **94**, 106480 (2020)
- Sooncharoen, S., Vitayasak, S., Pongcharoen, P., Hicks, C.: Development of a modified biogeography-based optimisation tool for solving the unequal-sized machine and multi-row configuration facility layout design problem. *ScienceAsia* **48**, 12–20 (2022)
- Thepphakorn, T., Pongcharoen, P.: Performance improvement strategies on Cuckoo Search algorithms for solving the university course timetabling problem. *Expert Syst. Appl.* **161**, 113732 (2020)
- Thepphakorn, T., Pongcharoen, P.: Modified and hybridised bi-objective firefly algorithms for university course scheduling. *Soft. Comput.* **27**(14), 9735–9772 (2023)
- Thepphakorn, T., Pongcharoen, P., Hicks, C.: An ant colony based timetabling tool. *Int. J. Prod. Econ.* **149**, 131–144 (2014)
- Thepphakorn, T., Pongcharoen, P., Hicks, C.: Modifying regeneration mutation and hybridising clonal selection for evolutionary algorithms based timetabling tool. *Math. Probl. Eng.* **2015**, 841748 (2015)
- Thepphakorn, T., Sooncharoen, S., Pongcharoen, P.: Academic operating costs optimisation using hybrid MCP SO based course timetabling tool. In: Cheung, S.K.S., et al. (eds.) *ICBL 2020*. LNCS, vol. 12218, pp. 338–350. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-51968-1_28.
- Thepphakorn, T., Sooncharoen, S., Pongcharoen, P.: Particle swarm optimisation variants and its hybridisation ratios for generating cost-effective educational course timetables. *SN Comput. Sci.* **2**(4), 264 (2021)
- Vitayasak, S., Pongcharoen, P.: Cooperative designing of machine layout using teaching learning based optimisation and its modifications. In: Luo, Y. (ed.) *CDVE 2020*. LNCS, vol. 12341, pp. 137–147. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60816-3_16
- Vitayasak, S., Pongcharoen, P., Hicks, C.: A tool for solving stochastic dynamic facility layout problems with stochastic demand using either a Genetic Algorithm or modified Backtracking Search Algorithm. *Int. J. Prod. Econ.* **190**, 146–157 (2017)



Exploring ChatGPT-Generated Assessment Scripts of Probability and Engineering Statistics from Bloom's Taxonomy

Christopher Chung Lim Kwan (✉)

Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University,
Hong Kong SAR, China
ceclkwan@polyu.edu.hk

Abstract. Subject Instructors, class teachers, and educational practitioners always devote much time to preparing assessment scripts, suggested solutions, and marking schemes such as mid-term and final examination scripts for assessing students' learning and performance as well as measuring their achievements of the subject learning outcomes. With precise prompts, ChatGPT seems to be able to work as an assistant to them in education, generating responses and deliverables in a quite structured and almost instant manner. In this paper, a ChatGPT-generated assessment script with its marking scheme and suggested solution of Probability and Engineering Statistics is explored based on Bloom's Taxonomy. It is found that the ChatGPT-generated assessment script is partially complete and one of the multiple-choice questions is incorrect. The total score of the assessment script is not consistent with that of its marking scheme. Its suggested solution to one of the questions is missing. In addition, there are many application-oriented questions but few analysis-based questions and no evaluation-based question at all in the assessment script as per Bloom's Taxonomy. Overall, ChatGPT-generated assessment scripts should further be reviewed and refined by educational practitioners to ascertain different levels of difficulty of questions which are in good alignment with the subject curriculum and the subject learning outcomes.

Keywords: Bloom's taxonomy · ChatGPT · learning outcome · probability and engineering statistics · subject curriculum

1 Introduction

Traditionally, assessment criteria and methods of a course or a subject such as assignments, mid-term tests, mini-projects, oral presentations, laboratory reports, final examinations and other forms of classroom participation are primarily devised by teaching professionals like subject instructors, class teachers, and educational practitioners to assess students' learning performance during the progression of their studies, to grade their overall performance and measure the achievement of the intended subject learning outcomes at the end of a course for quality assurance and accreditation purposes (Biggs

2003; Hong Kong Institution of Engineers 2013; Sazhin 1998). Correspondingly, appropriate teaching and learning activities are arranged in class for students to construct their knowledge and to facilitate the achievement of intended subject learning outcomes. In order to align well with the subject curriculum and the subject learning outcomes, assessment questions with different levels of difficulty are generally set according to Bloom's Taxonomy (Anderson et al 2001; Chandio et al. 2016; Krathwohl 2002; Mayer 2002). However, it is inevitable that they need to invest a lot of time in preparing assessment scripts, suggested solutions, and marking schemes such as mid-term and final examination scripts to come up with a finalized set in a correct and consistent manner (Lambert and Lines 2000).

With the recent popularity of generative artificial intelligence (GAI) applications such as ChatGPT, it is possible to use ChatGPT to pass some of the subjects related to engineering education, to work as a personal tutor at all times to help students study and to help them think about the structure of reflective and critical thinking-based writing as scaffolding activities (Nikolic et al. 2023). Johri et al. (2023) further outlined the potential impact of GAI on engineering education practices such as question creation, assignment and assessment generation related to teaching and learning. However, there is a lack of findings on the versatile use of ChatGPT in preparing assessment scripts, suggested solutions, and marking schemes, addressing the value of the present study of exploring ChatGPT-generated assessment scripts from Bloom's Taxonomy.

2 The Context of the Study

A sound knowledge of probability and engineering statistics is a prerequisite for mastering other civil engineering subjects such as construction management, geotechnics, hydraulics, structures, environmental and transportation engineering. The subject curriculum is devised to enable engineering students to master the basic concepts in probability theory and statistical analysis and is of particular relevance to applications in civil engineering. In particular, students learn fundamental concepts about probability, the probability distributions of discrete and continuous random variables, expectation and variance of random variables, jointly distributed random variables, marginal probability mass function, marginal probability density function, independence, covariance, correlation, sampling distribution, the Central Limit Theorem, estimation of parameters from samples, interval estimation, determination of sample size required, probability plots, goodness-of-fit test for distribution, linear regression and correlation analyses. Students will be able to achieve the following intended learning outcomes upon completion of the subject.

- summarize and present information effectively from data;
- design sampling plans for experiments and surveys;
- select and construct proper statistical models for engineering problems;
- apply the fundamentals of mathematics and science to formulate problems and obtain solutions in civil engineering;
- critically analyze and interpret the models formulated and solutions obtained to support the synthesis of logical and cost-effective solutions;

- integrate knowledge across different subject domains, including construction management, structures, geotechnics, hydraulics, environmental and transportation engineering;
- communicate solutions logically and lucidly through calculation, sketch, drawing and in writing

Exploring ChatGPT-generated assessment scripts of probability and engineering statistics from Bloom's Taxonomy is the focus of the present study. As per Bloom's Taxonomy (Anderson et al. 2001; Krathwohl 2002), it comprises both the knowledge dimension and the cognitive process dimension within its structure. Within the knowledge dimension, four types of knowledge such as factual, conceptual, procedural and metacognitive knowledge can be categorized. Within the cognitive process dimension, six levels of learning such as Remember, Understand, Apply, Analyze, Evaluate and Create can be set. A Taxonomy Table can be built upon these two dimensions. Based on the table, these four knowledge types and six learning levels can be distinguished and used to structure the intended subject learning outcomes, teaching and learning activities, and assessments of the course. By using the table, the intended subject learning outcomes of probability and engineering statistics are first mapped to the Taxonomy Table as shown in Table 1.

Table 1. The Taxonomy Table of the intended subject learning outcomes of probability and engineering statistics.

The Knowledge Dimension	The Cognitive Process Dimension					
	1. Remember	2. Understand	3. Apply	4. Analyze	5. Evaluate	6. Create
A. Factual Knowledge						
B. Conceptual Knowledge		X	X	X	X	X
C. Procedural Knowledge		X	X			
D. Metacognitive Knowledge						

It is noted from Table 1 that students are generally expected to demonstrate abilities to understand, apply, analyze, evaluate and create something from a low cognitive level to a high cognitive level, while grasping conceptual and procedural knowledge of probability and engineering statistics. In view of this, assessments should be designed and made to ascertain different levels of difficulty of questions that are in good alignment with the subject curriculum and the intended subject learning outcomes, as mapped to the Taxonomy Table.

The present study investigates the use of ChatGPT to prepare assessment scripts, suggested solutions, and marking schemes of probability and engineering statistics based on Bloom's Taxonomy, aiming to address the following research questions:

1. Is the ChatGPT-generated assessment script complete?
2. Are the questions of the ChatGPT-generated assessment script correct?
3. Is the total score of the ChatGPT-generated assessment script consistent with that of its marking scheme?
4. Are ChatGPT-generated solutions to the ChatGPT-generated assessment script complete?

5. What are the levels of difficulty of questions in the ChatGPT-generated assessment script based on Bloom's Taxonomy?

ChatGPT is thus employed to conduct this investigation in response to these five research questions.

3 Result

It is found that ChatGPT seems to be able to work as an assistant to them in education, generating responses and deliverables in a quite structured and almost instant manner, when it is prompted to prepare an examination paper of probability and engineering statistics for year 1 civil engineering students as shown in Fig. 1.

Figure 1 indicates that ChatGPT generates five multiple-choice questions in Section A and four descriptive questions in Section B. Sections A and B comprise 30 marks and 70 marks respectively. However, it is found that one of the multiple-choice questions is incorrect. Regarding Q.4 in Section A, the correct solution is 0 because there is not any student playing both football and basketball based on the intersection of two events in a sample space. The question is valid and the correct answer is b) 6/30 if Q.4 is revised to 24 students who play football. This result implies that the questions of the ChatGPT-generated assessment script are sometimes incorrect.

Secondly, ChatGPT is prompted to prepare a marking scheme for the examination paper. It generates the marking scheme in a quite structured and almost instant manner again as shown in Fig. 2. It is found that each question in Section A is worth 2 marks and there are only five multiple-choice questions to add up with a total of 10 marks. However, the allocated mark indicated in Section A of both the examination paper and the marking scheme is 30. Similarly, there are only four descriptive questions to add up with a total of 21 marks in Section B of the marking scheme but the allocated mark indicated in Section B of both the examination paper and the marking scheme is 70 instead. This implies that the total score of the ChatGPT-generated assessment script is not consistent with that of its marking scheme.

It is also found that the ChatGPT-generated assessment script and its marking scheme are not complete as shown in Figs. 1 and 2. As this is a 2-h examination, there should be more questions set to have a total of 100 marks.

ChatGPT is further prompted to prepare the suggested solution for the examination paper. It generates the suggested solution in a quite systematic, structured and nearly instant manner again as shown in Fig. 3. However, it is found that there are not any steps shown in Section A but full calculations are manipulated in Section B of the suggested solution.

The ChatGPT-generated solutions are clear and totally correct in Q.1, 2 and 3 of Section B as shown in Fig. 3. However, it is found that the solution to Q.4 of Section B is missing. This implies that ChatGPT-generated solutions to the ChatGPT-generated assessment script are not complete at all. The ChatGPT-generated assessment script and solution should be carefully vetted and revised if deemed necessary.



Could you prepare probability and engineering statistics examination paper for year 1 civil engineering students?

ChatGPT

Certainly! Here's a sample examination paper for Year 1 civil engineering students, focusing on probability and engineering statistics:

Probability and Engineering Statistics Examination Paper Duration: 2 hours

Instructions:

1. Read all questions carefully before answering.
2. Show all your working and provide clear explanations for your answers.
3. Use appropriate mathematical notation and units where necessary.
4. This examination paper consists of two sections: Section A and Section B.
5. Section A contains multiple-choice questions. Select the best answer for each question and mark it on the answer sheet provided.
6. Section B contains descriptive questions. Provide detailed solutions for each question in the space provided.

Section A: Multiple Choice Questions [30 marks]

Choose the best answer for each question and mark it on the answer sheet.

1. Probability is defined as:
 - a) The likelihood of an event occurring.
 - b) The number of favorable outcomes divided by the total number of outcomes.
 - c) The study of random variables.
 - d) The ratio of the number of favorable outcomes to the number of possible outcomes.
2. A fair six-sided die is rolled. What is the probability of rolling an odd number?
 - a) $1/6$
 - b) $1/3$
 - c) $1/2$
 - d) $2/3$
3. A bag contains 5 red marbles and 3 green marbles. If a marble is drawn at random, what is the probability of drawing a red marble?
 - a) $3/8$
 - b) $5/8$
 - c) $5/3$
 - d) $8/5$

Fig. 1. A ChatGPT-generated examination script of probability and engineering statistics.

ChatGPT is finally prompted to identify the level of difficulty of the ChatGPT-generated questions based on Bloom's Taxonomy. Figure 4 indicates that there are four application-based questions, two analysis-based questions and two memory-based (i.e. Remember) questions among eight questions. Even if both conceptual and procedural knowledge within the knowledge dimension seems to be covered in the ChatGPT-generated questions, the result however indicates there are not any evaluation-based and creation-based questions in the assessment script as per Bloom's Taxonomy. This implies that ChatGPT-generated assessment scripts should further be reviewed and amended to

4. In a class of 30 students, 18 students play football and 12 students play basketball. If a student is chosen at random, what is the probability that the student plays both football and basketball?
- 1/30
 - 6/30
 - 12/30
 - 18/30
5. The mean of a dataset is calculated by:
- Adding all the data points and dividing by the sample size.
 - Taking the median of the dataset.
 - Taking the mode of the dataset.
 - Finding the range of the dataset.

Section B: Descriptive Questions [70 marks]

Answer all questions in this section. Provide detailed solutions and show all your working.

- The lifetimes of a certain brand of light bulbs are normally distributed with a mean of 800 hours and a standard deviation of 50 hours. Compute the following:
 - What percentage of bulbs will last between 700 and 900 hours?
 - What is the probability that a bulb will last more than 950 hours?
- A civil engineer is analyzing the compressive strength of concrete. The engineer randomly selects 20 concrete samples and records their compressive strengths (in MPa) as follows:
20.1, 19.8, 19.9, 20.2, 20.3, 19.7, 20.0, 20.1, 20.4, 20.5, 19.7, 20.1, 19.9, 20.0, 20.3, 20.2, 20.0, 20.1, 19.8, 19.9
 - Calculate the sample mean and sample standard deviation for the compressive strengths.
 - Determine the 95% confidence interval for the mean compressive strength of the concrete.
- A civil engineering firm is conducting a survey on the heights of a population. A random sample of 100 individuals is taken, and their heights (in centimeters) are recorded. The sample mean height is found to be 165 cm, with a standard deviation of 10 cm. Calculate a 99% confidence interval for the population mean height.
- A contractor is bidding on a construction project. Based on historical data, the contractor believes that the number of construction workers required for the project follows a Poisson distribution with a mean of 12 workers per day. Calculate the probability that the contractor will need less than 10 workers on any given day.

End of Examination Paper

Please note that this is just a sample examination paper and the actual content and difficulty level of the questions may vary based on the curriculum and course requirements of your institution.

Fig. 1. (continued)

ascertain different levels of difficulty of questions aligned with the subject curriculum and the subject learning outcomes.



Could you prepare the marking scheme of the above examination paper?

ChatGPT

Certainly! Here's the marking scheme for the Probability and Engineering Statistics Examination Paper:

Section A: Multiple Choice Questions [30 marks]

Each question is worth 2 marks. Allocate 2 marks for each correct answer and 0 marks for each incorrect or unanswered question.

Section B: Descriptive Questions [70 marks]

Question 1:

a) 4 marks for correctly calculating the Z-scores and using the Z-table to find the probabilities.

b) 2 marks for correctly calculating the Z-score and using the Z-table to find the probability.

Total: 6 marks

Question 2:

a) 2 marks for correctly calculating the sample mean.

2 marks for correctly calculating the sample standard deviation.

b) 3 marks for correctly using the t-distribution and calculating the confidence interval.

Total: 7 marks

Question 3:

4 marks for correctly calculating the margin of error and using the t-distribution to find the confidence interval.

Total: 4 marks

Question 4:

4 marks for correctly using the Poisson distribution formula and calculating the probability.

Total: 4 marks

Total marks for the examination paper: 30 (Section A) + 40 (Section B) = 70 marks

Please note that this marking scheme is a suggestion and can be adjusted based on the specific requirements and marking criteria of your institution.

Fig. 2. A ChatGPT-generated marking scheme of probability and engineering statistics.

Could you prepare the suggested solution of the above examination paper?

● ChatGPT

Certainly! Here's the suggested solution for the Probability and Engineering Statistics Examination Paper:

Section A: Multiple Choice Questions

1. Probability is defined as:

Answer: a) The likelihood of an event occurring.

2. A fair six-sided die is rolled. What is the probability of rolling an odd number?

Answer: c) 1/2

3. A bag contains 5 red marbles and 3 green marbles. If a marble is drawn at random, what is the probability of drawing a red marble?

Answer: b) 5/8

4. In a class of 30 students, 18 students play football and 12 students play basketball. If a student is chosen at random, what is the probability that the student plays both football and basketball?

Answer: b) 6/30

5. The mean of a dataset is calculated by:

Answer: a) Adding all the data points and dividing by the sample size.

Section B: Descriptive Questions

1. The lifetimes of a certain brand of light bulbs are normally distributed with a mean of 800 hours and a standard deviation of 50 hours. Compute the following:

a) What percentage of bulbs will last between 700 and 900 hours?

Solution: To find the percentage of bulbs that will last between 700 and 900 hours, we need to calculate the z-scores for both values and use the Z-table to find the corresponding probabilities.

The z-score for 700 hours:

$$z_1 = (700 - 800) / 50 = -2$$

The z-score for 900 hours:

$$z_2 = (900 - 800) / 50 = 2$$

Using the Z-table, we find that the area to the left of $z = -2$ is 0.0228, and the area to the left of $z = 2$ is 0.9772.

The probability of bulbs lasting between 700 and 900 hours is the difference between these two probabilities:

$$P(700 \leq X \leq 900) = P(Z \leq 2) - P(Z \leq -2) = 0.9772 - 0.0228 = 0.9544$$

Therefore, 95.44% of the bulbs will last between 700 and 900 hours.

b) What is the probability that a bulb will last more than 950 hours?

Solution: To find the probability that a bulb will last more than 950 hours, we need to calculate the z-score for 950 hours and find the corresponding probability using the Z-table.

The z-score for 950 hours:

$$z = (950 - 800) / 50 = 3$$

Using the Z-table, we find that the area to the left of $z = 3$ is 0.9987. The probability of a bulb lasting more than 950 hours is the complement of this probability:

$$P(X > 950) = 1 - P(Z \leq 3) = 1 - 0.9987 = 0.0013$$

Therefore, the probability that a bulb will last more than 950 hours is 0.0013.

Fig. 3. A ChatGPT-generated solution of probability and engineering statistics.

2. A civil engineer is analyzing the compressive strength of concrete. The engineer randomly selects 20 concrete samples and records their compressive strengths (in MPa) as follows:
20.1, 19.8, 19.9, 20.2, 20.3, 19.7, 20.0, 20.1, 20.4, 20.5, 19.7, 20.1, 19.9, 20.0, 20.3, 20.2, 20.0, 20.1, 19.8, 19.9

a) Calculate the sample mean and sample standard deviation for the compressive strengths.

Solution:

Sample mean (\bar{x}) = $(20.1 + 19.8 + 19.9 + 20.2 + 20.3 + 19.7 + 20.0 + 20.1 + 20.4 + 20.5 + 19.7 + 20.1 + 19.9 + 20.0 + 20.3 + 20.2 + 20.0 + 20.1 + 19.8 + 19.9) / 20$

$\bar{x} = 399.8 / 20$

$\bar{x} = 19.99$ MPa (rounded to two decimal places)

To calculate the sample standard deviation, we need to find the deviations from the mean for each observation, square them, sum them, divide by (n-1), and take the square root. Deviations from the mean:

(20.1 - 19.99), (19.8 - 19.99), (19.9 - 19.99), (20.2 - 19.99), (20.3 - 19.99), (19.7 - 19.99), (20.0 - 19.99), (20.1 - 19.99), (20.4 - 19.99), (20.5 - 19.99), (19.7 - 19.99), (20.1 - 19.99), (19.9 - 19.99), (20.0 - 19.99), (20.3 - 19.99), (20.2 - 19.99), (20.0 - 19.99), (20.1 - 19.99), (19.8 - 19.99), (19.9 - 19.99)

Squared deviations:
(0.01)², (-0.19)², (-0.09)², (0.21)², (0.31)², (-0.29)², (0.01)², (0.11)², (0.41)², (0.51)², (-0.29)², (0.11)², (-0.09)², (0.01)², (0.31)², (0.21)², (0.01)², (0.11)², (-0.19)², (-0.09)²

Squared deviations:

(0.01)², (-0.19)², (-0.09)², (0.21)², (0.31)², (-0.29)², (0.01)², (0.11)², (0.41)², (0.51)², (-0.29)², (0.11)², (-0.09)², (0.01)², (0.31)², (0.21)², (0.01)², (0.11)², (-0.19)², (-0.09)²

Sum of squared deviations = 0.8094

Sample variance (s^2) = Sum of squared deviations / (n-1)

$s^2 = 0.8094 / (20-1)$

$s^2 = 0.045$

Sample standard deviation (s) = $\sqrt{s^2}$

$s = \sqrt{0.045}$

$s = 0.2121$ MPa (rounded to four decimal places)

b) Determine the 95% confidence interval for the mean compressive strength of the concrete.

Solution:

To determine the confidence interval, we will use the t-distribution since the population standard deviation is unknown and the sample size is small ($n = 20$).

The degrees of freedom (df) = $n - 1 = 20 - 1 = 19$

The critical value for a 95% confidence level and $df = 19$ is approximately 2.093 (obtained from the t-distribution table or calculator).

Standard error (SE) = s / \sqrt{n}

$SE = 0.2121 / \sqrt{20}$

$SE = 0.0474$ (rounded to four decimal places)

Fig. 3. (continued)

Margin of error (ME) = Critical value * SE

$$ME = 2.093 * 0.0474$$

$$ME = 0.099 \text{ (rounded to three decimal places)}$$

Confidence interval = Sample mean \pm Margin of error

$$CI = 19.99 \pm 0.099$$

$$CI = (19.891, 20.089) \text{ MPa (rounded to three decimal places)}$$

Therefore, the 95% confidence interval for the mean compressive strength of the concrete is approximately (19.891, 20.089) MPa.

3. A civil engineering firm is conducting a survey on the heights of a population. A random sample of 100 individuals is taken, and their heights (in centimeters) are recorded. The sample mean height is found to be 165 cm, with a standard deviation of 10 cm. Calculate a 99% confidence interval for the population mean height. Solution:
Since the population standard deviation is unknown, we will use the t-distribution to calculate the confidence interval.

The sample size (n) = 100

$$\text{Degrees of freedom (df)} = n - 1 = 100 - 1 = 99$$

The critical value for a 99% confidence level and df = 99 is approximately 2.626 (obtained from the t-distribution table or calculator).

Standard error (SE) = s / \sqrt{n}

$$SE = 10 / \sqrt{100}$$

$$SE = 10 / 10$$

$$SE = 1$$

Margin of error (ME) = Critical value * SE

$$ME = 2.626 * 1$$

$$ME = 2.626$$

Confidence interval = Sample mean \pm Margin of error

$$CI = 165 \pm 2.626$$

$$CI = (162.374, 167.626) \text{ cm}$$

Therefore, the 99% confidence interval for the population mean height is approximately (162.374, 167.626)

Fig. 3. (continued)

4 Conclusion and Future Works

The results of the ChatGPT-generated assessment script, its marking scheme and its suggested solution of Probability and Engineering Statistics are not very promising as the ChatGPT-generated assessment script is partially complete and one of the multiple-choice questions is not correct. The total score of the assessment script is not consistent with that of its marking scheme. Its suggested solution to one of the questions is missing. Furthermore, there are many application-oriented questions but few analysis-based questions and no evaluation-based questions at all in the assessment script as per Bloom's Taxonomy. A high cognitive level of questions in an assessment script is crucial and indispensable to enable students to develop logical thinking and problem-solving abilities (Kwan 2017). It is concluded that ChatGPT-generated assessment scripts should further be reviewed and refined by educational practitioners to ascertain different levels

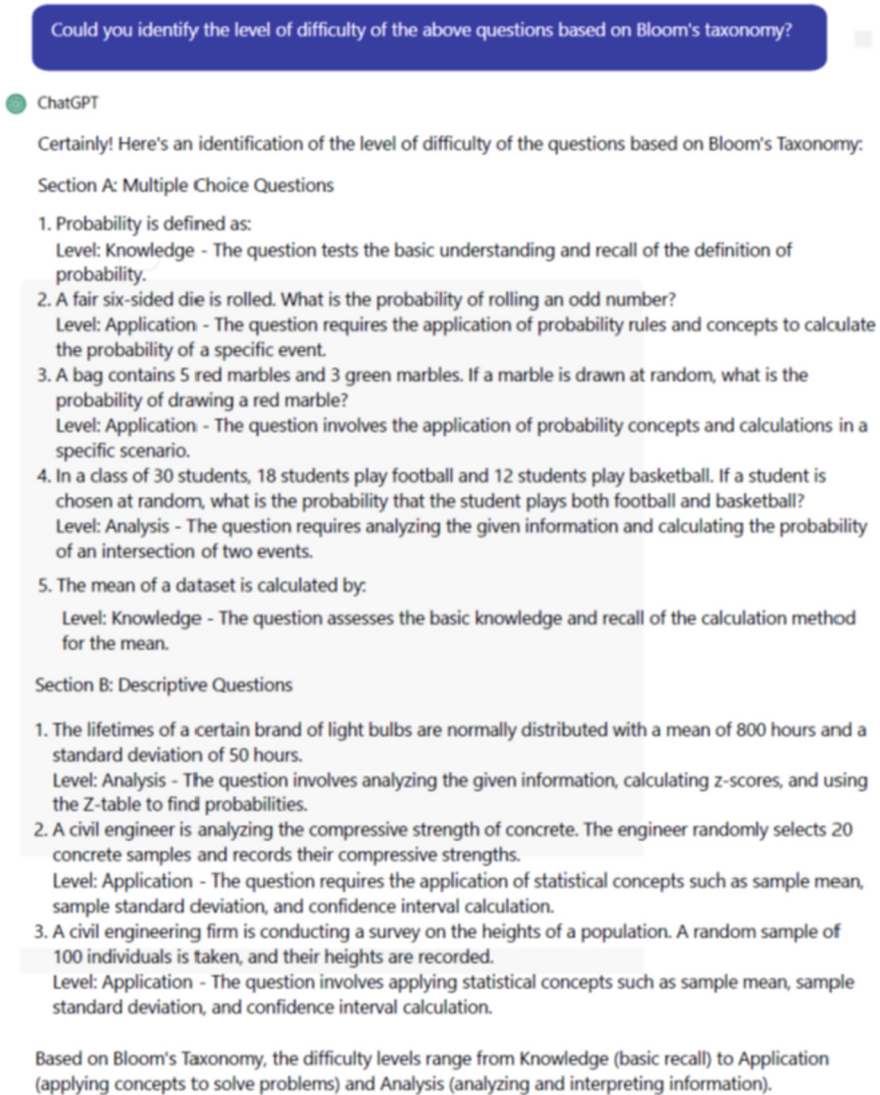


Fig. 4. ChatGPT-identified level of difficulty of the questions based on Bloom's Taxonomy.

of difficulty of questions that are in good alignment with the subject curriculum and the subject learning outcomes.

Further investigation needs to be carried out to ascertain whether the quality of ChatGPT-generated assessment scripts can be dramatically improved in terms of the knowledge dimension and the cognitive process dimension according to Bloom's Taxonomy in view of the rapid development of Generative AI. In addition, subject Instructors, class teachers, and educational practitioners will save a lot of time in preparing

assessment scripts, suggested solutions, and marking schemes if ChatGPT can do well in the near future.

References

- Anderson, L.W., et al.: *A Taxonomy for Learning, Teaching, and Assessing: a Revision of Bloom's Taxonomy of Educational Objectives*. Longman, New York (2001)
- Biggs, J.: *Teaching for Quality Learning at University*, 2nd edn. Society for Research into Higher Education & Open University Press (2003)
- Chandio, M.T., Pandhiani, S.M., Iqbal, R.: Bloom's taxonomy: Improving assessment and teaching-learning process. *J. Educ. Educ. Develop.* **3**(2), 203–221 (2016)
- Hong Kong Institution of Engineers. *Professional Accreditation Handbook (Engineering Degrees)*. Accreditation Board, pp. 1–35 (2013)
- Johri, A., Katz, A.S., Qadir, J., Hingle, A.: Artificial intelligence and engineering education. *J. Eng. Educ.* **112**(3), 572–577 (2023)
- Krathwohl, D.R.: A revision of bloom's taxonomy: an overview. *Theory Into Practice* **41**(4), 212–218 (2002)
- Kwan, C.L.C.: Findings from the assessment of students' learning outcomes in engineering mathematics. *Int. J. Learn. Teach.* **9**(2), 291–297 (2017)
- Lambert, D., Lines, D.: *Understanding Assessment: Purposes, Perceptions, Practices*. London, Routledge Falmer (2000)
- Mayer, R.E.: A taxonomy for computer-based assessment of problem solving. *Comput. Hum. Behav.* **18**(6), 623–632 (2002)
- Nikolic, S., et al.: ChatGPT versus engineering education assessment: a multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *Eur. J. Eng. Educ.* **48**(4), 559–614 (2023)
- Sazhin, S.: Teaching mathematics to engineering students. *Int. J. Eng. Educ.* **14**, 145–152 (1998)

Institutional Strategies and Practices



A Survey of Evaluation Approaches in STE(A)M Education

Billy T. M. Wong^(✉), Kam Cheong Li, and Hon Tung Chan

Institute for Research in Open and Innovative Education, Hong Kong Metropolitan University,
Ho Man Tin, Kowloon, Hong Kong SAR, China
{tamiwong, kcli, hotchan}@hkmu.edu.hk

Abstract. STEM and STEAM (or STE(A)M) education has been widely practised over the past two decades. It involves the disciplinary integration of science, technology, engineering, and mathematics, as well as arts. Evaluation plays an essential role in assessing and improving the effectiveness of STE(A)M educational practices. Despite a broad range of research on evaluating the effectiveness of STE(A)M education, little work has been conducted on summarising and analysing the evaluation approaches. This paper addresses the research gap by examining the literature in relation to the evaluation of STE(A)M education in terms of its research types, evaluation objectives, data collection and analysis methods, and approaches. The study covers a total of 250 related research articles published between 2015 and 2022 which were collected from the Scopus database and analysed using a content analysis approach. The findings reveal that a majority of the articles are experimental and interventional in nature, with an emphasis on examining the effects of STE(A)M education as well as the attitudes and perceptions of stakeholders. The results also show the common use of descriptive statistics, t-tests, ANOVA, and regression as quantitative data analysis approaches, whereas thematic and content analysis were often adopted for qualitative data analysis. The main approaches of STE(A)M education were found to be involved in the articles reviewed, namely problem-centred learning, STE(A)M education integration, inquiry-based learning, design-based learning, and co-operative learning. These findings contribute to enhancing our understanding about evaluation in STE(A)M education, as well as providing practical references for its evaluation practices.

Keywords: STEM education · STEAM education · evaluation · educational effectiveness

1 Introduction

STEM and STEAM (or STE(A)M) education, which involves the disciplinary integration of science, technology, engineering, and mathematics, as well as arts, was introduced in the early 2000s (Dugger 2010). Since then, STE(A)M education has been rapidly developing across the globe at different levels of education, including primary, secondary, tertiary, and vocational (Korbel 2016; Sapounidis et al. 2023). Governments worldwide

have been implementing policies and initiatives to support the incorporation of STE(A)M into curriculums. The Hong Kong Education Bureau, for example, has developed key learning areas in curriculums by connecting STE(A)M-related concepts with real life (EDB 2023). Similarly, the Ministry of Education in Denmark has promoted the integration of STE(A)M disciplines through a skills and application focused approach (Timms et al. 2018).

Evaluation plays an essential role in assessing and improving the effectiveness of STE(A)M educational practices. With the popularity of STE(A)M education, considerable attention has been also paid to its evaluation. Relevant work has addressed various aspects of these practices, such as their effectiveness in motivating and engaging students in learning (Chittum et al. 2017; Connors-Kellgren et al. 2013; Kennedy & Odell 2014), the affordances and challenges that teachers and students have in relation to STE(A)M education (Markworth et al. 2016), and their perceptions of and readiness for it (El-Deghaidy & Mansour 2015; Miorelli et al. 2015). The evaluation results have advanced our understanding of STE(A)M education, as well as provided pedagogical insights into ways to optimise it for students.

Despite a broad range of research involving evaluation of STE(A)M education, little work has been conducted on summarising and analysing the patterns of relevant evaluations. This paper addresses this research gap by reviewing the work in relation to the evaluation of STE(A)M education. The study focuses on the following research questions:

- i. What are the types of STE(A)M education research that involve evaluation?
- ii. What are the research issues addressed in the evaluation?
- iii. What are the data collection and analysis approaches used in the evaluation?
- iv. What are the STE(A)M educational approaches adopted in relevant work?

2 Related Literature

Research on the evaluation of STE(A)M education has addressed a range of areas. A major area concerns the application of technologies in STE(A)M education (Ajit et al. 2021; Çetin & Demircan 2020; Sirakaya & Sirakaya 2022; Xu & Ouyang 2022). Ajit et al. (2021), for example, examined publications on the benefits and challenges of the use of augmented reality in STEM education. They found that stimulating student learning achievements was most frequently reported as a benefit for this technology, while the issues of lighting and instability of augmented reality tools, as well as the usability issues resulting from heavy devices, were commonly reported as challenges. Çetin and Demircan (2020) analysed prior scholarship that provided children with programming experiences through robotics in order to identify the possible affordances of robotics programming for the integration of technology and engineering in STEM education. The authors observed that programming through robotics is an effective way to engage children with engineering and technology and therefore could be embedded into the engineering curriculum. Li and Wong (2020) surveyed case studies that reported the use of learning analytics in STE(A)M education. They identified the patterns and trends in relation to the types of data used, STE(A)M learning practices, and the benefits and limitations of learning analytics for STE(A)M education.

Another area focuses on examinations of ways to enhance the effectiveness of STE(A)M education (Abu Khurma et al. 2023; Jin 2021; Thibaut et al. 2018; van den Hurk et al. 2019). For example, Thibaut et al. (2018) analysed research on integrated STEM in secondary education. Their findings show the development of integrated STEM education, which is mainly based on social constructivist theories that view learning as socially constructed by students' existing experiences. Their work also reveals relevant instructional practices which focused primarily on STEM content integration, problem-solving learning, inquiry-based learning, design-based learning, and cooperative learning. Based on these findings, the authors proposed a theoretical framework for instructional practices for integrated STEM education. Van den Hurk et al. (2019) explored studies in relation to STEM-related interventions and found that only a few of these interventions were shown to be useful in increasing students' interest in STEM or their persistence in STEM education. Li and Wong (2021, 2023) reviewed the development of personalised learning in STE(A)M education. They identified the most common objective of personalised learning in this context which lay in catering for different learning styles, as well as the frequently addressed research issues and the aspects of practising personalisation.

Moreover, the research trends and gaps of STE(A)M education have been commonly evaluated (Bayanova et al. 2023; Darmawansah et al. 2023; Kayan-Fadlelmula et al. 2022). Darmawansah et al. (2023), for example, reviewed prior publications on robotics-based STEM (R-STEM) education. They found that much research has been conducted in the context of K-12 schools in the United States to examine teacher and student perceptions of R-STEM education, where LEGO and project-based learning were the most frequently adopted learning tool and approach, respectively. Their findings suggest ways to advance R-STEM education, such as using the concept of the zone of proximal development in R-STEM education research. Another study by Kayan-Fadlelmula et al. (2022) focused on STEM education in Gulf Cooperation Council countries (i.e., Bahrain, Kuwait, Qatar, and the United Arab Emirates "UAE"). Its findings reveal that most relevant publications focused on the UAE context and analysed the effectiveness of STEM education using quantitative research methods. They also identified research gaps such as the underrepresentation of women in STEM education, a lack of comparative or longitudinal approaches, and little attention being paid to potential enablers that may improve student willingness to participate in STEM-related fields and careers. Kwan and Wong (2021) reviewed the characteristics of and trends in STEAM education. The findings highlight emerging research on enhancing the effectiveness and broadening the diversity of STE(A)M education, as well as teachers' perceptions, development, and support.

Despite the various focuses of these studies, the evaluation of STE(A)M education practices has long been neglected in the literature. In fact, evaluation plays an essential role in improving education practice. As Saxton et al. (2014) explain, meaningful, effective, and iterative assessments provide schools, teachers, and districts with critical information and clear directions to improve the quality of STEM teaching. To address the research gap, this paper examines previous scholarship on the evaluation of STE(A)M

education practices. The study focuses on identifying how STE(A)M education practices were evaluated in the literature, as well as the patterns in relation to the STE(A)M educational approaches, research issues, data collection, and analysis approaches.

3 Methodology

3.1 Data Collection

Relevant publications for this study were collected from the Scopus database, which has been widely used for literature reviews on various topics. The keywords (“STEAM education” or “STEM education” or “STEAM learning” or “STEM learning” or “STEAM teaching” or “STEM teaching”) were used for searching the publications. The period of publication was set from 2015 to 2022. All the search results were screened based on the following selection criteria: (i) the paper reports on an empirical study on the evaluation of a STE(A)M education practice, (ii) it is written in English, and (iii) it is available in full text. Those publications which do not meet any of these criteria were excluded from the present study. A total of 250 papers were eventually selected for review and analysis.

3.2 Data Analysis

The selected papers were examined using a content analysis approach. For each paper, information related to the STE(A)M education research conducted, the research issues addressed, the data collection and analysis approaches used, and the educational approaches adopted was coded and categorised. The preliminary coding was performed by a researcher and then checked by one of the authors of this paper. Any inconsistency in the coding judgement was discussed until an agreement was reached. Based on the coding and categorisation results, the patterns in relation to the research questions were identified.

4 Results

4.1 Number of Publications

Figure 1 shows the number of publications between 2015 and 2022. The number remained relatively low in the earlier years. There was a clear increase in the number in the past two years. The results show a growing interest in the evaluation of STE(A)M education in recent years.

4.2 Reported STE(A)M Education Research

Figure 2 shows the types of STE(A)M education research reported in the publications. Six types of research were identified. The most common being experimental, which accounts for 51.2% of the research studies. This is followed by interventional (37.6%) and longitudinal (12.4%) research. Only a small proportion of studies adopted correlational (5.2%), descriptive (2.4%), and case study (2.4%) approaches while having an evaluation component.

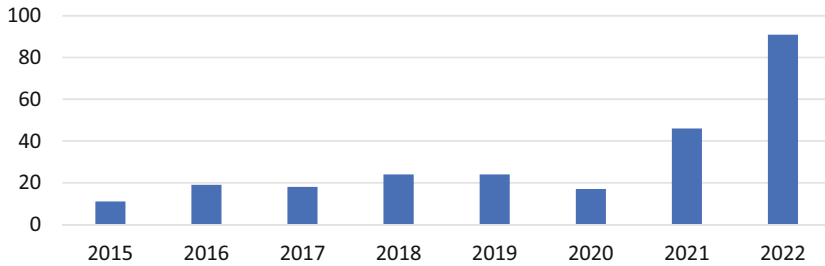


Fig. 1. Number of publications between 2015 and 2022

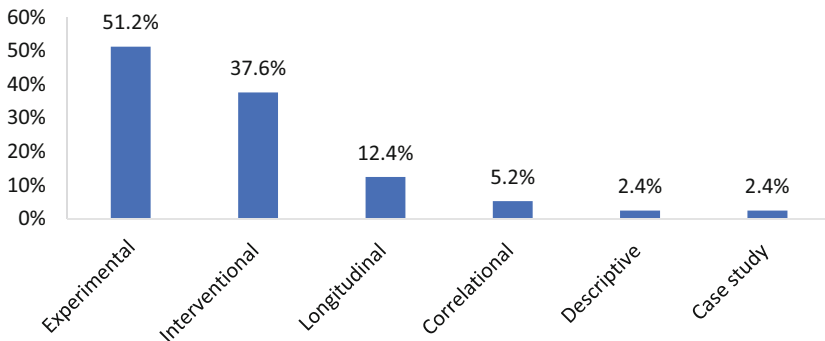


Fig. 2. STE(A)M education research reported in the publications

4.3 Research Issues Addressed

Figure 3 shows the research issues addressed in the evaluation of STE(A)M education. The effects of STE(A)M education have been most frequently evaluated, i.e., in 51.2% of the studies. This is followed by the attitudes and perceptions on STE(A)M education (20.4%) and the effectiveness of STE(A)M education (19.2%). Relatively small proportions of evaluation addressed the factors affecting STE(A)M learning (9.2%), STE(A)M learning processes (6.8%), potential for and challenges of STE(A)M initiatives (6.0%), student performance (5.6%), teacher preparedness (5.2%), teaching performance (3.6%), and student capability (2.0%). Overall, the results suggest a broad variety of research issues have attracted attention from researchers in the field of STE(A)M education.

4.4 Data Collection Approaches

Figure 4 shows the data collection approaches used for the evaluations. The use of surveys is the most common approach, covering 51.2% of the evaluations, followed by interviews (34.8%), tests (33.6%), observations (31.2%), and system log data (2.4%) such as online interaction and learning logs. Other approaches (13.2%) include the collection of course materials (e.g., lesson plans, learning materials, and rubrics) and assessment contents. The results suggest a preference for collecting data directly from study participants for the evaluations.

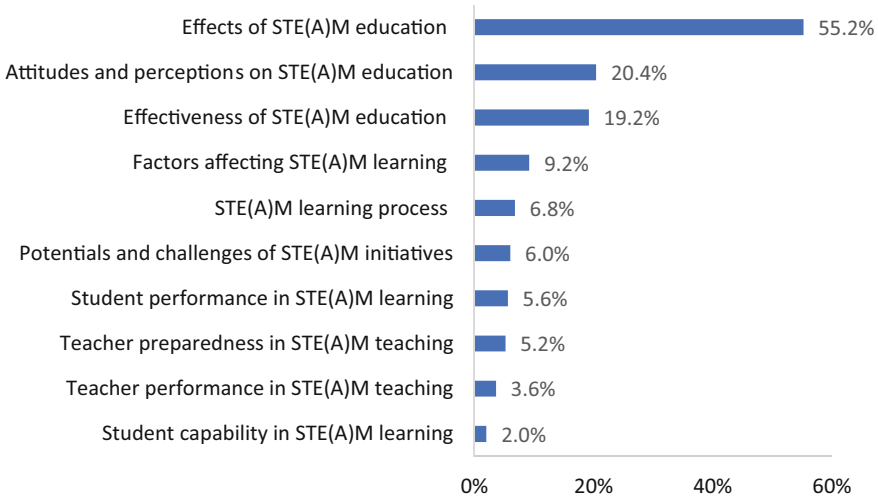


Fig. 3. Research issues addressed in the evaluations

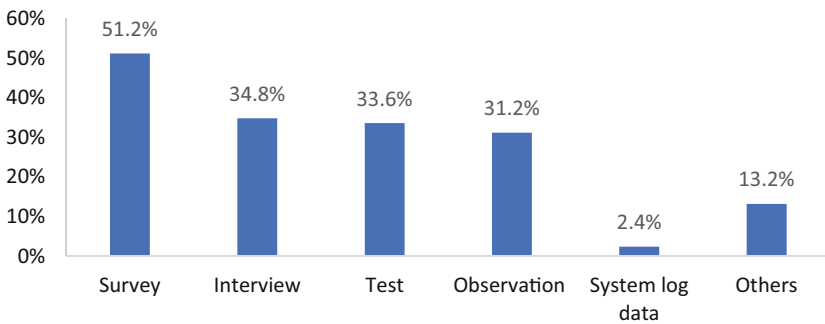


Fig. 4. Data collection approaches used for the evaluations

4.5 Data Analysis Approaches

A wide variety of data analysis approaches were identified for the evaluations. Figure 5 reports the top 10 most frequent data analysis approaches. Part of the evaluations applied more than one approach. Descriptive statistics was most commonly used in 60% of the evaluations, followed by *t*-tests (19.6%), ANOVA (17.2%), regression (11.2%), and thematic analysis (8.0%). These results show that quantitative approaches have been more commonly used for data analysis in the evaluations.

4.6 Educational Approaches

Figure 6 illustrates the educational approaches adopted in the relevant STE(A)M education practices that were evaluated. Problem-centred learning (28.8%), integration of STE(A)M content (28.4%), and inquiry-based learning (27.2%) were relatively more

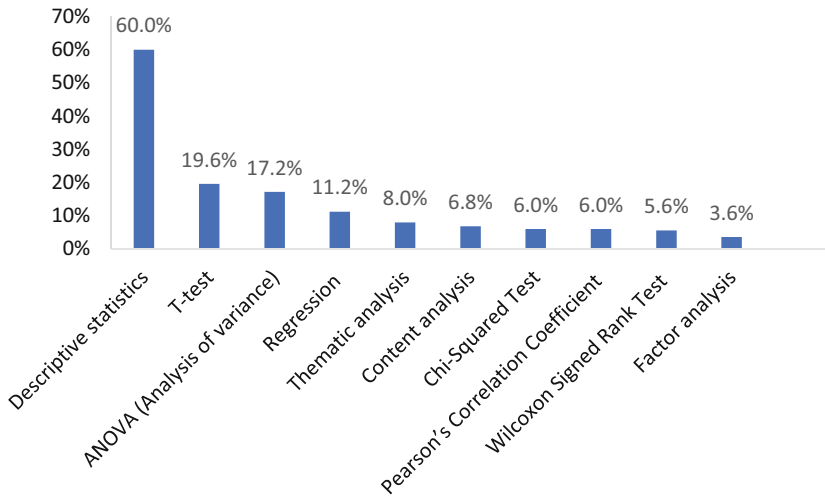


Fig. 5. Top 10 most frequent data analysis approaches for the evaluations

common. This is followed by design-based learning (21.6%) and cooperative learning (14.8%). There is also a small proportion (1.2%) of practices which adopted other approaches, including evidence-based learning, interactive learning, and task-centred learning. These results provide background on the educational contexts evaluated.

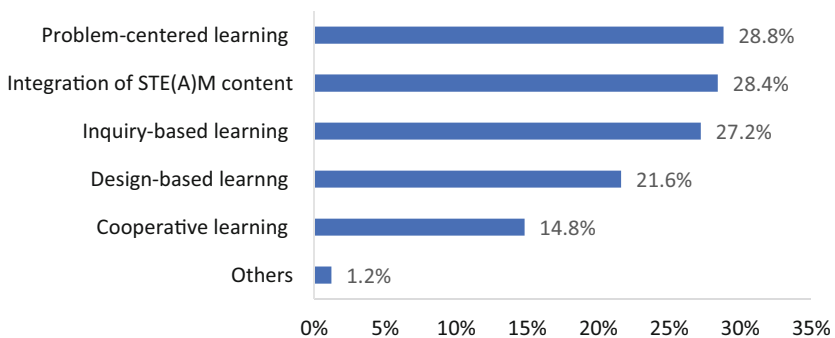


Fig. 6. Educational approaches adopted in STE(A)M education

5 Discussion and Conclusion

The findings of this study contribute to revealing the patterns of evaluation in STE(A)M education practices in terms of the types of STE(A)M education research, research issues, data collection and analysis methods, and educational approaches. These findings have advanced our understanding of evaluation practices in STE(A)M education, as well as

informed evaluation design for the future. They also supplement the results of relevant reviews and provide a comprehensive picture about the development of STE(A)M education.

The finding concerning the number of publications in the early years (low) compared to the number in recent years (an increase) is consistent with the findings of Kwan and Wong (2021), who show that early work of STE(A)M education was more concerned with the enhancement of our understanding and the potential of STE(A)M education, while the later work focused more on evaluating the effectiveness of STE(A)M practices. Furthermore, the increase of publications in recent years could be attributed to the COVID-19 pandemic, which attracted much attention in terms of examining how STE(A)M education was affected by social and physical distancing due to the pandemic (Li et al. 2023).

It was found that most of the evaluations were experimental in nature, with particular emphasis on examining the effects of and attitudes and perceptions on STE(A)M education. These findings are related to the frequent use of quantitative data collection and analysis approaches. With the increasing adoption of learning analytics in STE(A)M education (Li & Wong 2020), future evaluation work would benefit from relevant data processing and analytical techniques to address areas such as students' learning processes, which are difficult to evaluate.

The STE(A)M educational approaches were found to be closely related to problem-centred learning, integration of STE(A)M content, inquiry-based learning, design-based learning, and cooperative learning. These approaches have been recognised as essential in STE(A)M instruction (Connor et al. 2015; Kwan & Wong 2021). Their frequent use reveals the need to develop the diverse skills of students in STE(A)M. Accordingly, future studies should address how the different educational approaches facilitate the teaching and learning of those skills.

These findings also serve to inform potential future work on the evaluation of STE(A)M education. Further studies could focus on examining the differences, if any, in the evaluation practices at different times in the evolution of STE(A)M education. In addition, more qualitative evaluations could be carried out to obtain an in-depth understanding about the issues investigated in the evaluation in order to supplement the quantitative findings.

Acknowledgments. The work described in this paper was partially supported by grants from Hong Kong Metropolitan University (2021/011 and CP/2022/04).

References

- Abu Khurma, O., Al Darayseh, A., Alramamneh, Y.: A framework for incorporating the “Learning How to Learn” approach in teaching STEM education. *Educ. Sci.* **13**, 1 (2023)
- Ajit, G., Lucas, T., Kanyan, R.: A systematic review of augmented reality in STEM education. *Stud. Appl. Econ.* **39**(1), 1–22 (2021)
- Bayanova, A.R., Orekhovskaya, N.A., Sokolova, N.L., Shaleeva, E.F., Knyazeva, S.A., Budkevich, R.L.: Exploring the role of motivation in STEM education: a systematic review. *Eurasia J. Math. Sci. Technol. Educ.* **19**(4), em2250 (2023). <https://doi.org/10.29333/ejmste/13086>

- Çetin, M., Demircan, H.Ö.: Empowering technology and engineering for STEM education through programming robots: a systematic literature review. *Early Child Dev. Care* **190**(9), 1323–1335 (2020)
- Chittum, J.R., Jones, B.D., Akalin, S., Schram, A.B.: The effects of an afterschool STEM program on students' motivation and engagement. *Int. J. STEM Educ.* **4**(11), 1–16 (2017)
- Connor, A.M., Karmokar, S., Whittington, C.: From STEM to STEAM: strategies for enhancing engineering & technology education. *Int. J. Eng. Pedagogy* **5**(2), 37–47 (2015)
- Connors-Kellgren, A., Parker, C.E., Blustein, D.L., Barnett, M.: Innovations and challenges in project-based STEM education: lessons from ITEST. *J. Sci. Educ. Technol.* **25**, 825–832 (2013)
- Darmawansah, D., Hwang, G.J., Chen, M.R., Liang, J.C.: Trends and research foci of robotics-based STEM education: a systematic review from diverse angles based on the technology-based learning model. *Int. J. STEM Educ.* **10**, 12 (2023)
- Dugger, W.E.: Evolution of STEM in the United States. Knowledge in technology education. In: Proceedings of the 6th Biennial International Conference on Technology Education: Volume One (TERC 2010) (2010)
- Education Bureau (2023). Examples on STEAM learning and teaching activities. Retrieved from <https://www.edb.gov.hk/en/curriculum-development/kla/ma/res/STEMexamples.html>
- El-Deghaidy, H., Mansour, N.: Science Teachers' perceptions of STEM education: possibilities and challenges. *Int. J. Learn. Teach.* **1**(1), 51–54 (2015)
- Jin, Q.: Supporting indigenous students in Science and STEM education: a systematic review. *Educ. Sci.* **11**, 555 (2021)
- Kayan-Fadlelmula, F., Sellami, A., Abdelkader, N., Umer, S.: A systematic review of STEM education research in the GCC countries: Trends, gaps, and barriers. *Int. J. STEM Educ.* **9**, 2 (2022)
- Kennedy, T.J., Odell, M.R.L.: Engaging students in STEM education. *Sci. Educ. Int.* **25**(3), 246–258 (2014)
- Korbel, P.: Measuring STEM in vocational education and training. National Centre for Vocational Education Research (NCVER) (2016). Retrieved from <https://eric.ed.gov/?id=ED570653>. Accessed 22 August 2023
- Kwan, R., Wong, B.T.M.: Latest advances in STEAM education research and practice: a review of the literature. *Int. J. Innova. Learn.* **29**(3), 323–339 (2021)
- Li, K.C., Wong, B.T.M.: Trends of learning analytics in STE(A)M education: a review of case studies. *Interact. Technol. Smart Educ.* **17**(3), 323–335 (2020)
- Li, K.C., Wong, B.T.M.: Personalised learning in STE(A)M education: a literature review. In: Li, R., Cheung, S.K.S., Iwasaki, C., Kwok, L.F., Kageto, M. (eds.) *Blended Learning: Re-thinking and Re-defining the Learning Process.. ICBL 2021. Lecture Notes in Computer Science()*, vol. 12830. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-80504-3_12
- Li, K.C., Wong, B.T.M.: Personalisation in STE(A)M education: A review of literature from 2011 to 2020. *J. Comput. High. Educ.* **35**, 186–201 (2023)
- Li, K.C., Wong, B.T.M., Chan, H.T.: Teaching and learning innovations for distance learning in the digital era: a literature review. *Front. Educ.* **8** (2023). <https://doi.org/10.3389/feduc.2023.1198034>
- Markworth, K.A., Brobst, J., Ohana, C., Parker, R.: Elementary content specialization: models, affordances, and constraints. *Int. J. STEM Educ.* **3**, 16 (2016)
- Miorelli, J., Stambach, N., Moskal, B., Dwyer, J.: Improving faculty perception of and engagement in STEM education. In: IEEE Frontiers in Education Conference, El Paso, United States (2015)
- Sapounidis, T., Tselegkaridis, S., Stamovlasis, D.: Educational robotics and STEM in primary education: a review and a meta-analysis. *J. Res. Technol. Educ.* (2023)
- Saxton, E., et al.: A common measurement system for K-12 STEM education: adopting an educational evaluation methodology that elevates theoretical foundations and systems thinking. *Stud. Educ. Eval.* **40**, 18–35 (2014)

- Sirakaya, M., Sirakaya, D.A.: Augmented reality in STEM education: a systematic review. *Interact. Learn. Environ.* **30**(8), 1556–1569 (2022)
- Thibaut, L., et al.: Integrated STEM education: a systematic review of instructional practices in secondary education. *Eur. J. STEM Educ.* **3**(1), 2 (2018)
- Timms, M.J., Moyle, K., Weldon, P.R., Mitchell, P.: Challenges in STEM learning in Australian schools: Literature and policy review. Australian Council for Educational Research, Victoria (2018)
- van den Hurk, A., Meelissen, M., van Langen, A.: Interventions in education to prevent STEM pipeline leakage. *Int. J. Sci. Educ.* **41**(2), 150–164 (2019)
- Xu, W., Ouyang, F.: The application of AI technologies in STEM education: a systematic review from 2011 to 2021. *Int. J. STEM Educ.* **9**(59), 1–20 (2022)



Uncertainty-Based Metamorphic Testing for Validating Plagiarism Detection Systems

Pak Yuen Patrick Chan^(✉), Jacky Keung, and Zhen Yang

Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong SAR, China

ppychan2-c@my.cityu.edu.hk

Abstract. Plagiarism is a severe issue in academia, and uncertainty in plagiarism detection systems might lead to inconsistent detections. Thus, evaluating the system is essential; however, it is also a test oracle problem as it is challenging to distinguish correct behaviour from potentially incorrect behaviour of the system. To alleviate this challenge, we develop a feasible approach by applying an uncertainty matrix to identify the uncertainty of the plagiarism detection systems and derive metamorphic relations of metamorphic testing from the identified uncertainty for validation. We experimented with three plagiarism detection systems in a classroom scenario where students were hypothesized to use tools to generate answers for assignments. These answers were fed into the systems for validation by comparing the systems' similarity scores of the tool-generated answers. Results showed that the proposed approach can effectively validate plagiarism detection systems. Future studies can apply this approach to locate uncertainties to enhance systems' robustness.

Keywords: Validation · Metamorphic testing · Uncertainty · Natural language processing · Plagiarism detection

1 Introduction

Despite advanced technologies have been applied in detecting plagiarism to enhance performance (Kumar et al., 2022; Setha & Aliane, 2022; Wang & Shi, 2022; Zhao et al., 2022), other advanced technologies have also made detection more difficult (Verma, 2023; Y. Xiao et al., 2022). Thus, evaluating the performance of Plagiarism Detection Systems (PDSs) is essential because plagiarism is a severe issue in academia and incorrect detections will cause disputes (Verma, 2023) and reduce trust in using the systems (Zhang et al., 2022). Evaluating a PDS is a test oracle problem as it is difficult to distinguish correct behaviour from potentially incorrect behaviour of PDSs, so this paper proposed an approach by applying an Uncertainty Matrix (UMX) to systematically identify the uncertainty of PDSs and use identified uncertainty to derive Metamorphic Relations (MR) of Metamorphic Testing (MT) for validation. This paper experimented with three different PDSs. We simulated a classroom scenario where students used Artificial Intelligence Generated Content (AIGC) and paraphrasing tools to generate different assignment answers. The generated answers were fed into the subject systems separately for validation by comparing the systems' similarity scores of the tool-generated answers.

The main contributions of this paper are summarized as follows. 1. To the best of our knowledge, we are the first to apply UMX and MT together to validate PDSs. 2. The proposed approach can validate PDSs effectively, which shows that the approach could be extended into validating the uncertainty of other domains. 3. We created a systematic uncertainty-based validation approach for researchers and practitioners to follow and apply. 4. We experimented with applying Explainable Artificial Intelligence (XAI) to analyse the results for internal validity. 5. We outlined several research challenges and opportunities that researchers may encounter in the extension of the proposed approach for application in other areas.

The rest of this paper is structured as follows. Section 2 describes the background of this research under four aspects: plagiarism detection, uncertainty, metamorphic testing and the proposed approach. Section 3 provides an overview of the evaluation methodology. The results and discussion will be presented in Sect. 4. We conclude and discuss limitations and future research avenues in Sect. 5.

2 Literature Review

2.1 Plagiarism Detection

Plagiarism is often found in academia (Hartanto et al., 2021). Institute of Electrical and Electronics Engineers (IEEE) “defines plagiarism as the reuse of someone else’s prior processes, results, or words without explicitly acknowledging the original author and source”, and IEEE penalized offenders by disallowing publication in all IEEE copyrighted publications for up to five years (IEEE, 2021). Thus, plagiarism is a severe issue. Based on the definition of plagiarism, measuring the similarity of documents is an approach to detecting plagiarism. Several commonly used similarity calculation methods include Manhattan, Euclidean distance, Jaccard coefficients, and Cosine similarity (Achsan et al., 2022; Alobed et al., 2021; Du & Hu, 2022; Guan et al., 2022; Ha et al., 2021; Hartanto et al., 2021; Oo & Pa, 2020; Putra et al., 2022). The cosine similarity method is popular because it measures similarity between vectors, which has advantages over others (Guan et al., 2022).

Machine learning has been applied to improve the performance of detection of text documents such as fraudulent reviews, spam filtering, tag extraction and plagiarism (Kumar et al., 2022; Setha & Aliane, 2022; Wang & Shi, 2022; Zhao et al., 2022). A systematic review study has also found that plagiarism detection was improved by “Better semantic text analysis methods” and “Application of machine learning” (Foltýnek et al., 2019). To apply machine learning techniques to detecting text documents, the machine learning model typically works with Natural Language Processing (NLP). NLP is one of the computation techniques to capture the semantics of natural language texts and uses different knowledge representations for specific purposes (Cambria & White, 2014; dos Santos et al., 2021; Sonbol et al., 2022). Machine learning models can be based on representations of NLP to make predictions.

Term Frequency Inverse Document Frequency (TFIDF) is a term weighting scheme commonly used in NLP to represent textual documents as vectors (Akbar et al., 2022; Alobed et al., 2021; Oo & Pa, 2020; Putra et al., 2022; Sonbol et al., 2022). TFIDF is widely used and has been applied with cosine similarity method, machine learning

algorithms and word embedding techniques for text classifications, text similarity, and plagiarism detection (Achsan et al., 2022; Arabi & Akbari, 2022; Du & Hu, 2022; Oo & Pa, 2020; Putra et al., 2022; Ullah et al., 2021). Other advanced NLP techniques, such as Sentence-Transformer “SBERT”, “OpenAI Generative Pre-trained Transformer GPT2”, “word2vec”, and “doc2vec”, are also used with cosine similarity method and machine learning techniques for plagiarism detection, text categorization and noun phrase representation (Ajallouda et al., 2022; Akbar et al., 2022; Arabi & Akbari, 2022; Putra et al., 2022; Ramnarain-Seetohul et al., 2022; Setha & Aliane, 2022; Sonbol et al., 2022; Veisi et al., 2022).

Advanced technologies, such as AIGC, have also made detection difficult (Verma, 2023; Y. Xiao et al., 2022). No system can detect plagiarism 100% correctly (Verma, 2023), and there was a lack of performance evaluation of PDSs (Albluwi, 2019; Foltýnek et al., 2020). If the detection is incorrect, there will be severe consequences; thus, the uncertainty of PDSs is of crucial concern (Verma, 2023; Zhang et al., 2022).

2.2 Uncertainty

Uncertainty concerns with the variability of outcomes we are unsure about due to a lack of information, whereas risk refers to the probability of unexpected outcomes. It is necessary to assess the risks that uncertainty imposes on the safety of machine learning systems to increase the credibility of adopting systems (Zhang et al., 2022). Lupafya and Balasubramaniam (2022) proposed a framework to characterize and document the attributes in uncertainties of systems. Schulam and Saria (2019) introduced the Resampling Uncertainty Estimation algorithm to measure the uncertainty of machine learning models. Alwidian et al. (2020) proposed an uncertainty matrix based on three dimensions of uncertainty: location, level, and nature, to identify critical uncertainties. After identifying the uncertainty, we need an approach to validate so that the systems can offer performance without uncertainty, and MT is a valuable validating tool.

2.3 Metamorphic Testing (MT)

Metamorphic testing (MT) is one approach that can alleviate the problem of test oracles (Xie et al., 2020; Ying et al., 2021; Zhang et al., 2021; Zhou et al., 2018). MT has been used to validate the expected relationships between inputs and outputs of program executions. These expected relationships are expressed as metamorphic relations (MR). If the output of the program executions does not match the expected relationships, this violates the MR and fault is revealed (Segura et al., 2019; Sun et al., 2022; Xie et al., 2020; Zhou et al., 2018). So, by examining these expected relationships between inputs and outputs, we can validate the program under testing.

For example, a mathematical property of the *sine* function, which is “ $\sin(x) = \sin(\pi - x)$ ”. If the corresponding MR has two inputs, x_1 and x_2 , that satisfy “ $x_1 + x_2 = \pi$ ”, then two function outputs should be equal, i.e., $\sin(x_1) = \sin(x_2)$. Based on MR, x_2 is constructed based on x_1 and $\sin(x_1)$. Therefore, x_1 is the source input, and x_2 is the follow-up input. Such MR can be used to test the program as if the outputs of $\sin(x_1)$ and $\sin(x_2)$ are different, then the implementation of function “ $\sin()$ ” must have faults as it violates the above MR (Zhang et al., 2021).

MT has been successfully applied to alleviate test oracle problems in various domains (Segura et al., 2018), such as bioinformatic software (Stacy et al., 2022), machine learning classifiers (Ellis et al., 2021; Riccio et al., 2020; Saha & Kanewala, 2019; Xie et al., 2011), Deep learning compilers (D. Xiao et al., 2022), and test order generation (Zhang et al., 2021).

2.4 Proposed Approach

This study aims to develop an uncertainty-based validating approach using MT for PDSs and demonstrate the practicality of the proposed approach through a comprehensive scaled experiment. We adopted Alwidian et al. (2020)'s uncertainty matrix (Table 1) to identify the uncertainty of the PDSs. One of the "Scenario" level of uncertainties of PDSs is inconsistent detection, and the cause of this uncertainty can be located in "Context", "Model Structure", and "Input" (Table 1). This study focuses on the "Input" location because the subject PDSs were using default options and experimented in the same test scenario (Sect. 3), so we assumed "Context" and "Model Structure" were constant. Thus, we identified that the uncertainty of PDSs is due to having a range of values of inputs (Table 1).

After identifying the cause of uncertainty, we adopted the MT approach to validate PDSs. We derive the MR of MT based on the identified cause and create MR called "MR1: Paraphrasing object". This relation assumes that the similarity of the document should not be affected by paraphrasing. In other words, systems should identify the "Source" document and the "Follow-up" documents, which are created by MR1, as similar (Fig. 1). MR1 is the relation that should be held. If not, there is a violation with MR1 and uncertainty in the PDS.

Table 1. Alwidian et al. (2020)'s uncertainty matrix

Location	Level			Nature	
	Statistical	Scenario	Recognized Ignorance	Epistemic	Variability
Context	deviation from true values	External/future environment	Possible at any location	Possible at any location	Possible at any location
Model Structure	deviation from true values	different assumption about structure			
Input	measurement uncertainty	Range of values			
Model Outcome	accumulated uncertainty which results from uncertainty in location X level X nature				

Source doc = [(man eats food)]
 Follow-up doc = [(male human consumes edibles)] ← paraphrased by MR1
 Therefore, System predicts (Source doc, Follow-up doc) = similar

Fig. 1. Illustration of MR1: Paraphrasing object

3 Methodology

The main objective of this experiment is to attempt to answer **RQ1: Can we validate PDSs by MT with MR derived from the uncertainty matrix?** We derived the MR from the uncertainty of PDSs and evaluated the effectiveness of the MR in validating subject systems by counting the occurrence of violations related to MR.

3.1 Experiment Design

3.1.1 Experiment Scenario

Inspired by (Araujo et al., 2022)’s study, this experiment simulates a “single assignment question multiple assignment answers” classroom scenario. Three different assignment questions were given to three classes of twelve students. One Computer Science (CS) assignment was expected to return Java source codes. Two Law assignments (Law1 and Law2) were open questions and expected to return English text documents.

3.1.2 Subject Systems and Technologies Under Test (SUT)

Plagiarism Detection System (PDS). SBERT and TFIDF are commonly used word embedding techniques for plagiarism detection (Arabi & Akbari, 2022; Putra et al., 2022), so this study employed them. We focused on PDSs that measure the similarity between two documents, and three subject PDSs were set up. Inspired by Ha et al. (2021), the first PDS, “BERT_COS”, using SBERT Transformer with the pre-trained model “bert-base-uncased” and cosine similarity method. The second PDS, “NN-TFIDF”, follows Nursalman et al. (2018)’s unsupervised implementation using TFIDF with cosine similarity method and Nearest Neighbour algorithm. The third PDS, “Neu-SBERT”, follows Merchant (2020)’s implementation using SBERT Transformers, neural network algorithm and Stanford Natural Language Inference Corpus (SNLI). SNLI is a supervised dataset for sentence similarity with three labels: “contradiction” means the sentences share no similarity, “entailment” means the sentences have a similar meaning, and “neutral” means the sentences are neutral. Neu_SBERT is trained with SNLI corpus and derived semantically meaningful sentence embeddings for similarity prediction with the labels.

Artificial Intelligence Generated Content (AIGC). AIGC refers to using Artificial Intelligence (AI) to generate different kinds of content, including text, images, videos and audio (Lin et al., 2023; Wu et al., 2023). ChatGPT, one of the AIGC applications that can provide detailed responses based on users’ input (Wu et al., 2023), had been discussed its relationship with plagiarism (Verma, 2023; Y. Xiao et al., 2022). This study employed two different AIGC applications from “poe.com”: ChatGPT, powered by OpenAI’s GPT3.5 technology, and Claude, powered by Anthropic’s technology (POE.COM,

2023). We assumed each student would use one of the two AIGC applications to generate assignment answers.

Table 2. List of online free paraphrasing tools

ID	Site name	Web address
p1	Paraphrasing-Tool	https://paraphrasing-tool.com/
p2	Quillbot	https://quillbot.com/
p3	PrePostSeo	https://www.prepostseo.com/paraphrasing-tool
p4	Rewriter-tools	https://www.rewritertools.com/paraphrasing-tool
p5	SEOMagnifier	https://seomagnifier.com/online-paraphrasing-tool

Paraphrasing Tools. Online paraphrasing tools could make detection more difficult for PDSs (Prentice & Kinden, 2018), and we selected five different free online paraphrasing tools for MR1 (Table 2). In this study, each student would use zero or one of the paraphrasing tools to paraphrase the generated answers.

Explainable Artificial Intelligence (XAI). We selected Shapley XAI technology for internal validity. SHAP measures each feature’s contribution towards the prediction (Poeh et al., 2020). It is based on coalitional game theory, which explains prediction by “assuming that each feature value of the instance is a player in a game where the prediction is the payout”, and the Shapley values are calculated this way and explain how to distribute the prediction among the features (Molnar, 2020).

3.1.3 Configuration

This experiment environment included using “Intel Core” i7 CPU with 64 GB RAM, “Windows 10” operating systems, “Jupyter Notebook” development platform and “Python” programming language and related libraries, such as “scikit_learn” and “tensorflow”, for machine learning algorithms, “shap” is for XAI, and “transformer” is for language processing. Standard parameters and options were applied to SUTs.

3.1.4 Data

According to MR1, original and paraphrased questions or paraphrased and original answers should give similar results. For the CS assignment, as source codes might not be paraphrased because paraphrasing might lead to the source codes being uncompileable. Instead, we assumed students would paraphrase the CS assignment question to create slightly different answers to avoid plagiarism detection. So, we paraphrased the CS questions to create follow-up questions and then fed them into AIGC tools to generate CS answers. Including the model answer of CS, the total number of CS answers was thirteen (Table 3). For Law assignments, we fed the original question into AIGC tools to generate original LAW1 and LAW2 answers. Then, we use paraphrasing tools to

Table 3. Details of subject data

Type of assignment	CS	LAW1	LAW2
<i>Model answer</i>			
- original	std_answer-cs	std_answer-law1	std_answer-law2
<i>Generated answer</i>			
By ChatGPT			
- original	gpt-org-cs	gpt-org-law1	gpt-org-law2
- paraphrased by p1	gpt-p1-cs	gpt-p1-law1	gpt-p1-law2
- paraphrased by p2	gpt-p2-cs	gpt-p2-law1	gpt-p2-law2
- paraphrased by p3	gpt-p3-cs	gpt-p3-law1	gpt-p3-law2
- paraphrased by p4	gpt-p4-cs	gpt-p4-law1	gpt-p4-law2
- paraphrased by p5	gpt-p5-cs	gpt-p5-law1	gpt-p5-law2
By Claude			
- original	cla-org-cs	cla-org-law1	cla-org-law2
- paraphrased by p1	cla-p1-cs	cla-p1-law1	cla-p1-law2
- paraphrased by p2	cla-p2-cs	cla-p2-law1	cla-p2-law2
- paraphrased by p3	cla-p3-cs	cla-p3-law1	cla-p3-law2
- paraphrased by p4	cla-p4-cs	cla-p4-law1	cla-p4-law2
- paraphrased by p5	cla-p5-cs	cla-p5-law1	cla-p5-law2
Total	13	13	13

generate follow-up answers. Including the model answers of LAW1 and LAW2, the total number of Law1 and Law2 answers was twenty-six. Therefore, this experiment had thirty-nine answers (data) (Table 3).

3.1.5 Measurement

This study measures the correctness of plagiarism detections, so subject PDSs should identify which answers were generated by the same AIGC application, even after paraphrasing. We used the similarity score calculated by PDSs to rank the similarity of answers. If PDS identifies a ChatGPT-generated answer more similar to a Claude-generated answer, this would be counted as a violation as the paraphrased answer misled PDS (Table 4). The violation rate will be measured by the violations over the total number of comparisons, which is one hundred and thirty-two as Eq. (1).

$$\text{Violation rate} = \text{Violation count} / \text{Total number of comparisons} \quad (1)$$

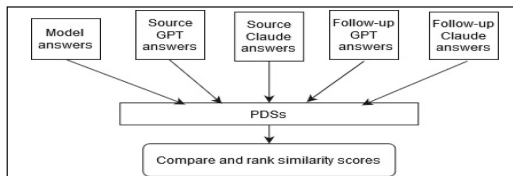
Table 4. Illustration of violating MR1

Answer	Similarity score with “gpt-org-law1”	Compare
cla-p5-law1	0.95	
gpt-p1-law1	0.91	Violation
gpt-p2-law1	0.96	
gpt-p3-law1	0.96	
gpt-p4-law1	0.96	
gpt-p5-law1	0.96	

3.2 Experiment Implementation

Step 1. For CS, an original question and five paraphrased questions, which were paraphrased by tools listed in Table 2 as MR1 stated, are inputted into two AIGC tools to generate answers. Answers from the original and paraphrased questions are marked in Table 3. For LAW1 and LAW2, original questions were inputted into two AIGC tools to generate original answers, and then the original answers were paraphrased by tools (Table 2) to create paraphrased answers (Table 3).

Step 2. We considered the generated answers without paraphrasing as the “Source” answers and the paraphrased answers as the “Follow-up” answers. Therefore, MR1 expects “Source” answers to be similar to “Follow-up” answers. CS and Law assignment answers were fed into PDSs separately. The results of similarity scores were used to compare and rank (Fig. 2) for counting violation of MR1 (Sect. 3.1.5).

**Fig. 2.** Implementation of experiment.

4 Discussion

4.1 Zero Violation

Table 5 shows that all PDSs performed better in CS than LAW1 and LAW2. BERT_COS did not violate MR1 with CS answers, which means BERT_COS could identify all the CS answers generated by the same AIGC tool as most similar to each other (Table 6). NN_TFIDF and Neu_SBERT did not violate MR1 with CS answers generated by Claude (Tables 5, 7, and 8). It is because CS answers were Java source codes that were similar

in size and highly structural for compilation. Thus, answering RQ1, MR1 can validate BERT_COS with CS answers generated by both AIGC tools and can validate NN_TFIDF and Neu_SBERT with CS answers generated by Claude.

Table 5. Violation counts in the experiment

	CS			LAW1			LAW2		
	BERT_COS	NN_TFIDF	Neu_SBERT	BERT_COS	NN_TFIDF	Neu_SBERT	BERT_COS	NN_TFIDF	Neu_SBERT
Violation(gpt)	0	6	6	3	8	6	2	6	13
Violation(cla)	0	0	0	1	8	9	4	2	6
Total violation	0	6	6	4	16	15	6	8	19
Total comparison	132	132	132	132	132	132	132	132	132
Violation rate	0%	4.55%	4.55%	3.03%	12.12%	11.36%	4.55%	6.06%	14.39%

Table 6. Matrix of similarity scores of BERT_COS with CS

CS-assignment-sim-score														
	std answer-cs	cla-org-cs	gpt-org-cs	cla-p1-cs	cla-p2-cs	cla-p3-cs	cla-p4-cs	cla-p5-cs	gpt-p1-cs	gpt-p2-cs	gpt-p3-cs	gpt-p4-cs	gpt-p5-cs	
std answer-cs		0.9598	0.9872	0.9598	0.9598	0.9598	0.9598	0.9598	0.9872	0.9872	0.9872	0.9872	0.9872	0.9872
cla-org-cs	0.9598		0.9264	1.0000	1.0000	1.0000	1.0000	1.0000	0.9264	0.9264	0.9264	0.9264	0.9264	0.9264
gpt-org-cs	0.9872	0.9264		0.9264	0.9264	0.9264	0.9264	0.9264	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
cla-p1-cs	0.9598	1.0000	0.9264		1.0000	1.0000	1.0000	1.0000	0.9264	0.9264	0.9264	0.9264	0.9264	0.9264
cla-p2-cs	0.9598	1.0000	0.9264	1.0000		1.0000	1.0000	1.0000	0.9264	0.9264	0.9264	0.9264	0.9264	0.9264
cla-p3-cs	0.9598	1.0000	0.9264	1.0000	1.0000		1.0000	1.0000	0.9264	0.9264	0.9264	0.9264	0.9264	0.9264
cla-p4-cs	0.9598	1.0000	0.9264	1.0000	1.0000	1.0000		1.0000	0.9264	0.9264	0.9264	0.9264	0.9264	0.9264
cla-p5-cs	0.9598	1.0000	0.9264	1.0000	1.0000	1.0000	1.0000		0.9264	0.9264	0.9264	0.9264	0.9264	0.9264
gpt-p1-cs	0.9872	0.9264	1.0000	0.9264	0.9264	0.9264	0.9264	0.9264		1.0000	1.0000	1.0000	1.0000	1.0000
gpt-p2-cs	0.9872	0.9264	1.0000	0.9264	0.9264	0.9264	0.9264	0.9264	1.0000		1.0000	1.0000	1.0000	1.0000
gpt-p3-cs	0.9872	0.9264	1.0000	0.9264	0.9264	0.9264	0.9264	0.9264	1.0000	1.0000		1.0000	1.0000	1.0000
gpt-p4-cs	0.9872	0.9264	1.0000	0.9264	0.9264	0.9264	0.9264	0.9264	1.0000	1.0000	1.0000		1.0000	1.0000
gpt-p5-cs	0.9872	0.9264	1.0000	0.9264	0.9264	0.9264	0.9264	0.9264	1.0000	1.0000	1.0000	1.0000		1.0000

Table 7. Matrix of similarity scores of NN_TFIDF with CS

CS-assignment-sim-score														
	std answer-cs	cla-org-cs	gpt-org-cs	cla-p1-cs	cla-p2-cs	cla-p3-cs	cla-p4-cs	cla-p5-cs	gpt-p1-cs	gpt-p2-cs	gpt-p3-cs	gpt-p4-cs	gpt-p5-cs	
std answer-cs		0.7657	0.9650	0.7705	0.7728	0.7728	0.7728	0.7728	0.8542	0.9682	0.9833	0.9833	0.9833	
cla-org-cs	0.7657		0.7308	0.9937	0.9931	0.9931	0.9931	0.9931	0.5898	0.6969	0.7092	0.7092	0.7092	0.7092
gpt-org-cs	0.9650	0.7308		0.7354	0.7372	0.7372	0.7372	0.7372	0.8313	0.9458	0.9810	0.9810	0.9810	0.9810
cla-p1-cs	0.7705	0.9937	0.7354		0.9993	0.9993	0.9993	0.9993	0.5935	0.7012	0.7137	0.7137	0.7137	0.7137
cla-p2-cs	0.7728	0.9931	0.7372	0.9993		1.0000	1.0000	1.0000	0.5946	0.7028	0.7152	0.7152	0.7152	0.7152
cla-p3-cs	0.7728	0.9931	0.7372	0.9993	1.0000		1.0000	1.0000	0.5946	0.7028	0.7152	0.7152	0.7152	0.7152
cla-p4-cs	0.7728	0.9931	0.7372	0.9993	1.0000	1.0000		1.0000	0.5946	0.7028	0.7152	0.7152	0.7152	0.7152
cla-p5-cs	0.7728	0.9931	0.7372	0.9993	1.0000	1.0000	1.0000		0.5946	0.7028	0.7152	0.7152	0.7152	0.7152
gpt-p1-cs	0.8542	0.5898	0.8313	0.5935	0.5946	0.5946	0.5946	0.5946		0.8823	0.8552	0.8552	0.8552	0.8552
gpt-p2-cs	0.9682	0.6969	0.9458	0.7012	0.7028	0.7028	0.7028	0.7028	0.8823		0.9659	0.9659	0.9659	0.9659
gpt-p3-cs	0.9833	0.7092	0.9810	0.7137	0.7152	0.7152	0.7152	0.7152	0.8552	0.9659		1.0000	1.0000	1.0000
gpt-p4-cs	0.9833	0.7092	0.9810	0.7137	0.7152	0.7152	0.7152	0.7152	0.8552	0.9659	1.0000		1.0000	1.0000
gpt-p5-cs	0.9833	0.7092	0.9810	0.7137	0.7152	0.7152	0.7152	0.7152	0.8552	0.9659	1.0000	1.0000		1.0000

4.2 Violation with MR1 and BERT_COS

BERT_COS had the lowest violation rate with Law assignment answers (Table 5). LAW1 and LAW2 questions were analysed type with no specific structure for answers. The low

Table 8. Matrix of similarity scores of Neu_SBERT with CS

CS -assignment-sim-score (score of entailments)													
	std answer-cs	cla-org-cs	gpt-org-cs	cla-p1-cs	cla-p2-cs	cla-p3-cs	cla-p4-cs	cla-p5-cs	gpt-p1-cs	gpt-p2-cs	gpt-p3-cs	gpt-p4-cs	gpt-p5-cs
std answer-cs		0.2181	0.6635	0.2198	0.2181	0.2181	0.2181	0.2181	0.2302	0.2302	0.2302	0.2302	0.2302
cla-org-cs	0.3659		0.3911	0.5859	0.6009	0.6009	0.6009	0.6009	0.2511	0.2511	0.2511	0.2511	0.2511
gpt-org-cs	0.4554	0.2508		0.2475	0.2508	0.2508	0.2508	0.2508	0.2354	0.2354	0.2354	0.2354	0.2354
cla-p1-cs	0.3674	0.5997	0.3872		0.5997	0.5997	0.5997	0.5997	0.2522	0.2522	0.2522	0.2522	0.2522
cla-p2-cs	0.3659	0.6009	0.3911	0.5859		0.6009	0.6009	0.6009	0.2511	0.2511	0.2511	0.2511	0.2511
cla-p3-cs	0.3659	0.6009	0.3911	0.5859	0.6009		0.6009	0.6009	0.2511	0.2511	0.2511	0.2511	0.2511
cla-p4-cs	0.3659	0.6009	0.3911	0.5859	0.6009	0.6009		0.6009	0.2511	0.2511	0.2511	0.2511	0.2511
cla-p5-cs	0.3659	0.6009	0.3911	0.5859	0.6009	0.6009	0.6009		0.2511	0.2511	0.2511	0.2511	0.2511
gpt-p1-cs	0.3772	0.4606	0.4544	0.4486	0.4606	0.4606	0.4606	0.4606		0.6669	0.6669	0.6669	0.6669
gpt-p2-cs	0.3772	0.4606	0.4544	0.4486	0.4606	0.4606	0.4606	0.4606	0.6669		0.6669	0.6669	0.6669
gpt-p3-cs	0.3772	0.4606	0.4544	0.4486	0.4606	0.4606	0.4606	0.4606	0.6669	0.6669		0.6669	0.6669
gpt-p4-cs	0.3772	0.4606	0.4544	0.4486	0.4606	0.4606	0.4606	0.4606	0.6669	0.6669	0.6669		0.6669
gpt-p5-cs	0.3772	0.4606	0.4544	0.4486	0.4606	0.4606	0.4606	0.4606	0.6669	0.6669	0.6669	0.6669	

violation rate represented that SBERT with the pre-trained model had built a semantic relationship that could identify most LAW answers generated by the same AIGC tool as most similar to each other, even after paraphrasing. In addition, Claude generated LAW1 and LAW2 answers, which were about 330 words and 280 words, respectively. On the other hand, ChatGPT generated LAW1 and LAW 2 answers, which were about 700 words and 440 words, accordingly. The size differences in LAW1 answers were significant between ChatGPT-generated and Claude-generated, which made it easier to distinguish. However, BERT_COS performed only 1.55% better in LAW1 than in LAW2 (Table 5), so the size difference effect was insignificant. XAI could not be applied in BERT_COS because it did not use a machine learning technique.

4.3 Violation with MR1 and NN_TFIDF

Table 7 shows NN_TFIDF identified some ChatGPT-generated CS answers were more similar to model answer, such as gpt-org-cs had a higher similarity score with std_answer-cs (0.9650) than with gpt-p1-cs (0.8313). As NN_TFIDF uses term frequencies to measure similarity and CS answers were highly structural and short, NN_TFIDF is expected to perform as well as BERT_COS. So, we used SHAP to interpret one of the violated cases of NN_TFIDF. Figure 3(a) and (b) show that “particle” was a negative feature of std_answer-cs and gpt-org-cs, whereas Fig. 3(c) shows “particle” was a positive feature of gpt-p1-cs. They explained the violation of NN-TFIDF in CS and showed that ChatGPT could generate source codes that made NN_TFIDF difficult to detect plagiarism.

Despite NN_TFIDF violated MR1 with LAW1 and LAW2, NN_TFIDF performed better in LAW2 (Table 5). LAW1 answers are longer and paraphrased, which might affect the term frequencies, leading to increase violations. SHAP interpreted one of the violated cases in NN_TFIDF with LAW1. Figure 4(a) and (c) show that “international” was a positive feature of cla-org-law1 and gpt-p1-law, whereas “international” was a negative feature of gpt-org-law1. In addition, SHAP interpreted one violated case in LAW2. Figure 5(a) and (b) show “international” was a negative feature of cla-p3-law2 and gpt-p2-law2; however, a positive feature of gpt-p5-law2. Apart from explaining the violations, Fig. 5(a), (b) and (c) had more common features than Fig. 4(a), (b) and (c), which can explain why NN_TFIDF performed worse in LAW1 than LAW2. It is because LAW1 answers, which were longer and paraphrased, contained more different terms which affected the term frequencies of NN_TFIDF to detect plagiarism.

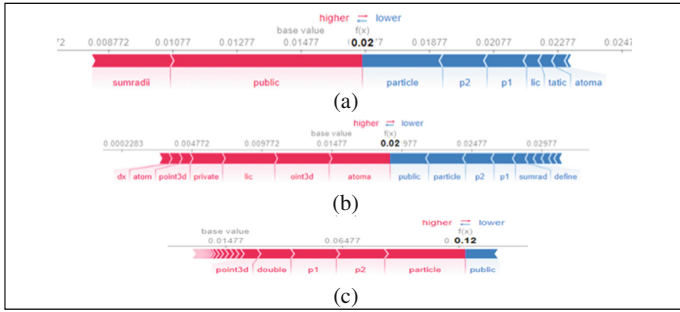


Fig. 3. SHAP interpretation on features that influenced NN_TFIDF’s similarity prediction with (a) std_answer-cs, (b) gpt-org-cs, (c) gpt-p1-cs.

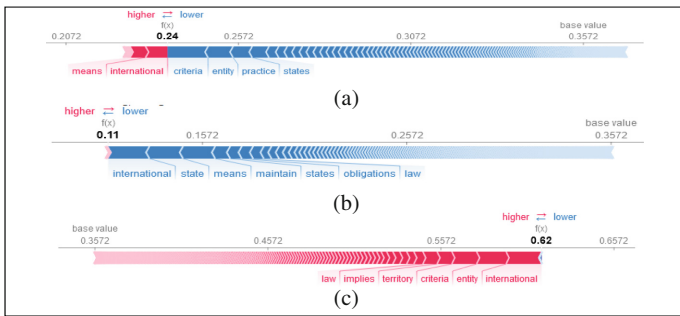


Fig. 4. SHAP interpretation on features that influenced NN_TFIDF’s similarity prediction with (a) cla-org-law1, (b) gpt-org-law1, (c) gpt-p1-law1.

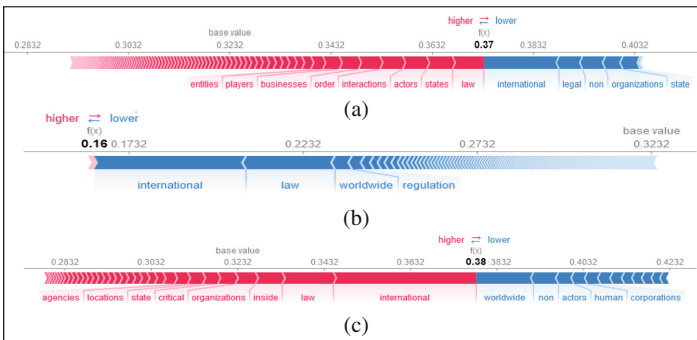


Fig. 5. SHAP interpretation on features that influenced NN_TFIDF’s similarity prediction with (a) gpt-p2-law2, (b) cla-p3-law2, (c) gpt-p5-law2.

4.4 Violation with MR1 and Neu_SBERT

Table 8 shows Neu_SBERT identified some ChatGPT-generated CS answers to be more similar to model answer and Claude-generated answers, such as all five paraphrased

ChatGPT-generated CS answers had a higher similar score with cla-p1-cs (0.2522) than with gpt-org-cs (0.2354). Neu_SBERT also used SBERT transformer as BERT_COS and CS answers were highly structural and short, Neu_SBERT is expected to perform as well as BERT_COS. We used SHAP to interpret one of the violated cases of Neu_SBERT and found that the term “public double x” is a negative feature of gpt-p1-cs and gpt-org-cs, whereas cla-p1-cs did not have “public double x”. It made Neu_SBERT identify gpt-p1-cs was more similar to cla-p1-cs than gpt-org-cs. The interpretations explained Neu_SBERT’s violations and showed that its semantic relationship to detect similarities between ChatGPT-generated CS answers was inaccurate. Thus, ChatGPT could generate source codes that made Neu_SBERT challenging to detect plagiarism.

Unexpectedly, Neu_SBERT performed poorly with both Law assignment answers (Table 5), even though it could build semantic relationships between words and sentences to detect similarities. We used SHAP to interpret one of the violated cases in LAW1 and found the features that contribute to the decisions of “entailment” of cla-p5-law1, gpt-p1-law1 and cla-org-law1. The term “that decide whether” was a negative feature of cla-p5-law1. cla-org-law1 contains the term “that determine whether”, which is semantically similar to the term “that decide whether”. On the other hand, gpt-p1-law1 does not contain any term similar to “that decide whether”, which made Neu_SBERT identify cla-p5-law1 was more similar to gpt-p1-law1 than cla-org-law1. The interpretations explained Neu_SBERT’s violations and showed that Neu_SBERT could not build an accurate semantic relationship with legal contents for predicting similarities.

4.5 Further Analysis

The experiment showed that using the uncertainty to derive MR is useful as this approach can identify the core concern of PDSs. If PDS do not violate this MR, the PDS is considered reliable. As BERT_COS is insensitive to MR1 with CS, we can assure the performance of BERT_COS in detecting plagiarism of CS.

Understandably, BERT_COS outperformed NN_TFIDF because SBERT, with a pre-trained model, can build semantic relationships between words and sentences, which is more suitable in LAW1 and LAW2. However, even though CS answers were highly structural and shorter than Law answers, results show BERT_COS outperformed NN_TFIDF in CS, which shows SBERT with a pre-trained model can also understand source code structure and outperform TFIDF.

At last, BERT_COS outperformed Neu_SBERT in the study, even though they used an SBERT transformer. The critical difference is the training corpus. While BERT_COS uses a pre-trained model, Neu_SBERT uses the SNLI corpus. SNLI corpus could have contained fewer source codes and legal sentences for training than the pre-trained model. Thus, PDSs should choose a pre-trained model over SNLI corpus in detecting plagiarism of CS and LAW assignment answers.

4.6 Validity

Different input parameters would lead to totally different results. In this experiment, Standard parameters and options (Sect. 3) were applied to all subject technologies and systems. We build subject systems using Python libraries and five online tools (Sect. 3)

to create the “follow-up” datasets. These tools and technologies should be considered reliable since they have been widely used. White-box testing has also been conducted so the source codes are error-free. Finally, this study focused on the relationship between MR and measurements mentioned in Sect. 3. The outcomes were listed and checked to ensure that the measurements mentioned in Sect. 3 were calculated accurately. In addition, XAI was applied to interpret violated cases, and the explanations matched the results. Thus, the potential threat to the internal validity of this study is low.

MT was widely used, and this study is based on a generic MT approach. We applied the approach to systems that were used in other published studies. Although the datasets used for assessment may vary case by case, the subject systems assessed and evaluated by our approach are rather general. Thus, the potential threat to the external validity of this study is also low.

5 Conclusion

The reliability of PDSs is essential; however, evaluating a PDS is a test oracle problem as it is difficult to distinguish correct behaviour from potentially incorrect behaviour of PDSs. Thus, this study proposed an approach that applies an uncertainty matrix to systematically identify the uncertainty of PDSs and uses the identified uncertainty to derive the MR of MT for validating PDSs. We experimented the approach with three different PDSs and used AIGC and paraphrasing tools, as MR stated, to generate different answers for validation by comparing the systems’ similarity scores of the answers.

Results show BERT_COS did not violate MR1 with all CS answers, and NN_TFIDF and Neu_SBERT did not violate MR1 with CS answers generated by Claude. It is because CS answers were source code programs that were well-structured and similar in size. Results also show that PDS using SBERT with a pre-trained model outperformed PDS using TFIDF and SBERT trained by SNLI corpus. It is because the semantic relationship built by BERT_COS with a pre-trained model can understand source code and legal-related sentences better than the other two PDSs. In conclusion, this study proved that the proposed approach can systematically and effectively identify uncertainty and derive associated MR for validating PDSs. We recommend using the proposed approach for validating PDSs using SBERT with a pre-trained model and cosine similarity method in the plagiarism detection of highly structural and similar-in-size documents.

There are a certain number of aspects that could extend this research. Firstly, future studies could extend our approach to different domains or technologies and derive MR from critical uncertainty to ensure the performance of validated targets. Secondly, future studies could experiment with the proposed approach with different sizes or types of documents to find the optimal combination for PDSs.

This study has also encountered several limitations. Firstly, we did not include human-created answers in the study. Secondly, this study focuses on the similarity of two documents, and to the best of our knowledge, the SNLI corpus was the only supervised dataset for sentence and document similarity.

Acknowledgements. This work is supported in part by the General Research Fund of the Research Grants Council of Hong Kong and the research funds of the City University of Hong Kong (6000796).

References

- Achsan, H.T.Y., Kurniawan, D., Purnama, D.G., Barcah, Q.K.D., Astoria, Y.Y.: Application of natural language processing using cosine-similarity algorithm in making chatbot information on the new capital city of the Republic of Indonesia. In: 2022 7th International Workshop on Big Data and Information Security (IWBIS), pp. 1–6 (2022)
- Ajalloua, L., Najmani, K., Zellou, A.: Doc2Vec, SBERT, InferSent, and USE which embedding technique for noun phrases? In: 2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), pp. 1–5. IEEE (2022)
- Akbar, R., Suharsono, T.N., Indrapriyatna, A.S.: Essay test based E-testing using cosine similarity vector space model. In: 2022 International Symposium on Information Technology and Digital Innovation (ISITDI), pp. 80–85. IEEE (2022)
- Albluwi, I.: Plagiarism in programming assessments: a systematic review. *ACM Trans. Comput. Educ.* **20**(1), 1–28 (2019)
- Allobed, M., Altrad, A.M., Bakar, Z.B.A.: A comparative analysis of Euclidean, Jaccard and Cosine similarity measure and arabic wordnet for automated arabic essay scoring. In: 2021 Fifth International Conference on Information Retrieval and Knowledge Management (CAMP), pp. 70–74. IEEE (2021)
- Alwidian, S., Dhaouadi, M., Famelis, M.: A vision towards a conceptual basis for the systematic treatment of uncertainty in goal modelling. In: Proceedings of the 12th System Analysis and Modelling Conference, pp. 139–142 (2020)
- Arabi, H., Akbari, M.: Improving plagiarism detection in text document using hybrid weighted similarity. *Exp. Syst. Appl.* **207**, 118034 (2022)
- Araujo, M.D., de Almeida, G.F., Nunes, J.L.: Epistemology goes AI: a study of GPT-3's capacity to generate consistent and coherent ordered sets of propositions on single-input-multiple-outputs basis. SSRN 4204178 (2022)
- Cambria, E., White, B.: Jumping NLP curves: a review of natural language processing research. *IEEE Comput. Intell. Mag.* **9**(2), 48–57 (2014)
- dos Santos, V., et al.: Using natural language processing to build graphical abstracts to be used in studies selection activity in secondary studies. In: 2021 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), pp. 1–8. IEEE (2021)
- Du, L., Hu, C.: Text similarity detection method of power customer service work order based on TFIDF algorithm. In: 2022 IEEE 5th International Conference on Information Systems and Computer Aided Education (ICISCAE), pp. 978–982. IEEE (2022)
- Ellis, J.D., Iqbal, R., Yoshimatsu, K.: Verification of the neural network training process for spectrum-based chemical substructure prediction using metamorphic testing. *J. Comput. Sci.* **55**, 101456 (2021)
- Foltýnek, T., et al.: Testing of support tools for plagiarism detection. *Int. J. Educ. Technol. High. Educ.* **17**(1), 1–31 (2020)
- Foltýnek, T., Meuschke, N., Gipp, B.: Academic plagiarism detection: a systematic literature review. *ACM Comput. Surv.* **52**(6), 1–42 (2019)
- Guan, D., Liu, D., Zhao, W.: Adversarial detection based on local cosine similarity. In: 2022 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), pp. 521–525. IEEE (2022)
- Ha, T.-T., Nguyen, V.-N., Nguyen, K.-H., Nguyen, K.-A., Than, Q.-K.: Utilizing sbert for finding similar questions in community question answering. In: 2021 13th International Conference on Knowledge and Systems Engineering (KSE), pp. 1–6. IEEE (2021)
- Hartanto, A.D., Pristyanto, Y., Saputra, A.: Document similarity detection using Rabin-Karp and cosine similarity algorithms. In: 2021 International Conference on Computer Science and Engineering (IC2SE), pp. 1–6. IEEE (2021)

- IEEE. IEEE Publication Services and Products Board Operations Manual 2021 (2021). <https://www.ieee.org/content/dam/ieee-org/ieee/web/org/pubs/pspb/opsmanual.pdf>
- Kumar, A., Gopal, R.D., Shankar, R., Tan, K.H.: Fraudulent review detection model focusing on emotional expressions and explicit aspects: investigating the potential of feature engineering. *Decis. Support Syst.* **155**, 113728 (2022)
- Lin, Y., et al.: Blockchain-aided secure semantic communication for AI-generated content in metaverse. *IEEE Open J. Comput. Soc.* **4**, 72–83 (2023)
- Lupafya, C., Balasubramaniam, D.: A framework for considering uncertainty in software systems. 2022 In: IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC), pp. 1519–1524. IEEE (2022)
- Merchant, M.: Semantic Similarity with BERT (2020). https://colab.research.google.com/github/keras-team/keras-io/blob/master/examples/nlp/ipynb/semantic_similarity_with_bert.ipynb. Accessed 18 May 2023
- Molnar, C.: *Interpretable Machine Learning*. lulu.com (2020)
- Nursalman, M., Kusnendar, J., Fadhila, U.F.: Implementation of k-nearest neighbor with cosine similarity for classification abstract international journal of computer science. In: 2018 International Conference on Information Technology Systems and Innovation (ICITSI), pp. 43–48. IEEE (2018)
- Oo, H.M., Pa, W.P.: Myanmar news retrieval in vector space model using cosine similarity measure. In: 2020 IEEE Conference on Computer Applications (ICCA), pp. 1–5. IEEE (2020)
- POE.COM. About. (2023). <https://poe.com/about>. Accessed 18 May 2023
- Poth, A., Meyer, B., Schlicht, P., Riel, A.: Quality assurance for machine learning – an approach to function and system safeguarding. In: 2020 IEEE 20th International Conference on Software Quality, Reliability and Security (QRS), pp. 22–29. IEEE (2020)
- Prentice, F.M., Kinden, C.E.: Paraphrasing tools, language translation tools and plagiarism: an exploratory study. *Int. J. Educ. Integr.* **14**(1), 1–16 (2018)
- Putra, S.J., Gunawan, M.N., Hidayat, A.A.: Feature engineering with Word2vec on text classification using the K-nearest neighbor algorithm. In: 2022 10th International Conference on Cyber and IT Service Management (CITSM), pp. 1–6. IEEE (2022)
- Ramnarain-Seetohul, V., Bassoo, V., Rosunally, Y.: Work-in-progress: computing sentence similarity for short texts using transformer models. In: 2022 IEEE Global Engineering Education Conference (EDUCON), pp. 1765–1768. IEEE (2022)
- Riccio, V., Jahangirova, G., Stocco, A., Humatova, N., Weiss, M., Tonella, P.: Testing machine learning based systems: a systematic mapping. *Empiric. Softw. Eng.* **25**(6), 5193–5254 (2020)
- Saha, P., Kanewala, U.: Fault detection effectiveness of metamorphic relations developed for testing supervised classifiers. In: 2019 IEEE International Conference on Artificial Intelligence Testing (AITest), pp. 157–164 (2019)
- Schulam, P., Saria, S.: Can you trust this prediction? Auditing pointwise reliability after learning. In: The 22nd International Conference on Artificial Intelligence and Statistics, PMLR, pp. 1022–1031 (2019)
- Segura, S., Durán, A., Troya, J., Ruiz-Cortés, A.: Metamorphic relation patterns for query-based systems. In: 2019 IEEE/ACM 4th International Workshop on Metamorphic Testing (MET), pp. 24–31. IEEE (2019)
- Segura, S., Towey, D., Zhou, Z.Q., Chen, T.Y.: Metamorphic testing: testing the untestable. *IEEE Softw.* **37**(3), 46–53 (2018)
- Setha, I., Aliane, H.: Enhancing automatic plagiarism detection: using Doc2vec model. In: 2022 International Conference on Advanced Aspects of Software Engineering (ICAASE), pp. 1–5. IEEE (2022)
- Sonbol, R., Rebdawi, G., Ghneim, N.: The use of NLP-based text representation techniques to support requirement engineering tasks: a systematic mapping review. arXiv preprint [arXiv:2206.00421](https://arxiv.org/abs/2206.00421) (2022)

- Stacy, B., Hauzel, J., Lindvall, M., Porter, A., Pop, M.: Metamorphic testing in bioinformatics software: a case study on metagenomic assembly. In: 2022 IEEE/ACM 7th International Workshop on Metamorphic Testing (MET), pp. 31–33. IEEE (2022)
- Sun, C.-A., Liu, B., Fu, A., Liu, Y., Liu, H.: Path-directed source test case generation and prioritization in metamorphic testing. *J. Syst. Softw.* **183**, 111091 (2022)
- Ullah, F., Jabbar, S., Mostarda, L.: An intelligent decision support system for software plagiarism detection in academia. *Int. J. Intell. Syst.* **36**(6), 2730–2752 (2021)
- Veisi, H., Golchinpour, M., Salehi, M., Gharavi, E.: Multi-level text document similarity estimation and its application for plagiarism detection. *Iran J. Comput. Sci.* **5**(2), 143–155 (2022)
- Verma, P.: A professor accused his class of using ChatGPT, putting diplomas in jeopardy. *Washington Post* (2023). <https://www.washingtonpost.com/technology/2023/05/18/texas-professor-threatened-fail-class-chatgpt-cheating/>. Accessed 19 May 2023
- Wang, R., Shi, Y.: Research on application of article recommendation algorithm based on Word2Vec and TFIDF. In: 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), pp. 454–457. IEEE (2022)
- Wu, T., et al.: A brief overview of ChatGPT: the history, status quo and potential future development. *IEEE/CAA J. Automat. Sin.* **10**(5), 1122–1136 (2023)
- Xiao, D., Liu, Z., Yuan, Y., Pang, Q., Wang, S.: Metamorphic testing of deep learning compilers. *Proc. ACM Meas. Anal. Comput. Syst.* **6**(1), 15 (2022)
- Xiao, Y., Chatterjee, S., Gehringer, E.: A new era of plagiarism the danger of cheating using AI. In: 2022 20th International Conference on Information Technology Based Higher Education and Training (ITHET), pp. 1–6. IEEE (2022)
- Xie, X., Ho, J.W.K., Murphy, C., Kaiser, G., Xu, B., Chen, T.Y.: Testing and validating machine learning classifiers by metamorphic testing. *J. Syst. Softw.* **84**(4), 544–558 (2011)
- Xie, X., Zhang, Z., Chen, T.Y., Liu, Y., Poon, P.-L., Xu, B.: METTLE: a metamorphic testing approach to assessing and validating unsupervised machine learning systems. *IEEE Trans. Reliab.* **69**(4), 1293–1322 (2020)
- Ying, Z., Towey, D., Bellotti, A., Zhou, Z.Q., Chen, T.Y.: Preparing SQA professionals: metamorphic relation patterns, exploration, and testing for big data. In: Proceedings of the International Conference on Open and Innovation Education (ICOIE 2021), pp. 22–30 (2021)
- Zhang, M., Keung, J.W., Chen, T.Y., Xiao, Y.: Validating class integration test order generation systems with Metamorphic Testing. *Inf. Softw. Technol.* **132**, 106507 (2021)
- Zhang, X., Chan, F. T., Yan, C., Bose, I.: Towards risk-aware artificial intelligence and machine learning systems: an overview. *Decis. Supp. Syst.* 113800 (2022)
- Zhao, C., et al.: An improved term frequency-inverse document frequency method solving multi-text label problem. In: 2022 Global Conference on Robotics, Artificial Intelligence and Information Technology (GCRAIT), pp. 400–404. IEEE (2022)
- Zhou, Z.Q., Sun, L., Chen, T.Y., Towey, D.: Metamorphic relations for enhancing system understanding and use. *IEEE Trans. Software Eng.* **46**(10), 1120–1154 (2018)



Measuring the Effect of Family Structure on High-School Students' Stress and Enthusiasm for Learning Using Tourism-Related Questionnaire

Pongthorn Ruksorn, Suwitchaporn Witchakul, and Naraphorn Paoprasert^(✉)

Department of Industrial Engineering, Faculty of Engineering, Kasetsart University,
50 Ngamwongwan Road, Lat Yao, Chatuchak 10900, Bangkok, Thailand
naraphorn.p@ku.th

Abstract. This research study was to measure the effect of family structure on high-school students' stress and enthusiasm in learning via tourism questionnaire. The questionnaire was designed to collect the students' opinions in regards to leisure-travelling and information regarding the students' family structure which in our study was classified into two types: nuclear family and non-nuclear family in the context of Thailand. Binary logistic regression and the Statistical Package of the Social Sciences (SPSS) were selected as a tool to perform statistical analysis. The binary logistic regression model analysis revealed that students who live with nuclear families tend to be more eager in learning outside of the classroom during their leisure-travel time. On the other hand, students who live with non-nuclear families were more likely to be stressed in general resulting in the desire for an escape from studying and the need and/or physical mental relaxation.

Keywords: family structure · high-school student · tourism · nuclear family

1 Introduction

Nuclear family is a type of family in which kids stay with both of their parents (Randall, 2021; Hoffmann, 2023; Khudyakova et al., 2016). In Thailand, the trend for the number of nuclear families shows a declining signal as illustrated in Fig. 1 (Samutchak et al., 2022).

According to the Kom Chad Luek News (2019), the fact that children do not live with both parents can have various impacts on them, such as delayed growth or slower development compared to their age (Office of the Health Promotion Fund, 2019). Furthermore, if kids stay with their grandparents, when there is a significant age gap between the elderly and grandchildren, it can create a generation gap, which may lead to relationship issues that result in stress. It was also found that children in economically disadvantaged households also have fewer opportunities to live with both parents compared to those in non-disadvantaged households in Thailand (Thai Health Promotion Foundation, 2020).

Looking at another angle of the argument, academic performance pressure from parents can stress out children in nuclear families. This is true even if with single parent families. For example, Deb et al. (2015) indicated that the educational levels of the parents, especially student of fathers possessing a relatively low-education level, tend to perceive more pressure from the parent(s) for better academic performance in Kolkata, India. In South Korea where shadow education was extremely popular, Jarvis et al. (2022) found that children with single parent and unstable two-parent married families would have less shadow education than those in stereotyped nuclear families. This may also due to the matter of the overall household income since a family with both parents is most likely to earn more than a family with a single parent.

Nonetheless, many studies (e.g., Yang & Jiang, 2023) still argued that children staying in non-nuclear families seemed to have more behavioral and mental health problem than those living with nuclear families. In Thailand, even if the stress of expectation to enter public universities for high-school students is not as high as in many other non-western countries, it is worthwhile to pay explore stress-level of high-school students in Thailand.

Focusing at the mental health issue, this current research study aimed to explore how high-school students living with nuclear families or non-nuclear families dealt with stress via tourism behavior. Since Thailand is one of the resourceful countries with many tourist attractions, it is interesting to explore whether children in high-school age use tourism as a tool to help relaxing themselves from stress. There have been a number of past researches studies that explored the relationships among motivation, satisfaction, positive emotion and tourist behavioral intention (Bayih & Singh, 2020; Asmelash & Kumar,

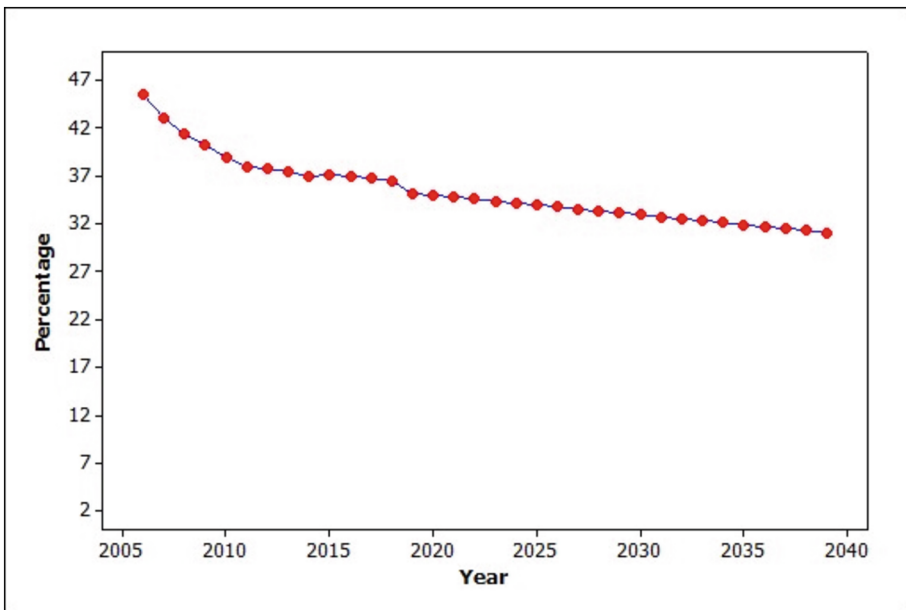


Fig. 1. The percentage of nuclear families in Thailand and its prediction

2019; Munar & Jacobsen, 2014; Pestana et al., 2020; Otoo et al., 2021; Zhang & Walsh, 2020; Kucukergin & Gurlek, 2020). This study adopted questions from the existing literatures and intended to measure and analyze the effect of family structure on high-school students' stress and enthusiasm in learning via tourism questionnaire.

2 Data Collection and Research Methodology

2.1 Questionnaire Survey

Participants were twelfth-grade students from science-and-mathematics program since we assume that this is in general the most difficult program in any typical high school. The sample group was students from a public school in Ayutthaya, a province in Thailand. The questionnaire was distributed via an online platform. The participants were from various family structure backgrounds. In this study, we classified family structure into two types consisted of a nuclear family and a non-nuclear family. Five-point Likert scale was used in the questionnaire. Nine questions (independent variables) adapted from existing literatures (Bayih & Singh, 2020; Zhang & Yin, 2020; Diallo et al., 2022; Lee & Jan, 2022; Gardiner et al., 2022; Pestana et al., 2020) were used to ask students in this study as shown below.

X_1 : Did you travel to escape from your study?

X_2 : Did you travel to relax yourself physically and mentally?

X_3 : Did you travel to get freedom?

X_4 : Did you search to travel to a new destination?

X_5 : During the trip, was it interesting for you to learn about different cultures and lifestyles?

X_6 : During the trip, did you often escape from everything and do self-exploration?

X_7 : During the trip, was it interesting for you to learn any new thing?

X_8 : Did you travel to get a once-in-a-lifetime experience?

X_9 : Did you like nature?

SPSS software was used to provide descriptive statistics, multicollinearity test and binary logistic regression. The descriptive statistics in this study shows the age, gender, tourism frequency (Zhang et al., 2020), companion (Fenitra et al., 2022) and family structure (Randall, 2021; Hoffmann, 2023; Khudyakova et al., 2016).

2.2 Multicollinearity Test

The multicollinearity test was used to check the subsistence of correlation among nine independent variables. The formulae for variance inflation factor (*VIF*) and tolerance (*TOL*) were shown in Eqs. (1) and (2).

$$VIF = \frac{1}{1 - R_i^2} \quad (1)$$

$$TOL = 1 - R_i^2 \quad (2)$$

From Eqs. (1) and (2), R_i^2 is the unadjusted coefficient of determination by regressing the i^{th} independent variable on all other independent variables. Generally, if the value of

VIF is more than the threshold value of 10, it is an indication for high multicollinearity. Likewise, if the value of *TOL* is less than the threshold value of 0.1, this implies that the variables are correlated (Hair et al., 2019).

2.3 Logistic Regression Model

Logistic Regression model was chosen as a model to fit the data in this study since the response was binary (nuclear family vs. non-nuclear family) (Kleinbaum & Klein, 2010; Hilbe, 2015). The binary logistic regression model is as shown in Eq. (3) below.

$$p = \frac{e^z}{1 + e^z} \quad (3)$$

From Eq. (3), p is the probability of occurrence which range between zero and one, e is an exponential function, and z is a linear combination as shown below:

$$z = \beta_0 + \sum_{i=1}^n \beta_i X_i \quad (4)$$

From Eq. (4), X_i = independent variable ($i = 1, 2, 3, \dots, n$), β_0 is a constant, and β_i = coefficient of logistic regression model representing the weight of X_i ($i = 1, 2, 3, \dots, n$).

In addition Eq. (5) defines the odds as:

$$odds = \frac{p}{1 - p} \quad (5)$$

By taking the natural logarithm of Eq. (5), this yields Eq. (6) as follow.

$$\text{logit} = \log(odds) = \ln\left(\frac{p}{1 - p}\right) = z = \beta_0 + \sum_{i=1}^n \beta_i X_i \quad (6)$$

The range for the above logit function is between negative infinity and infinity. Recently, logistic regression model analysis has been applied in various studies (Sun et al., 2018; Sahani & Ghosh, 2021; Khan et al., 2020; Thakur et al., 2022; Jawa, 2022; Hashemi et al., 2022).

3 Result and Discussion

In this study, 81 twelfth-grade students (mean age: 17.53 years) answered the online questionnaire. The results indicated that the most of the respondents were female (54.3%), followed by male (29.6%) and LGBT (12.3%). Most of the respondents were found in the tourism frequency between 1 and 3 time per year (50.6%). The descriptive statistics are shown in Table 1 below.

Table 1. The descriptive statistics of respondents

Sample n = 81 Items (characteristics)	Frequency	Percentage (%)
Gender		
Male	24	29.6
Female	44	54.3
LGBT	10	12.3
Others	3	3.7
Tourism frequency		
1 to 3 times/year	41	50.6
4 to 6 times/year	27	33.3
More than 6 times/year	13	16.0
Companion		
Alone	8	9.9
Family	55	67.9
Friends	11	13.6
Other	7	8.6
Do you live with nuclear family?		
Yes	64	79.0
No	17	21.0

Likert scale was used to explain the main variables in the study. Table 2 showed the percentages of respondents in their attitudes and the means of the nine questions. The majority of the participants answered that they want to relax physically and mentally (67.9% strongly agree); they want to travel to get a once-in-a-lifetime experience and they want to explore nature (60.5% strongly agree); they want to learn new thing (58.0%). Some participants showed positive attitude about freedom and the exploration of new destination (53.1%).

The multicollinearity test of the independent variables was examined by the *VIF* and *TOL*. The results in Table 3 showed that the *VIF* values were lower than the threshold value of 10. In addition, the *TOL* values exceeded the recommended threshold value of 0.1, showing no collinearity exists among all the independent variables.

Binary logistic regression analysis was used to analyze the impact of independent variables on the dependent variable (whether students live with both of their parents or not). The results are shown in Table 4 in which there are only three independent variables including X_1 , X_2 and X_5 that have statistically significant impact on the students family structure (at a 93.5% confidence level). Note that even though the p-value for X_2 is a little over 0.05 that makes it impossible to achieve 95% confidence level in the first model, it is believed to be sufficiently small to be included in the final model. Table 5. Shows the results of the final binary logistic regression model with only three significant independent variables and Eq. (7) shows the equation of the final model.

$$\text{logit} = 2.524 - 0.791X_1 - 0.724X_2 + 1.322X_5 \quad (7)$$

Table 2. Percentages of respondents regarding their attitudes toward.

Independent variables	Mean	Standard deviation	Attitudes				
			Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
X_1	3.57	1.30	11.1	7.4	25.9	24.7	30.9
X_2	4.41	1.00	2.5	2.5	14.8	12.3	67.9
X_3	4.22	1.02	3.7	1.2	17.3	24.7	53.1
X_4	4.19	1.03	1.2	6.2	18.5	21.0	53.1
X_5	3.98	1.02	2.5	3.7	27.2	27.2	39.5
X_6	3.95	1.09	4.9	2.5	24.7	28.4	39.5
X_7	4.30	0.98	2.5	1.2	18.5	19.8	58.0
X_8	4.40	0.88	1.2	1.2	14.8	22.2	60.5
X_9	4.31	1.02	2.5	3.7	14.8	18.5	60.5

Table 3. Multicollinearity test of the independent variables used in this study.

Independent Variable	Collinearity Statistics	
	Tolerance	VIF
X_1	0.688	1.454
X_2	0.491	2.036
X_3	0.442	2.265
X_4	0.254	3.931
X_5	0.279	3.588
X_6	0.273	3.665
X_7	0.242	4.139
X_8	0.401	2.491
X_9	0.298	3.351

The Wald test of variable X_5 indicated that learning about different cultures and lifestyles affected significantly on the students living with nuclear family since p-value was much smaller than the significant level. The odds ratio of the students who live with nuclear families was 3.751 units with 95% confident interval (1.682, 8.365) compared to that of the students who live with non-nuclear families. This result implied that from the collected samples, students who lived with nuclear families were more likely to be enthusiastic to learn anything outside of their classrooms more than students in the other group.

However, since the odds ratios of variables X_1 and X_2 were less than 1 (odds ratio of $X_1 = 0.453$ and odds ratio of $X_2 = 0.485$), this implied reverse relationship between the question and the response. Hence, the students who lived with non-nuclear families had a higher tendency to need travel to escape from their studies and higher need to relax themselves physically and mentally during their travelling compared to the students living with nuclear families.

Table 4. Coefficients of binary logistic regression model with all independent variables

Independent Variable	Coefficient	S.E	Wald	df	p-value	Odds ratio	95% CI for Odds ratio	
							Lower	Upper
Constant	2.207	2.004	1.023	1	0.312	7.590	0.156	0.954
X ₁	-0.952	0.461	4.253	1	0.039**	0.386	0.139	1.063
X ₂	-0.957	0.519	3.397	1	0.065***	0.384	0.819	6.885
X ₃	0.865	0.543	2.536	1	0.111	2.375	0.134	2.462
X ₄	-0.555	0.743	0.558	1	0.455	0.574	2.042	24.684
X ₅	1.960	0.636	9.505	1	0.002*	7.100	0.157	2.235
X ₆	-0.524	0.677	0.597	1	0.440	0.592	0.322	5.858
X ₇	0.318	0.740	0.184	1	0.668	1.374	0.505	4.845
X ₈	0.448	0.577	0.602	1	0.438	1.564	0.150	1.766
X ₉	-0.665	0.629	1.115	1	0.291	0.514		

Dependent variable was the students who live with nuclear family (Binary; Yes = 1 and No = 0)

Note: * p-value < 0.01; ** p-value < 0.05; *** p-value < 0.10

Table 5. Coefficients of binary logistic regression model with significant independent variables.

Independent Variable	Coefficient	S.E	Wald	df	p-value	Odds ratio	95% CI for Odds ratio	
							Lower	Upper
Constant	2.524	1.935	1.702	1	0.192	12.482	0.218	0.953
X ₁	-0.791	0.373	4.486	1	0.034**	0.453	0.211	1.115
X ₂	-0.724	0.425	2.904	1	0.088***	0.485	1.682	8.365
X ₅	1.322	0.409	10.439	1	0.001*	3.751		

Dependent variable was the students who live with nuclear family (Binary; Yes = 1 and No = 0)

Note: * p-value < 0.01; ** p < 0.05; *** p < 0.10

The Cox& Snell R Square and Nagelkerke R Square were equal to 0.191 and 0.298, respectively, showing that the data were well-explained by the model. Moreover, with our percentage of predicted probability of 84%, this means that the response was properly predicted by the independent variables (using a 50% criteria).

Table 6 shows the contingency obtained from the Hosmer-Lemeshow test. In this table, the observed and expected numbers of nuclear family and non-nuclear family were shown. Then from Table 6, the Hosmer and Lemeshow goodness-of-fit statistics was calculated base on Eq. (8). In this equation, it can be explained that if the differences between the expected numbers and the observed numbers are small, the model fits well.

$$\chi^2 = \sum_{i=10}^{10} \frac{(Observed_i - Expected_i)^2}{Expected_i} \tag{8}$$

After obtaining the chi-square statistics in Eq. (8), the following hypotheses were tested:

H₀: the model was a goodness of fit.

H₁: the model was not a goodness of fit.

The result in Table 7 indicates that the p-value for the Hosmer-Lemeshow test was 0.924 which was greater than significance level of 0.05. Therefore, there were no difference between the expected values and observed values. Thus, the model in this study fits well.

Table 6. Contingency results from the Hosmer-Lemeshow test.

	Non-nuclear family		Nuclear family		Total
	Observed	Expected	Observed	Expected	
1	5	5.551	3	2.449	8
2	3	2.601	3	3.399	6
3	2	2.248	7	6.752	9
4	0	0.431	2	1.569	2
5	4	3.293	15	15.707	19
6	2	1.007	5	5.993	7
7	1	0.971	7	7.029	8
8	0	0.623	8	7.377	8
9	0	0.223	8	7.777	8
10	0	0.052	6	5.948	6

Table 7. Hosmer-Lemeshow test

Chi-square	df	p-value
3.157	8	0.924

4 Conclusion

From the intention to measure mental health issue of high-school students (12th grade) via tourism-related questionnaire, it was interesting to observe that students staying with non-nuclear families tend to be more stressful than those staying with nuclear families. Another interesting finding in our study was that students living with nuclear families tend to be more enthusiastic to learn new things even during their leisure travelling.

Not only that our results were consistent with many existing studies, we also showed an interesting approach by attempting to use opinions about leisure behavior to measure stress level in students living in two big different types of family structure. As discussed in the introduction, since the number of nuclear families in Thailand is gradually decreasing, in the following years, schools may expect to have higher level of stress among students

especially students in the 12th grade who are expecting to have a big transition in their lives by either attending colleges or leaving school for work. One suggestion that could be adopted from our study is that schools may establish a program to support students who live with non-nuclear family.

Nonetheless, the number of samples collected in this study was relatively small. Also, it was collected from only one province in Thailand. It could be interesting to collect more samples from the other provinces, especially Bangkok which is the capital city of Thailand with highly-densed population. Another extension to this study was by collecting other information regarding students' background such as family income, after-school activity, parent(s)/guardian occupation, parent(s)/guardian education, peer-pressure related questions, etc.

References

- Asmelash, A.G., Kumar, S.: The structural relationship between tourist satisfaction and sustainable heritage tourism development in Tigray, Ethiopia. *Heliyon* **5**, e01335 (2019)
- Bayih, B.E., Singh, A.: Modeling domestic tourism: motivations, satisfaction and tourist behavioral intentions. *Heliyon* **6**, e04839 (2020)
- Deb, S., Strodl, E., Sun, J.: Academic stress, parental pressure, anxiety and mental health among Indian high school students. *Int. J. Psychol. Behav. Sci.* **5**(1), 26–34 (2015)
- Diallo, M.F., Diop-Sall, F., Leroux, E., Vachon, M.: How do tourism sustainability and nature affinity affect social engagement propensity? The central roles of nature conservation attitude and personal tourist experience. *Ecol. Econ.* **200**, 107503 (2022)
- Fenitra, R.M., Premananto, G.C., Sedera, R.M.H., Abbas, A., Laila, N.: Environmentally responsible behavior and Knowledge-Belief-Norm in the tourism context: the moderating role of types of destinations. *Int. J. Geoheritage Parks* **10**(2), 273–288 (2022)
- Gardiner, S., Vada, S., Yang, E.C.L., Khoo, C., Le, T.H.: Recreating history: the evolving negotiation of staged authenticity in tourism experiences. *Tour. Manag.* **91**, 104515 (2022)
- Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E.: *Multivariate Data Analysis*, 8th edn. Cengage Learning EMEA (2019)
- Hashemi, H., Bargegol, I., Hamed, G.H.: Using logistic regression and point-biserial correlation, an investigation of pedestrian violations and their opportunities to cross at signalized intersections. *IATSS Res.* **46**, 388–397 (2022)
- Hilbe, J.M.: *Practical Guide to Logistic Regression*. CRC Press Taylor & Francis Group (2015)
- Hoffmann, J.P.: Family structure, unstructured socializing, and delinquent behavior. *J. Crim. Just.* **87**, 102086 (2023)
- Jarvis, J.A., Read, A.R., Dufur, M.J., Pribesh, S.: Impacts of family structure on shadow education and educational achievement among South Korean youth. *Int. J. Educ. Dev.* **89**, 102529 (2022)
- Jawa, T.M.: Logistic regression analysis for studying the impact of home quarantine on psychological health during COVID-19 in Saudi Arabia. *Alex. Eng. J.* **61**, 7995–8005 (2022)
- Khan, S., Halder, H.R., Rashid, M., Afroja, S., Islam, M.: Impact of socioeconomic and demographic factors for underweight and overweight children in Bangladesh: a polytomous logistic regression model. *Clin. Epidemiol. Glob. Health* **8**(4), 1348–1355 (2020)
- Khudyakova, T.L., Gridyaeva, L.N., Klepach, Y.V.: Specific features of gender identity formation in children from single-parent and nuclear families in ontogenesis. *Proc. Soc. Behav. Sci.* **233**, 393–396 (2016)
- Kleinbaum, D.G., Klein, M.: *Logistic Regression a Self-Learning Text*, 3rd edn. Springer (2010)
- Kom Chad Luek News. Only 28% of Thai kids live in nuclear families, support locals and change of care giving styles (2019). <https://www.komchadluek.net/scoop/370171>

- Kucukergin, K.G., Gurlek, M.: 'What if this is my last chance?': developing a last-chance tourism motivation model. *J. Destin. Mark. Manag.* **18**, 100491 (2020)
- Lee, T.H., Jan, F.: How does tourism image affect visitor perceptions of a festival's identity? Influence analysis of three aboriginal festivals in Taiwan. *J. Destin. Mark. Manag.* **24**, 100704 (2022)
- Munar, A.M., Jacobsen, J.Kr.S.: Motivations for sharing tourism experiences through social media. *Tourism Manag.* **43**, 46–54 (2014)
- Otoo, F.E., Kim, S., King, B.: African diaspora tourism - How motivations shape experiences. *J. Destin. Mark. Manag.* **20**, 100565 (2021)
- Pestana, M.H., Parreira, A., Moutinho, L.: Motivations, emotions and satisfaction: the keys to a tourism destination choice. *J. Destin. Mark. Manag.* **16**, 100332 (2020)
- Randall, B.A.: School administrators' leadership practices with non-nuclear families: a phenomenological study. Doctor of Education thesis. Liberty University, Lynchburg (2021)
- Sahani, N., Ghosh, T.: GIS-based spatial prediction of recreational trail susceptibility in protected area of Sikkim Himalaya using logistic regression, decision tree and random forest model. *Eco. Inform.* **64**, 101352 (2021)
- Samutchak, P., Rittirong, J., Chuanwan, S., Tianlai, K., Nitnara, P., Somta, S.: Future Thai Families in 2040. Mahidol University, Bangkok, Thailand (2022)
- Sun, L., Miao, C., Yang, L.: Ecological environmental early-warning model for strategic emerging industries in China based on logistic regression. *Ecol. Ind.* **84**, 748–752 (2018)
- Thai Health Promotion Foundation. Worry about the increase in the number of non-nuclear families (2020). <https://www.thaihealth.or.th/ห่วงครั้วเรือนแห่งกลาง/>
- Thakur, S., Ranjitkar, P., Rashidi, S.: Investigating evacuation behaviour under an imminent threat of volcanic eruption using a logistic regression-based approach. *Saf. Sci.* **149**, 105688 (2022)
- Yang, Y., Jiang, J.: Influence of family structure on adolescent deviant behavior and depression: the mediation roles of parental monitoring and school connectedness. *Publ. Health* **217**, 1–6 (2023)
- Zhang, J., Walsh, J.: Tourist experience, tourist motivation and destination loyalty for historic and cultural tourists. *Pertanika J. Soc. Sci. & Hum.* **28**(4), 3277–3296 (2020)
- Zhang, T., Yin, P.: Testing the structural relationships of tourism authenticities. *J. Destin. Mark. Manag.* **18**, 100485 (2020)
- Zhang, Z., et al.: Analysis of the island tourism environment based on tourists' perception—a case study of Koh Lan, Thailand. *Ocean Coastal Manag.* **197**, 105326 (2020)



Towards Education 4.0: A Holistic Approach for Performance Evaluation of German and Thai Universities

Jettarat Janmontree¹, Hartmut Zadek¹, and Woramol C. Watanabe²(✉)

¹ Institute of Logistics and Material Handling Systems, Faculty of Mechanical Engineering, Otto-Von-Guericke University, Magdeburg, Germany
jettarat.janmontree@ovgu.de

² Department of Logistics and Digital Supply Chain, Naresuan University, Phitsanulok, Thailand

Abstract. This paper presents a framework to evaluate the readiness and performance of educational systems concerning Education 4.0. The proposed framework integrates the concept of Analytic Hierarchy Process (AHP) and Data Envelopment Analysis (DEA) methods. Pairwise comparisons were conducted by experts in education and academic areas from Germany and Thailand. After that, the DEA method was used to determine the performance value based on selected criteria. The result showcases the benchmark values of various universities from both Germany and Thailand. Overall both countries demonstrate good performance regarding Education 4.0. Nonetheless, one university from each country shows lower performance which could be improved by focusing more on the area of technology skills and personalized learning. The result of this paper contributes to improving the process of evaluating the performance and readiness of universities towards Education 4.0.

Keywords: Education 4.0 · Performance Measurement · Education System

1 Introduction

Nowadays, information technology plays a crucial role in the educational system. The infamous term “Industry 4.0” was coined to describe the integration of advanced technologies into the physical and biological worlds (Hussin, 2018). Led by technologies such as Artificial Intelligence, the Internet of Things, robotics, and mobile connectivity, Industry 4.0 enables new innovative approaches in communication and collective work. The emergence of Industry 4.0 has affected not only the industry but also everyone involved to a great extent. Workers in the industry are now expected to be equipped with the right skill sets that enable them to develop, operate, monitor, maintain, and manage. (Ellingrud et al. 2020).

Influenced by the concept of Industry 4.0, Education 4.0 integrates advanced technology and new concepts in learning and teaching in order to equip learners with the right skill sets. Education 4.0 is defined as the transformation of the educational system using advanced technologies to foster, promote, and enable new possibilities for learners

(Ramírez-Montoya et al. 2022). It has shifted the learning and teaching process from oral dictation and one-way communication to a less rigid, more innovative, and process-oriented approach (Agrawal et al. 2021). Furthermore, the concept of Education 4.0 also suggests that teachers and instructors will benefit by utilizing advanced technologies to create lessons more flexibly and creatively. (Hussin, 2018). Skill sets and learning methods are two main focuses in Education 4.0 (WorldEconomicForum, 2023).

Oliveira and Souza (2021) identified skill sets that are necessary to solve real-life and work-related problems. These skill sets range from typical soft skills such as communication, creativity, and social and cultural awareness, to technology-related hard skill sets, including digital technology design, technology resource management, and computational thinking. Each student is an individual, and there is no single teaching method that can best educate all students. In order to equip students with a wide range of skill sets, new learning methods must be implemented. The concept of blended learning models, such as the flipped classroom introduced in the 1980s, has recently gained much attention. It requires students to prepare for each lesson beforehand and allows them to be more flexible in learning at their own pace (Sierra-Fernández et al. 2023). This learning model helps promote their critical thinking skills and reinforces their engagement in the classroom (Hew and Lo, 2018; Nouri, 2016). Furthermore, advanced technology such as smart devices has been integrated heavily in Education 4.0. It can be used to monitor, control, optimize, and automate to create a better learning system and experience. A more recent development involves introducing Artificial Intelligence (AI) and Machine Learning (ML) into teaching methods. Haderer and Ciolacu (2022) proposed a system that integrates AI into learning and task planning. This system generates individual plans for each student, allowing them to monitor their learning progress and forecast learning success.

It is certain that we have entered the era of Education 4.0 for many years. However, universities and higher education institutes are struggling to adapt to new changes. Alongside the disruption caused by the global pandemic in 2020, many universities have undertaken a comprehensive reassessment of their practices with a focus on establishing a more robust and resilient educational system. Since Education 4.0 consists of many elements, transforming the education system into Education 4.0 is a challenging task. The rapid development of advanced technologies has a significant influence on higher education. Integrating new technologies into teaching methods and designing curricula for Education 4.0 requires significant effort. Therefore, it is important for universities to evaluate and assess their performance before developing an action plan toward Education 4.0.

2 Literature Review

2.1 History of Education 4.0

Education has always been influenced by the current trends in society. It was designed to prepare learners by arranging lessons into multiple tiers where the complexity of content is increased. Higher education represents the highest tier of knowledge, aiming for students to develop deductive reasoning, abstract thought, systematic problem-solving skills, metacognitive abilities, and the capacity to cultivate hypothetical or counterfactual

thinking (Jones and Sharma, 2021). The education system consists of four main variables, including content, students, teachers, and institutions. *Students* seek education (*content*) from *teachers* through methods such as lecturing, which is the most common form of passing on knowledge. Students are rewarded with grades or graduation degrees as proof of their qualifications for the job market. These activities are managed and overseen by *institutions* such as universities. Besides internal factors, the education system is often influenced by external factors such as technology, politics, the economy, or the requirements of the job market (González-Pérez and Ramírez-Montoya, 2022). Although there is no concrete evidence regarding the timeline of the development from Education 1.0 to Education 4.0, many researchers have studied the influence of the *Industrial Revolution* and the evolution of the education system (Miranda et al. 2021).

The advancements in technologies such as 5G, online learning platforms, cloud computing, machine learning, and artificial intelligence marked the dawn of the new era known as Industry 4.0. This term was coined in 2016 by Klaus Schwab. Industry 4.0 is closely linked to Education 4.0 and has strengthened the concept of learning and teaching to be more individualized and effective. Connectivity and collaboration between individuals became easier. Students were no longer solely dependent on teachers, but rather driven by self-determination and engaged in highly autonomous self-learning. What was once confined to traditional classroom settings has now expanded into a vast learning space that spans across the globe and shifted the teaching method from a teacher-centric approach to a more interactive one, fostering greater communication between teachers and students.

2.2 Analytical Hierarchy Process (AHP)

The Analytical Hierarchy Process (AHP) is a Multi-Criteria Decision Making (MCDM) tool that is widely used due to its flexibility and functionality (Janmontree, 2021). The main advantage of AHP is its capability to evaluate qualitative problems quantitatively (Raco et al. 2022). Furthermore, the approach of AHP can be integrated and combined with other techniques such as the Delphi method, Linear Programming (LP), or Data Envelopment Analysis (Ho, 2008; Janmontree, 2021). AHP works by breaking down decision-making problems into criteria that contribute to and influence the outcome. The problems are structured in multiple hierarchical levels where decision-makers can compare each criterion using pairwise comparisons. (Ho, 2008). Given that the concept of Education 4.0 is broad and can be subjective, several researchers have utilized AHP to identify the significant factors of Education 4.0 (Raco et al. 2022; Ulufer et al., 2023).

Generally, there are three steps in performing AHP. First, the criteria related to the problem and overall goal must be identified. Two-level hierarchy is a commonly used format. Next, experts engage in pairwise comparisons to assess the relationships between each criterion. Saaty's scale (ranging from 1 to 9) is typically used, denoting degrees of importance from equally important to extremely more important. After verifying the consistency of these judgments, the global and local weights of each criterion are calculated based on the hierarchy's structure. This process is then extended to compare each alternative. Finally, the priority of the criteria and attributes of each alternative is determined (Ho, 2008).

2.3 Application of DEA in Education Sector

Data Envelopment Analysis (DEA) is a tool used to assess the efficiency and performance of educational institutions within the education sector. The unique attributes of universities which are involving multiple inputs and/or outputs and the absence of market prices. And the results of education depend on what society thinks, there's a point that considers the outcomes acceptable, like good quality, learning well, and fairness. These reasons position DEA as a more dependable technique compared to other parametric methods (Barra and Zotti, 2016; Berbegal-Mirabent, 2018; Greene, 1980). DEA effectively transforms input into outputs, treating the process as a 'black box' (Chaowarat et al. 2014; Patitad and Watanabe, 2022) and identifying inputs that aren't optimally utilized (Shero et al., 2022). By employing DEA, it becomes possible to identify performance gaps, determine weaknesses, and facilitate benchmarking (Aparicio and Monge, 2022).

A study used DEA to discern the impact of high and low-quality teachers on school performance (Santín and Sicilia, 2018). In addition, DEA has been utilized to evaluate the technical efficiency of prominent public universities, with a focus on teaching and research endeavors (Barra & Zotti, 2016). Based on the previous research, it is evident that DEA is an unequivocally practical tool for measuring performance within the education sector.

This study focuses on Education 4.0 and how learners' skills and teaching methods interact using advanced technology. This mix leads to outcomes that are discernibly subjective character. Consequently, AHP was selected as a clarifying the output of Education 4.0, as discussed earlier.

Conversely, the inputs used in education are relatively stable and concrete. Considering this context, it makes sense to begin by examining the inputs, which pave the way for a more thorough evaluation of the wider educational picture. As highlighted earlier, a considerable number of research initiatives have been dedicated to investigating DEA in the educational section. The summary of the input is described as follows;

- Expenditure, referring to the municipal funding directed towards the university (Haelermans and Ruggiero, 2013; Mayston, 2014; Mota and Meza, 2020);
- Human resource, indicating the count of university lecturers (Endovitsky and Komendenko, 2022; Mota and Meza, 2020; Santín and Sicilia, 2018); and
- Support resources, denoting the number of faculties and members at the university (Alam et al. 2019; Firsova and Chernyshova, 2019; Zoghbi et al. 2013).

3 Methodological Framework

The research proposed a methodological framework to evaluate the performance and readiness of universities concerning Education 4.0. It consists of three phases. Firstly, literature research and expert interviews were conducted to determine the initial criteria relevant to Education 4.0. After that, the unimportant criteria were neglected using AHP. A group of experts consists of members from academic institutes in Germany and Thailand. The selected criteria were used as output indicators. The DEA method was then employed to determine the performance value of each university based on input-output indices. Finally, the results of each university were compared and recommendations

for enhancing performance toward Education 4.0 were proposed. Figure 1 provides an illustration of the methodology adopted in this research.

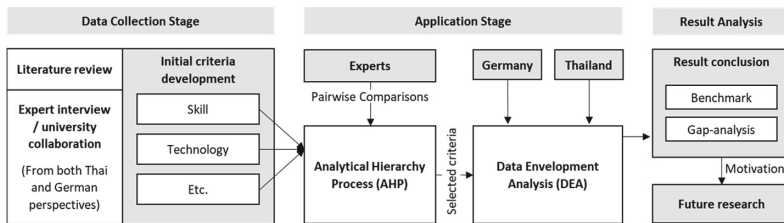


Fig. 1. Methodological framework

4 Result

4.1 Development of Education 4.0 criteria

According to the World Economic Forum, Education 4.0 should focus on developing diverse skills in students through well-designed content, utilizing innovative teaching methods, and integrating advanced technology into teaching (WorldEconomicForum, 2023). There are elements that define Education 4.0 that cover both quantitative and qualitative aspects. In this study, a literature review was conducted to identify criteria concerning Education 4.0. Table 1 shows the summary and brief description of each criterion.

As seen in Table 1, there are in total 15 criteria covering 4 areas of education 4.0. The first category of Education 4.0 is Teaching methods. It focuses mainly on methods and techniques used in teaching. Using creative and teaching methods can greatly enhance both the learning and teaching experience.

The next category is Infrastructure and Supporting facilities. University plays a crucial role in Education 4.0 as an intermediate part that connects other players in the system. Universities are also responsible for providing the supporting infrastructure necessary for a seamless practice of Education 4.0. In contrast to the past where classrooms were the sole platform for teacher-student interaction, the digital era has introduced alternative options like online classrooms and virtual laboratories.

Education 4.0 advocates for pedagogical changes, moving from traditional lectures towards more student-centric methods. This change requires effective and efficient communication systems. Apart from the institutional level, advanced technologies used in classrooms are crucial. These technologies enable teachers to monitor students' development and analyze results in real time, reinforcing creative teaching methods like personalized and self-paced learning.

Given the rapid change in advanced technologies, students need to develop not only basic knowledge but also advanced skill sets and competencies. The last category centers on creating new curricula and content that foster the development of skills and competencies that cover both basic and advanced technology skills. This should help

Table 1. Summary of criteria concerning Education 4.0

Category	Criteria	Short Description
Teaching Method	Distance learning	Distance learning enables the ability for students and teacher to conduct and participate in each lesson anywhere (Akimov et al. 2023; González-Pérez and Ramírez-Montoya, 2022; Mansor et al. 2020; Rienties et al.; WorldEconomicForum, 2020)
	Personalized learning	Personalized learning focuses on developing teaching methods to be more flexible, supporting students from different backgrounds and enabling them to learn at their own interests and pace. (Mansor et al. 2020; Rienties et al.; WorldEconomicForum, 2020)
	Project-based learning	Project-based learning promotes essential skills for students, including collaborative skills, project management, critical thinking, and problem-solving skills (Akimov et al. 2023; González-Pérez and Ramírez-Montoya, 2022; Mansor et al. 2020; Rienties et al.; WorldEconomicForum, 2020)
	Hands-on learning	Hands-on learning primarily focuses on internship programs or collaborations with companies. This approach allows students to learn from real-world experiences (Akimov et al., 2023; Mansor et al., 2020)
Infrastructure and supporting facilities	Communication Systems	Communication systems are fully developed to allow offline/online communication between teachers and students (Akimov et al. 2023; González-Pérez and Ramírez-Montoya, 2022)

(continued)

Table 1. (continued)

Category	Criteria	Short Description
	University Infrastructures	University focus on developing smart and sustainable infrastructure. Internet of Things, smart boards, or sensors are integrated into teaching and learning processes (Akimov et al. 2023; González-Pérez and Ramírez-Montoya, 2022)
	Software and Technical Support	University should provide support for all necessary hardware and software related to education (Akimov et al., 2023; González-Pérez and Ramírez-Montoya, 2022; Mansor et al. 2020; Rienties et al.; WorldEconomicForum, 2020)
Advanced teaching tools	Interactive teaching materials	Interactive teaching materials such as remote or virtual laboratories, augmented reality, interactive books, or 3D models are utilized in the classroom (Mansor et al. 2020)
	Advanced teaching analytic	University uses AI or ML in teaching such as enrolment forecasting, or predicting the outcome of students. These results are used to adapt and redesign lessons to better suit students (González-Pérez and Ramírez-Montoya, 2022)
	Online teaching tools	University ensures that teachers and students have access to essential tools and platforms for online teaching (González-Pérez and Ramírez-Montoya, 2022; Mansor et al. 2020)
	Safety-system	University focuses on a safety system in Education 4.0. Safety measures and protocols were developed to mitigate potential risks for both hardware and software aspects (Mansor et al. 2020)

(continued)

Table 1. (continued)

Category	Criteria	Short Description
Educational and curriculum design	Technology skills	New curricula should be developed to incorporate content that promotes digital skills such as programming, digital responsibility, and the use of technology (González-Pérez and Ramírez-Montoya, 2022; Mansor et al. 2020; WorldEconomicForum, 2020)
	Basic knowledge	Students should be encouraged to acquire and learn basic knowledge, including reading and writing, mathematical proficiency, financial literacy, natural science knowledge, information and computer skills, cultural awareness, and civic literacy (Akimov et al. 2023; González-Pérez and Ramírez-Montoya, 2022)
	Innovative skills	Curricular should include content that fosters innovative and creative skills. It should promote critical thinking, innovative problem-solving skills, and the ability to work in teams and collaborate effectively (WorldEconomicForum, 2020)
	Global citizenship skills	Students should learn to be aware of the global community, focus on sustainability, and develop a sense of responsibility to take on an active role in the future (González-Pérez and Ramírez-Montoya, 2022; WorldEconomicForum, 2020)

promote lifelong learning that prepares students for the future and teaches them how to use technology responsibly.

4.2 Analytical Hierarchy Process (AHP)

To determine the important criteria concerning Education 4.0, The AHP technique was used. Based on the initial criteria from the previous section, a hierarchical structure of the criteria was created (Fig. 2).

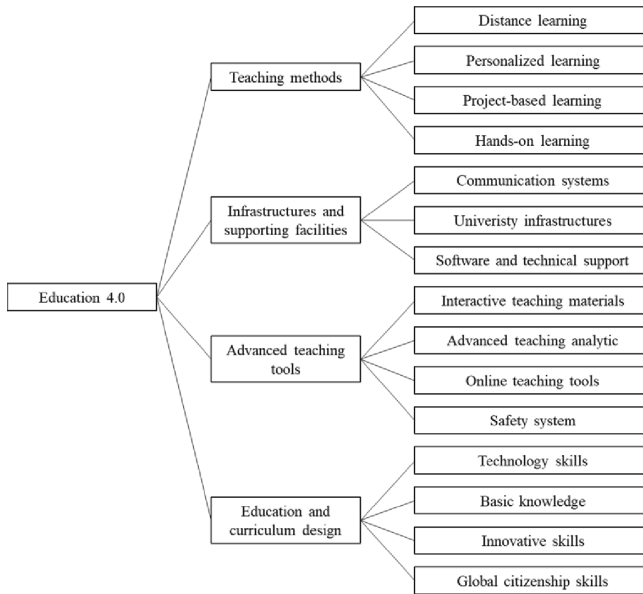


Fig. 2. AHP Hierarchical Structure

Questionnaires in the form of pairwise comparisons between each criterion were sent out to experts. In total, 6 experts, representing both Germany and Thailand, participated in the survey. After the pairwise comparison of each expert, a consistency check was performed to verify whether the judgment was consistent. The global and local weights of each criterion are shown in Table 2. Using the Pareto principle, unimportant criteria were neglected. The selected criteria, highlighted in Table 2, were used as an output indicator for the next step.

4.3 DEA Model

The DEA model requires the establishment of a realistic set of input-output indices. The literature review in Sect. 2.3 identifies core input measures, including finance, human resources, and support factors. Based on the result from AHP, eight output indicators are selected using the Pareto principle. Taking these insights and data availability into account, this paper constructs an efficiency index structure for Education 4.0, considering both input and output dimensions. This model consists of two primary-level metrics, seven secondary-level metrics, and twelve tertiary-level metrics. The framework is shown in Table 3.

4.4 Performance Evaluation

The dataset utilized for conducting DEA, which includes both input and output variables, is illustrated in Table 4. This dataset incorporates data from five universities, including three from Germany and two from Thailand. The input data has been procured

Table 2. AHP weight calculation

Main Criteria	Weight	Sub Criteria	Local weight	Global weight
Teaching methods	0.373	Distance learning	0.121	0.045
		Personalized learning	0.406	0.152
		Project-based learning	0.286	0.107
		Hands-on learning	0.187	0.070
Infrastructure & supporting facilities	0.155	Communication systems	0.348	0.067
		University infrastructures	0.166	0.039
		Software & technical support	0.236	0.049
Advanced teaching tools	0.184	Interactive teaching material	0.405	0.087
		Advanced teaching analytic	0.181	0.045
		Online teaching tools	0.076	0.026
		Safety-system	0.081	0.027
Education and Curriculum Design	0.288	Technology skills	0.384	0.111
		Basic knowledge	0.283	0.081
		Innovative skills	0.245	0.070
		Global citizenship skills	0.088	0.025

from university reports, while the output data was obtained using Education 4.0 score-cards. Additionally, in-depth interviews with university officials have been integrated to enhance the credibility of the output data.

To calculate the performance of each university regarding Education 4.0, the R programming language and the “DeaR” package were used. The chosen analytical approach takes on an output-oriented perspective, with the primary objective of revealing strategies to elevate the levels of output in alignment with the principles of Education 4.0, consequently refining operational efficiency. Acknowledging the inherent complexities linked to diminishing university inputs, the examination predominantly revolves around amplifying the specified outputs. The outcomes derived from this analytical process are summarized in Table 5.

4.5 University Comparisons

The efficiency scores shown in Table 5 represent the relative performance of different universities (DMUs). In the DEA output-oriented approach, values greater than 1 suggest

Table 3. DEA Model

Primary indicator	Secondary indicators	Sym	Third-level indicators
Input indicators	Expenditure	I1	Municipal funding directed towards the university
	Human resource	I2	Number of lecturers
	Support resource	I3	Number of faculties
I4		Number of staffs	
Output indicators	Teaching methods	O1	Level of personalized learning
		O2	Level of project-based learning
		O3	Level of hands-on learning
	Infrastructure & supporting facilities	O4	Level of communication systems
	Advanced teaching tools	O5	Level of interactive teaching material
	Educational and curriculum design	O6	Level technology skills
		O7	Level basic knowledge
		O8	Level innovative skills

Table 4. Data for DEA

DMU	Input indicators				Output indicators							
	I1	I2	I3	I4	O1	O2	O3	O4	O5	O6	O7	O8
G1	305.00	597	17	8,219	4	4	4	5	4	3	4	3
G2	563.20	351	9	7,243	5	4	5	5	5	4	4	5
G3	162.90	195	9	2,050	4	4	4	4	4	3	4	5
T1	63.70	1,506	21	5,037	4	3	5	4	3	2	3	3
T2	130.02	2.838	23	8,419	4	4	5	4	4	3	4	4

Table 5. Efficiency result

DMUs	G1	G2	G3	T1	T2
Efficiency score	1.502	1.000	1.000	1.000	1.048

inefficiency, as higher values indicate that the DMU can potentially increase outputs without requiring more of any observed input values to become more efficient.

Table 5 shows that three universities demonstrate operational efficiency, while the remaining two universities are lagging. A recurrent pattern observed within inefficient universities centers on their pronounced faculty and staff numbers, despite their capability to produce nearly equal levels of Education 4.0 outputs. Conversely, these universities with suboptimal performance could enhance their performance in specific Education 4.0 dimensions such as Technology Skills and Innovative Skills. By focusing exclusively on the inefficiencies, a significant comparison comes into focus, particularly between T1 and G1. Although the inputs remain almost identical except for funding and lecturer count (where G1 has more funding but fewer lecturers compared to T2), T2 sustains nearly comparable output levels. This leads to T2 achieving a higher efficiency score in this specific context (Fig. 3).

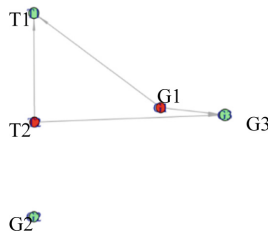


Fig. 3. University comparison

Figure 3 illustrates the projection of DMUs (universities). The green circle designates universities operating at an efficient level, while the red circle represents universities exhibiting inefficiency in their operations. In pursuit of enhancement towards efficacy, inefficient universities have the potential to align themselves with neighboring frontier universities. As an illustrative example, a T2 university can consider both T1 and G3 universities as benchmark institutions, endeavoring to elevate their own operational standards to mirror comparable levels of input and output as these benchmark entities.

5 Conclusion

One positive lesson we've gained from the recent global pandemic is the human ability to adapt and change. We were forced to change how we learn and teach altogether and therefore shift from the traditional learning methods. Students are the main focus of education 4.0. Technology plays a crucial part in Education 4.0, enabling effective connection, collaboration, and learning. This research focuses on evaluating the performance and readiness of universities in the context of Education 4.0.

Overall, both countries' universities exhibit good performance. Nonetheless, one university from each country lags behind. To enhance their performance, both institutions could concentrate on refining their curriculum, particularly in the areas of technology skills and personalized learning.

This research contributes to the field of Education 4.0 evaluation. It introduces a methodological framework that can be used to evaluate and compare universities' performances concerning Education 4.0 using a combination of AHP and DEA methods. Even though the result demonstrates positive progress towards Education 4.0, rapid progress in advanced technology urges us to focus on continuous improvement and fast implementation.

References

- Agrawal, S., Sharma, N., & Bhatnagar, S.: Education 4.0 to industry 4.0 vision: current trends and overview. In: Agrawal, R., Jain, J.K., Yadav, V.S., Manupati, V.K., Varela, L.: *Recent Advances in Smart Manufacturing and Materials* Singapore (2021)
- Akimov, N., et al.: Components of education 4.0 in open innovation competence frameworks: systematic review. *J. Open Innovation: Technol. Market Complex.* **9**(2), 100037 (2023). <https://doi.org/10.1016/j.joitmc.2023.100037>
- Alam, T.E., González, A.D., Raman, S.: Benchmarking of academic departments using data envelopment analysis (DEA). *J. Appli. Res. Higher Educ.* (2019)
- Aparicio, J., Monge, J.F.: The generalized range adjusted measure in data envelopment analysis: Properties, computational aspects and duality. *Euro. J. Operat. Res.* **302**(2), 621–632 (2022). <https://doi.org/10.1016/j.ejor.2022.01.001>
- Barra, C., Zotti, R.: Measuring efficiency in higher education: an empirical study using a bootstrapped data envelopment analysis. *Int. Adv. Econ. Res.* **22**(1), 11–33 (2016). <https://doi.org/10.1007/s11294-015-9558-4>
- Berbegal-Mirabent, J.: The influence of regulatory frameworks on research and knowledge transfer outputs: an efficiency analysis of Spanish public universities. *J. Eng. Technol. Manag.* **47**, 68–80 (2018). <https://doi.org/10.1016/j.jengtecman.2018.01.003>
- Chaowarat, W., Suto, H., Piboonrungrroj, P.: Extended Supply chain DEA for considering replaceable DMUs. *Adv. Mater. Res.* **974**, 305–309 (2014). <https://doi.org/10.4028/www.scientific.net/AMR.974.305>
- Ellingrud, K., Gupta, R., Salguero, J.: *Building the vital skills for the future of work in operations* (2020)
- Endovitsky, D.A., Komendenko, S.N.: Application of DEA to measure the productivity of the academic staff. *Vestnik Voronezhskogo gosudarstvennogo universiteta. Ser.: Ekonomika i upravlenie = Proceedings of Voronezh State University. Series: Economics and Management* (2022)
- Firsova, A., Chernyshova, G.Y.: *Mathematical Models for Evaluation of the Higher Education System Functions with DEA Approach. Izvestiya of Saratov University. New Series. Series: Mathematics. Mechanics. Informatics* (2019)
- González-Pérez, L.I., Ramírez-Montoya, M.S.: Components of education 4.0 in 21st century skills frameworks: systematic review. *Sustainability* **14**(3), 1493 (2022). <https://www.mdpi.com/2071-1050/14/3/1493>
- Haderer, B., Ciolacu, M.: Education 4.0: Artificial intelligence assisted task- and time planning system. *Proc. Comput. Sci.* **200**, 1328–1337 (2022). <https://doi.org/10.1016/j.procs.2022.01.334>
- Haelermans, C., Ruggiero, J.: Estimating technical and allocative efficiency in the public sector: a nonparametric analysis of Dutch schools. *Eur. J. Oper. Res.* **227**, 174–181 (2013). <https://doi.org/10.1016/j.ejor.2012.12.004>
- Hew, K.F., Lo, C.K.: Flipped classroom improves student learning in health professions education: a meta-analysis. *BMC Med. Educ.* **18**(1), 38 (2018). <https://doi.org/10.1186/s12909-018-1144-z>

- Ho, W.: Integrated analytic hierarchy process and its applications – a literature review. *Eur. J. Oper. Res.* **186**, 211–228 (2008). <https://doi.org/10.1016/j.ejor.2007.01.004>
- Hussin, A.: Education 4.0 made simple: ideas for teaching. *Inter. J. Educ. Literacy Stud.* **6**, 92 (2018). <https://doi.org/10.7575/aiac.ijels.v6n.3p.92>
- Janmontree, J.: Sustainability Performance Measurement Framework for the Product Life Cycle An Application for the Wind Turbine Industry Otto-von-Guericke University (2021)
- Jones, K., Sharma, R.: Higher Education 4.0: The Digital Transformation of Classroom Lectures to Blended Learning (2021). <https://doi.org/10.1007/978-981-33-6683-1>
- Mansor, N., Abdullah, N., Abd Rahman, H.: Towards electronic learning features in education 4.0 environment: Literature study. *Indonesian J. Elect. Eng. Comput. Sci.* **19**, 442 (2020). <https://doi.org/10.11591/ijeecs.v19.i1.pp442-450>
- Mayston, D.J.: Effectiveness analysis of quality achievements for university departments of economics. *Appl. Econ.* **46**, 3788–3797 (2014)
- Miranda, J., et al.: The core components of education 4.0 in higher education: three case studies in engineering education. *Comput. Elect. Eng.* **93**, 107278 (2021). <https://doi.org/10.1016/j.compeleceng.2021.107278>
- Mota, T.R.A., Meza, L.A.: The use of DEA as a tool to evaluate public expenditure on education: an analysis of the cities of the state of Rio de Janeiro. *Anais da Academia Brasileira de Ciencias* **92**(2), e20190187 (2020)
- Nouri, J.: The flipped classroom: for active, effective and increased learning – especially for low achievers. *Int. J. Educ. Technol. High. Educ.* **13**(1), 1 (2016). <https://doi.org/10.1186/s41239-016-0032-z>
- Oliveira, K., Souza, R.: Digital transformation towards education 4.0. *Inform. Educ.* **21** (2021). <https://doi.org/10.15388/infedu.2022.13>
- Patitad, P., Watanabe, W.C.: DEA Model for considering relationship between supply chain members. *Naresuan Univ. J. Sci. Technol. (NUJST)* **30**(4) (2022). <https://doi.org/10.14456/nujst.2022.34>
- Raco, J., et al.: Soft Skill of higher education industry in industry 4.0 era using buckley’s fuzzy AHP. *Inter. J. Analytic Hierarchy Proc.* **14**, 21–37 (2022). <https://doi.org/10.13033/ijahp.v14.i1.943>
- Ramírez-Montoya, M.S., Castillo-Martínez, I.M., Sanabria-Z.J., Miranda, J.: Complex Thinking in the framework of education 4.0 and open innovation—a systematic literature review. *J. Open Innovat. Technol. Market Compl.* **8**(1), 4 (2022). <https://www.mdpi.com/2199-8531/8/1/4>
- Rienties, B., et al.: Education 4.0 in higher education and computer science: A systematic review. *Comput. Appli. Eng. Educ.* n/a(n/a). <https://doi.org/10.1002/cae.22643>
- Santín, D., Sicilia, G.: Using DEA for measuring teachers’ performance and the impact on students’ outcomes: evidence for Spain. *J. Prod. Anal.* **49**, 1–15 (2018)
- Shero, J.A., Al Otaiba, S., Schatschneider, C., Hart, S.A.: Data envelopment analysis (DEA) in the educational sciences. *J. Exp. Educ.* **90**(4), 1021–1040 (2022). <https://doi.org/10.1080/0020973.2021.1906198>
- Sierra-Fernández, C.R., Alejandra, H.-D., Trevethan-Cravioto, S.A., Azar-Manzur, F.J., Mauricio, L.-M., Garnica-Geronimo, L.R.: Flipped learning as an educational model in a cardiology residency program. *BMC Med. Educ.* **23**(1), 510 (2023). <https://doi.org/10.1186/s12909-023-04439-2>
- Ulufier, N.B., et al.: Digitalization maturity model development for higher education. In: Kahraman, C., Haktanir, E. (eds.) *Intelligent Systems in Digital Transformation: Theory and Applications*, pp. 471–488. Springer International Publishing (2023). https://doi.org/10.1007/978-3-031-16598-6_20

- WorldEconomicForum. Schools of the Future: Defining New Models of Education for the Fourth Industrial Revolution (2020)
- WorldEconomicForum. Defining Education 4.0: A Taxonomy for the Future of Learning (2023)
- Zoghbi, A., Rocha, F., Mattos, E.: Education production efficiency: Evidence from Brazilian universities. *Econ. Model.* **31**, 94–103 (2013). <https://doi.org/10.1016/j.econmod.2012.11.018>



Digital Technology for the Ageing Population: Perspectives from the Younger Generation

Miloslava Černá¹, Anna Borkovcová¹, and Simon K. S. Cheung²(✉)

¹ Faculty of Informatics and Management, University of Hradec Kralove, Rokitanskeho 62,
500 03 Hradec Kralove, Czech Republic

{Miloslava.Cerna,anna.borkovcova}@uhk.cz

² Information Technology Office, Hong Kong Metropolitan University, Homantin, Hong Kong,
Kowloon, China

kscheung@hkmu.edu.hk

Abstract. The paper investigates digital technology for the ageing population from the perspectives of the younger generation. Based on the inter-generational mutual interaction, sharing and enrichment in the use of digital technology, our study involved the students from the Faculty of Informatics and Management, University of Hradec Kralove as the representatives of the younger generation in reviewing the use of digital technology for the elderly, understanding their needs and examining their digital literacy. Qualitative research was applied in the study. The students were divided into three focus groups, taking different approaches to investigating the topic. The importance of digital technology for the well-being of the elderly was highlighted, and the utilization of ICT according to their usage was reviewed. It is revealed that digital literacy of the elderly has become a burning issue. They faced difficulties in meeting basic needs, especially during the COVID-19 pandemic. Our study also showed how the elderly and younger generations can benefit from the mutual cooperation.

Keywords: digital technology · digital literacy · ageing population · social technology · COVID-19 pandemic

1 Introduction

Unquestionably, everyone's daily life had been greatly affected by the unprecedented COVID-19 pandemic. The use of digital technology might help to some extent, such as online meetings, online shopping and e-banking. However, many elderly people are at a disadvantage as it is demanding for them to keep pace with the development of technologies. Coping with the ageing society is a part of both transnational policies [1] and national policies [2]. The paper investigates digital technology for the ageing population from the perspectives of the younger Generation Z. Generation Z refers to those who were born between 1997 and 2012 - the generation that was raised in the era of the Internet and social media [3].

Globally, the number of people aged 60 and over was 1 billion in 2019. By 2050, the number will double to 2.1 billion. This historically significant change in the global

population requires an adaptation of the structure of all sectors like medical and social care or urban planning to keep pace with the changing demography. Aging brings challenges and opportunities. Companies that adapt to this changing demography and invest in healthy aging can enable individuals to live longer and healthier lives, and companies will thrive. Increased demand for primary health care and long-term care will require a larger and better trained workforce and an intense need for greater adaptation of the physical and social environment to age [4].

The Czech population is aging demographically. The growing number and share of seniors in the second decade of the 21st century was due to the transfer of people from strong generations born after World War II to the age of 65, as well as the long-term increase in life expectancy. The standard of living can be assessed from multiple perspectives, but logically the income side dominates. In terms of the average level of income, the Czech Republic is slightly above average in relation to Eastern European countries. The highest expenditures of seniors in the Czech Republic are spent on food and housing. This limits the possibilities of other spending.

In Czech Republic, 9% of employees are currently 60 + and they are quickly approaching retirement age. However, a slightly higher standard of living can also be caused by the fact that a large part of the population in the 60 + still goes to work, so they have a stable full income. In addition, there still may be the influence of the former regime and the revolution, when these people were able to buy property under very favorable conditions. Thus, this standard of living may not indicate that a similar standard could remain for future generations.

According to the Czech Statistical Office, the number of people aged 65 and over is growing in the population of the Czech Republic. Between 2010 and 2019, the number of elderly (aged 65 or over) increased by almost half a million from 1.64 to 2.13 million (as of 31 December 2019), and they already make up one fifth of the Czech population [5]. The number of elderly is growing in all sub-age groups. The largest group is traditionally represented by the youngest elderly aged between 65 and 69, they formed at the end of 2019 about one third of all elderly. Almost 30% were aged between 70 and 74. The remaining two-fifths represented two relatively similar groups, one aged between 75 and 79 and the other aged 80 and over.

The topics regarding the ageing population and the uses of digital technology for the elderly has been embedded in the study programmes in the Faculty of Information and Management, University of Hradec Králové [6]. Lessons learnt from the COVID-19 pandemic suggested actions need to be taken so that the elderly could participate at least partially, if not fully, in the digital space without feeling of being isolated. It is therefore desirable to motivate the elderly in attaining digital competence, and provide necessary assistance to them.

It would be the right time to explore the topics. For this reason, a study was carried out by the Faculty of Information and Management, University of Hradec Králové in order to investigate digital technology for the ageing population in Czech Republic, understand the needs of the elderly people in using ICT, and explore ways to help them attain digital literacy. Our study target is the elderly in Czech Republic. Specifically, we focus on the Baby boomers who were born between 1946 and 1964, currently aged between 59 and 77.

Our study was conducted during the period when the COVID-19 pandemic was almost over. We aimed to derive unique results and findings on the topics from the perspectives of the younger generation, or more specifically, the Generation Z who were born between 1997 and 2012, currently aged between 11 and 26. It is believed that the results and findings are unique in the literature, as similar studies are rarely, if not, reported.

In conducting our study, we engaged the students from the Faculty of Information and Management, University of Hradec Králové, who all came from the younger Generation Z. Hence, the results and findings could accurately represent the viewpoints from the younger generation. It is also found both the younger and elder generations could benefit from this inter-generation interaction and cooperation. This paper serves to reports the results and findings.

The rest of this paper is structured as follows. Following this introductory section, Sect. 2 describes our research method and setting. Section 3 reports the results and findings. Section 4 concludes this paper with discussions.

2 Research Method and Setting

Our study emphasizes the importance of inter-generational cooperation. There are positive impacts of intergenerational cooperation, which foster the interaction between different ages, help overcome age-related stereotypes is taken from central Italy [7]. Another good source on inter-generational studies is the Journal of Intergenerational Relationships which focuses on the intergenerational field from a practical, theoretical, and social policy perspective [8]. Our idea also evolved from our previous promising experience of taking the inter-generational approach to organizing computer courses for the elderly [9, 10].

Defining generations as X, Y, Z is found in scholarly articles [11, 12]. Our study focus on two generations, namely, Baby boomers and Generation Z [13]. Baby boomers were born between 1946 and 1964. They are currently aged between 59 and 77, representing the elderly generation. Generation Z were born between 1997 and 2012. They are currently aged between 11 and 26, representing the younger generation. [14]. Throughout this paper, the elderly generation refers to the baby boomers while the younger generation refers to the Generation Z. Information and communication technology (ICT) is a term that refers to devices that can be connected to the Internet such as computers, smartphones, tablets.

Elderly people are usually considered vulnerable, and they were much more susceptible to the COVID-19 pandemic. What is worth highlighting is that, in many other aspects, they are stronger and be able to enrich life for younger generations. Patra claimed that mental health of the elderly is better than that of younger generations, as they have much more life experience, where wisdom brought them to moderate prudence and rationality [15]. There is less depression, anxiety and stress in the elderly. They have a greater ability to adapt and regulate emotions and stress. We still have a lot to learn from them. They are able to look at problems from different angles and in the longer term.

Generation Z differs from the previous generations in the ways they perceive world, communicate, get and process information, interact and learn. Generation Z has grown

up with the Internet and social media. They are technologically savvy that are able to learn and get information quickly, which seems positive however, with demand of instant information. There are troubles with longer concentration and patience. When dealing with computer literacy from an inter-generational perspective, considerable pitfalls could be expected.

Our study involved the students of Generation Z in investigating the topics and working with the elderly people to help them acquire digital literacy and computer skills. Based on the knowledge of Generation Z characteristics, we cannot expect that every one of them could be a patient teacher who can systematically explain the “secrets” of computers and the Internet. However, they were given a free hand to decide their approaches, such as using suitable resources and materials from the Internet that are suitable for the elderly people.

Qualitative research is adopted in our study. The research sample was made up from the students from the Faculty of Informatics, University of Hradec Králové. The involved students have a number of attributes. They had personal experience with the elderly in a close family circle or close surroundings, with assisting them in mastering modern technologies from phones to creating a profile on a social network. They are students of Generation Z, corresponding to full-time students with specialization in ICT. They had a deeper awareness of ICT because it was their major in their studies, so it can be assumed that they may know the potential pitfalls that might occur, for example, online shopping and payment.

The process of sampling is described as follows. First, a total of 78 sophomore students of Applied information and Information Management, undertaking Bachelor degree programs and attending Professional English classes formed the population. The age of the participants is in the range of 20–23 years old, thus belonging to Generation Z along with their full-time studentship.

Then, a sampling frame consisting of three steps was created. Step 1 is a voluntary sampling. Out of 78 students 43 students took part in the voluntary task on virtual platforms for the elderly and submitted their assignment into the learning management system Blackboard. Step 2 is a non-probability purposive sampling based on stated research sample criteria. 17 students fit the criteria. Step 3 divides these students into three focus groups. Each focus group discussed the topic for about an hour during consultation hours, the discussion was run in English. The discussion ended with the evaluation; one student with the greatest potential or the most interesting elaboration of the task was selected from each focus group.

Data were collected from all phases, including pre-discussion, submitted tasks in the e-course, presentation of assignments and a follow-up discussion. In working the tasks, the students might apply various techniques, including analogy, comparative, explorative and thematic analysis. Heuristic teaching method was applied to facilitate a constructive solution of the stated task on the elderly and specialized virtual platforms on the Internet designed for this target group of people. The characteristic of the heuristic approach fits philosophy of the ways of the task accomplishment which is based on the active involvement of participants who construct their knowledge by discovering the issue themselves [16].

3 Results and Findings

This section summarizes the results and findings from the three focus groups which took different approaches to investigating the topics and proposing solutions for the elderly people.

In the following, one representative student of each focus group would describe how he or she considered the problems, interacted with the elderly people, and made their proposals or suggestions.

The student from the first focus group take an approach to investigate the issues from the point of view of how she would speak to her grandparents. So, she speaks to them in the first person, trying to outline to them what the use of the Internet can bring to them. She realizes that these technologies can be scary or unfamiliar to them, so she tries to motivate them to overcome fear. She showed them ways how the Internet could be beneficial to them, starting with listing the basic functions and tools, such as e-mail, mobile phone (calling, SMS and as a supplement to the Internet, for example when shopping online), then illustrating the list of useful information that a senior can find on the Internet. These are mainly contact details of their medical doctors and office hours, latest news from their home cities and news from around the world, and the weather forecast.

Then, the students offered the possible uses of online games, which can help in the prevention of mental illness (Alzheimer's, dementia, etc.). She informed them that, in Czech Republic, they could have the opportunity to obtain a prescription for regularly used medicines using a mobile phone and would not always have to visit their general practitioner or specialist. She also offered them the possible uses of online shopping, both for everyday use (groceries, drugstores, etc.) and shopping for gifts. She informed that they could on-line shopping with a phone call, although they usually prefer to talk. Lastly, she highlighted to them that they could use the Internet for leisure activities or obtaining information about hobbies such as painting, gardening, crocheting through discussion forums or thematically focused sites.

The approach taken by the second focus group is different from the approach taken by the first focus group. As reported by the representative student, he grasped the task rather in a statistical way. He selected a total of 21 websites that could potentially be of interest to the ageing population and evaluates them according to their usefulness and usability. The contents of these websites were of the elderly's interests. The websites also include different kinds of study courses, such as computer literacy courses and hobby courses, and instructional videos, etc.

The students evaluated the selected websites in accordance with the following attributes, sorted by priority, namely: user-friendly interface, no registration, free content, organization, course quality, primary focus on elder, low amount advertisement, Czech language, variety on topics, and information only. He gave these attributes a weight of

between 50% and 100%. After evaluation, the websites with best ratings include: www.youtube.com, www.getsetup.io, beconnected.esafety.gov.au, www.duolingo.com, and www.udemy.com. The beconnected.esafety.gov.au had its primary focus on the elderly population. Websites also presented in the Czech language were www.youtube.com and www.duolingo.com. While for other attributes, these webpages were comparable.

The third focus group took a more personal approach. As reported by the representative student, she tried to understand and even empathize with elderly population in using the Internet and browsing websites. Based on her knowledge and understanding of the elderly needs, she defined the requirements that the website should meet. These include vision, coordination, memory, knowledge, help, advertisement, and personal data collection.

Vision means that the elderly should be able to zoom in on the page, the design should be simple, uncluttered, the pages should be uniform and consistent, contrasting colours should be used, but in limited quantities. Coordination means that the elderly should have as much area as possible for clickables, it should be very clear that they are making a choice, or that they can make a choice, and time limits should be waived so that the senior has time to process information and react. Memory is related to a possible memory problem. There should be no complicated or over-structured sentences on the website. The contents should be stylistically consistent, without ambiguity.

Ideally, there should be a help button on the page, which should be highlighted for the elderly to find support. A chatbot could be a suitable alternative. Knowledge is associated with how much knowledge the senior will take from the site. The pages should be organized, the information should be thematically divided, and it may be appropriate to use icons with informational value. Telephone support could be more suitable for seniors, as they prefer human voice. Besides, websites should limit the advertisements that might be confusing or even overwhelming to the elderly. It should be clearly stated in the websites what personal data would be collected, and whether the websites could be truthful.

The representative student recommended a number of websites for the elderly population, namely: <https://www.kiwico.com>, providing delivery service for all ages, however more focused on children; ohhdeer.com, also providing delivery services, but more focused on art; www.domestika.org, which is a community for people interested in creativity; www.coursera.org, providing with a wide range of online courses; www.zivot90.cz/cs/komunita/online-kurzy and e-senior.czu.cz, also providing online courses but especially for seniors.

4 Discussion and Conclusion

This section summarizes and concludes this paper with discussion on the results and findings from our study.

Generation Z is probably the generation with the greatest potentials to help the elderly population acquire digital literacy. The students from the Faculty of Informatics, University of Hradec Králové, as the representation of Generation Z, participated in our study. Being experts in ICT to some extent, the students investigated various digital technology for the elderly population, and tried to identify the elderly's needs and sort

out effective ways to help them acquire digital literacy. This study provided valuable experience for the students to better understand the issue of digital literacy for the elderly people.

From this study, the students learnt that too much skill-based learning might be problematic, especially for the beginners in the elderly population. They also learnt that basic knowledge and concepts as well as simple and straight-forward skills could not be considered as taken for granted for the elderly.

Through a heuristic approach to the exploration of the topic, our study promoted critical and imaginative thinking for students by engaging them in proposing effective and appropriate solutions to help elderly acquire digital literacy via self-determined problem-based learning. Pre-discussion with students served to define the problem issues for the students to investigate and solve. They need to learn something about psychology, something about medicine, something about the structure of portals and services, something about the policy applicable in elderly. They involved their own experience in search of a solution.

The issue of digital literacy in the elderly was considered from the perspectives of the students - how they perceive it, what they believe could work and what personal experiences they have. There were interactions between the students and the elderly people, which turned into direct intergenerational discussion. The students could better understand the topic, and propose solutions that reflect these unique viewpoints correspond to the findings. Both affective and cognitive determinants are of key importance in learning [17]. When intergenerational approach is applied, desired digital skills acquisition is being enriched with social acquisition. The students might benefit from reverse mentoring or service-learning [18].

Given the diversity in approaches taken by the three focus groups, we can expect that different approaches have their own advantages and shortcomings. It is therefore important to identify the right combination of approaches that could lead to the most promising results. They share one common observation – it is very important for one to fully understand why elderly people use and do not use the Internet. These helped derive appropriate solutions to in tackling the issue.

It is undeniable that digital literacy is the basis for the limited functioning of a society being lockdown. This has been experienced during the COVID-19 pandemic, where the elderly people were found to possess a lower level of digital literacy. Now, the pandemic is almost over, it is the right time for the society to review the lessons learnt. There are strong needs to find a way to help the elderly attain digital literacy. We believe this is valid not only in Czech Republic but also other countries and regions over the globe.

In this paper, we shared the motivation of our study, and presented the results and findings on this topic from the perspectives of Generation Z. We believe our unique results and findings would be of useful for researchers and practitioners in the field to make reference in conducting research studies, formulating policies and deploying solutions to promote the use of digital technology for the ageing population and help the elderly people attain digital literacy.

Acknowledgments. This study is supported by the SPEV project 2023, run at the Faculty of Informatics and Management, University of Hradec Kralove, Czech Republic.

Conflicts of Interest. The authors declare no conflict of interest.

References

1. Szewczenko, A.: The concept of smart city in terms of improving the quality and accessibility of urban space for the elderly: literature review. *Archit. Civil Eng. Environ.* **13**(2), 27–35 (2020). <https://www.sciendo.com/article/>, <https://doi.org/10.21307/acee-2020-015>
2. Tymkiewicz, J.: The role of an architect in creating the image of an elderly-friendly sustainable smart City. *Building* **9**(10), 223 (2020). <https://www.mdpi.com/2075-5309/9/10/223>
3. Chillakuri, B.: Understanding generation z expectations for effective onboarding. *J. Organizat. Change Manag.* **33**(7), 1277–1296 (2020). <https://www.emerald.com/insight/content/>, <https://doi.org/10.1108/JOCM-02-2020-0058/full/html>
4. United Nations, World Population Ageing 2020 Highlights: Living Arrangements of Older Persons. Department of Economic and Social Affairs, Population Division, United Nations (2020). <https://doi.org/10.18356/9789210051934>
5. Czech Statistical Office. Home page of Czech Statistical Office, Czech Republic. <https://www.czso.cz/csu/czso/home>
6. Univerzita Hradec Králove, EU Project - COST Action CA16226, Indoor Living Space Improvement: Smart Habitat for the Elderly 2017–2021, Univerzita Hradec Králove, (2023). <https://www.uhk.cz/cs/fakulta-informatiky-a-managementu/veda-a-vyzkum/vav-projekty/evropske-projekty>
7. Santini, S., Tombolesi, V., Baschiera, B.: Intergenerational programs involving adolescents, institutionalized elderly, and older volunteers: results from a pilot research-action in Italy. *BioMed. Res. Inter.* **2018**, 4360305 (2018). <https://doi.org/10.1155/2018/4360305>
8. Jarrott, S. (ed.) *J. Intergenerational Relation.* (2023). <https://www.tandfonline.com/toc/wjir20/current>
9. Cerna, M., Poulouva, P., Svobodova, L.: The elderly in SMART cities. In: Uskov, V.L., Howlett, R.J., Jain, L.C., Vlacic, L. (eds.) *KES SEEL-18 2018. SIST*, vol. 99, pp. 224–233. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-92363-5_21
10. Svobodova, L., Cerna, M.: Benefits and pitfalls in utilization of the internet by elderly people. In: Kar, A.K., Vigneswara Ilavarasan, P., Gupta, M.P., Dwivedi, Y.K., Mäntymäki, M., Janssen, M., Simintiras, A., Al-Sharhan, S. (eds.) *Digital Nations – Smart Cities, Innovation, and Sustainability: 16th IFIP WG 6.11 Conference on e-Business, e-Services, and e-Society, I3E 2017, Delhi, India, November 21–23, 2017, Proceedings*, pp. 135–146. Springer International Publishing, Cham (2017). https://doi.org/10.1007/978-3-319-68557-1_13
11. Bresman, H., Rao, V.D.: A survey of 19 countries shows how generations X, Y, and Z are - and aren't – different. *Harvard Bus. Rev.* **25** (2017). <https://hbr.org/2017/08/a-survey-of-19-countries-shows-how-generations-x-y-and-z-are-and-arent-different>
12. Dewanti, P., Indrajit, R.E.: The Effect of XYZ Generation Characteristics to e-Commerce C-to-C: A Review. *Ikraith Informatika* **2**(2), 56–60 (2018). <https://journals.upi-yai.ac.id/index.php/ikraith-informatika/article/view/204>
13. Dimock M., “Defining Generations: Where Millennials End and Generation Z Begins”, *Generations*, Research Topic, 17 January 2019, Pew Research Centre, 2019. Available from: <https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/>
14. Kasasa, P.: Boomers, Gen X, Gen Y, Gen Z, and Gen A Explained, Internet Resources, Kasasa (2021). <https://www.kasasa.com/exchange/articles/generations/gen-x-gen-y-gen-z>
15. Patra S.: COVID-19: An opportunity to learn resilience from the elderly. *J. Geriatric Care Res.* **8**(1), 8–10, Geriatric Care and Research Organisation (2021). https://instituteofinsight.org/wp-content/uploads/2021/07/Journal_of_Geriatric_Care_and_Research_2021_06.pdf

16. Khafizova, A., Zakirova, V.: Effective-Practical and Creative (Heuristic) Pedagogical Technologies of Students' Self-Development. Kazan Federal University, Russia, pp. 1037–1042 (2019) [cited 31 Aug 2021]. <https://ap.pensoft.net/article/22156/>
17. Cerna, M.: Psychodidactic approach in the development of language competences in university students within blended learning. *Open Learn. J. Open Distance e-Learn.* **33**(2), 142–154 (2018). <https://www.tandfonline.com/doi/full/>, <https://doi.org/10.1080/02680513.2018.1454834>
18. Tambaum, T.: Intergenerational learning in action. In: Evans, K., Lee, W.O., Markowitsch, J., Zukas, M. (eds.) *Third International Handbook of Lifelong Learning*, pp. 1–24. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-67930-9_35-1



Exploring Student Profile Features and Their Impact on Learning Performance in Secondary School

Yicong Liang¹, Haoran Xie², Di Zou³, Xinyi Huang⁴, and Fu Lee Wang¹ (✉)

¹ School of Science and Technology, Hong Kong Metropolitan University, Hong Kong, China
{s1304812, pwang}@hkmu.edu.hk

² Department of Computing and Decision Sciences, Lingnan University, Hong Kong, China
hrxie@ln.edu.hk

³ Centre for English and Additional Languages and School of Graduate Studies, Lingnan University, Hong Kong, China
dizou@ln.edu.hk

⁴ Society Hub, The Hong Kong University of Science and Technology, Guangzhou, China
caraxyhuang@hkust-gz.edu.cn

Abstract. Emerging technologies have allowed researchers to easily access educational data, conduct data analysis, and predict students' learning performance. However, the factors that are essential for the predictive model have not been identified. In the present research, based on the information entropy framework, we firstly identify the factors that influence students' academic learning performance. Then, we adopt the explainable machine learning frameworks, which are based on logistic regression and support vector machines, to predict student learning achievements. The experiment was conducted on the real-world dataset from the secondary school within two subjects. The results reveal that the feature of the failure records from students' past performance is a significant factor related to learning achievements. The predictive model based on student profiles achieves up to 86% accuracy for the prediction of learning outcome related to the final grade.

Keywords: Student Learning Performance · Data Analysis in Education · Student Profiles · Learning Achievement Predictive Model

1 Introduction

Massive Open Online Courses (MOOCs) have become a popular and important learning mode and increasingly attracted researchers' interest in recent years (Liu et al. 2021). It is because they can easily gain educational data from the learning platform to conduct learning analysis. MOOCs can support self-regulated learning as they allow learners to choose and control the learning resources, learning path, and learning schedule, providing a high level of self-regulation (Rabin et al. 2019; Zhang et al. 2022). It also promotes personalized learning for students by bringing new strategies and techniques that meet

students' preferences (Aljaraidh, 2019). With MOOCs, students can access learning materials without time and space limits, enabling students to continue their learning during the pandemic time. On the other hand, challenges like low completion rates and high dropout rates in MOOCs still exist (Feng et al. 2019; He et al. 2015; Prenkaj et al. 2020). To facilitate student retention and increase the ratio of course completion, the instructors must provide timely intervention to the at-risk students (Hlosta et al. 2017) and identify these students in early stages (Tang et al. 2015).

Emerging techniques, especially learning analytics and machine learning, can be applied to effectively locating at-risk students and predicting their learning performances. In particular, these algorithms classified learners as either "high risk" or "low risk" for dropping out from the course by analyzing their learning activities and context information, i.e., course and user information (Feng et al. 2019). It allowed educators to predict students' final exam performance by studying students' performance at the early stages and their learning activities (Wolff et al., 2014).

Further, demographic information of students, such as gender, is also crucial for predicting dropout rates of online courses. In Liang et al. (2023), it was found that gender affected course selection, with female students preferring social science courses and male students interested in STEM. Similarly, female learners had a higher probability of dropping science-based courses, whereas male students were more likely to quit non-science courses (Feng et al. 2019). Other demographic information like age (Tabassum and Akhter, 2020), socio-economic status (Farooq et al., 2011), and parental occupational level (Odoh et al. 2017) also had significant impacts on students' learning performances. Other important data sources, which have been widely employed for constructing student profiles, are from the learning logs in the online learning system (Zou and Xie, 2018) and interaction histories in the social media platforms (Xie et al. 2015).

Therefore, it was challenging to identify features that could fit into machine learning models for analysis or employed to student profiles as it required expert domain knowledge and the number of candidate feature sets is very large. To fill this gap, we adopted the explainable machine learning framework based on Logistic Regression (LR) (Tolles and Meurer, 2016) and Support Vector Machines (SVM) (Cortes and Vapnik, 1995) techniques in this paper to predict students' final achievement. We aimed to understand the various features, which can be employed for establishing students' profiles in facilitating the personalized and self-directed learning in MOOCs, and investigated how these features might interact and influence secondary school students' academic learning performance.

2 Literature Review

Machine learning techniques are effective and often applied to develop prediction models (Feng et al. 2019; Tang et al. 2015; Wang et al. 2020; Wang et al. 2021). For example, Wang et al. (2020) used convolutional neural networks to obtain user context information (e.g., demographics and online behavior) and attention mechanisms to predict learners' performances. In Tang et al. (2015), they developed a framework to identify students with higher chances to drop out with the decision tree model as the classification method. Wang et al. (2021) proposed a graph neural network (i.e., Top-N personalized

Recommendation with Graph Neural Network (TP-GNN)) to achieve personalized recommendations. Several studies have proved the feasibility of using binary classification such as dropout or non-dropout (Feng et al. 2019) and pass or fail (Kuzilek et al. 2015) to predict learning outcomes. It was observed that the initial submission of assignments (Hlosta et al. 2017) and the average time that students spent on watching instructional videos were essential in identifying at-risk students. With these techniques and models, the instructors could easily identify which students have a higher tendency to quit the courses (Prenekaj et al. 2020). It allowed the instructors to have adequate information about students' learning situations and make interventions accordingly before students leave the course without notice (Jayaprakash et al. 2014; Wolff et al. 2014).

To identify which attributes should be included in the predictive model, some studies are concerned about the factors that affect learning performance. For example, Feng et al. (2019) found that gender and friend relationships had significant impacts on learners' dropout probabilities. Liang et al. (2023) focused on the gender factor and found an inherent bias existed in online course registration, with female learners showed a preference for social science courses and male for STEM. Additionally, Aldowah et al. (2020) concluded that high dropout rates in MOOCs were related to low academic skills and unpleasant prior experience. Chen et al. (2020) found that course assessments, learning tools and platforms, and critique were the aspects that students have most negative comments towards MOOCs, which might affect students' intention of dropping out. Similarly, Chen et al. (2021) found that grading, problem solving and practices, course quality, and textbook design were the main attributes of students' satisfaction in MOOC learning. Further, Tao et al. (2023) were concerned about students' emotional behaviors while learning MOOCs. They revealed that engagement is the most significant indicator of students' performance. Compared to the previous studies (Xie et al., 2017; Xie et al., 2019), which focused more on the relevant features among online learners, this paper focused on the data drawn from students' profiles and investigated how family, school, and social relationships affect students' learning achievement.

3 Methodology

This work uses Student Performance (Cortez, 2014) as the dataset in the experiment for exploration because it includes a variety of information in student profiles that are in accordance with our research aims. This dataset contains different attributes from each student, related to three aspects: demographics, school-related, and social-related. For example, demographic features include the student's home address, parent's education, and occupation status information. For school-related features, studying time per week, desire to take higher education, and past failure records are also included. The frequency of going out with friends and alcohol consumption during workdays and weekends are examples of social-related features (Cortez and Silva, 2008).

Overall, by exploring the real-world dataset related to secondary school students' academic performance, our research is guided by the research questions as follows:

RQ1: What is the relatedness between student profile features and learning achievements in the secondary school within two different subjects?

RQ2: What is the prediction performance from the predictive model based on secondary school students' profiles?

3.1 Relatedness Measurement with Learning Performance

The information entropy can be used to measure the uncertainty of a random variables (MacKay, 2003). This study adopts information entropy method to identify relevant features connecting with learning performance. In the context of learning achievements in a course, learning outcome in terms of pass or fail can be considered as a discrete random variable¹ denoted as Y . The entropy of learning outcome $H(Y)$ can be computed as follows:

$$H(Y) = -p \log p - (1 - p) \log(1 - p) \quad (1)$$

p denotes the probability of pass. Since students may have different profiles, e.g., come from different family background, different personalities etc., these factors can have different impact on the students' learning performance. In other words, the uncertainty of learning outcome can be altered when conditioned on different student feature values. Hence, this uncertainty can be measured by the entropy of conditional probability distribution of learning outcome and it can be calculated as follows:

$$H(Y|X_j = k) = -p_{jk} \log p_{jk} - (1 - p_{jk}) \log(1 - p_{jk}) \quad (2)$$

$H(Y|X_j = k)$ is the entropy of the random variable Y (i.e., learning outcome) conditioned on another discrete random variable X_j (i.e., the j -th profile feature) taking a certain value k . And p_{jk} denotes the conditional probability of pass when $X_j = k$. For example, $H(Y|X_{Fjob} = \text{'teacher'})$ represents the entropy of learning outcome when the feature related to the student's father whose job is a teacher.

According to the information theory (Shannon, 1948), the entropy achieves the maximum when the random variable follows a uniform distribution. In the context of measuring the uncertainty of student learning performance (a binary random variable), the entropy reaches its maximum when the probability of pass is one-half. If the entropy of the conditional distribution related to the learning outcome decreases, it suggests that the uncertainty of learning outcome is also reduced. In other words, the feature that can minimize the uncertain of learning outcome is considered as relevant feature. Intuitively, the entropy conditioning on the student whose father works as a teacher is less compared to the student whose father has no job.

Furthermore, the conditional entropy (CE) can quantify the uncertainty of learning performance after the whole feature information is given (MacKay, 2003). The conditional entropy given a random variable X can be computed in the following:

$$H(Y|X) = \sum_{i=1}^n p(X = x_i) H(Y|X = x_i) \quad (3)$$

n is the number of unique values in X . Note that $H(Y|X)$ is the result of the expectation of entropy $H(Y|X = x)$ over all possible values that feature X may take.

It is known that information gain (IG) is the reduction in information entropy from a prior state to another state taking some observed information (Quinlan, 1986). In

¹ A binary random variable with 1 for 'pass' and 0 for 'fail'.

particular, the IG of a certain feature X is computed as follows:

$$g(Y, X) = H(Y) - H(Y|X) \quad (4)$$

This nature of uncertainty reduction in IG helps to measure the relatedness of a selected random variable feature X towards the concerned random variable Y . Given a set of features in student profiles, this work adopts information gain to indicate salient features that could impact student's learning outcome.

Although the relevance of a feature is dependent on information gain, it has bias on those features taking a large number of distinct values (Quinlan, 1986). Information gain ratio (IGR) proposed in (Quinlan, 1986) can alleviate this problem. In particular, the information gain ratio of a certain feature can be computed as follows:

$$g_R(Y, X) = g(Y, X)/H_X(Y) \quad (5)$$

$H_X(Y)$ is the split information entropy of feature X .

3.2 Predictive Model for Student Learning Achievement

For RQ2, a machine learning framework based on LR (Tolles & Meurer, 2016) and SVM (Cortes & Vapnik, 1995) was used to build the predictive model. Given an input n dimensional feature vector \mathbf{x} representing a student's profile with n related features, the predictive model is going to predict the binary target label y , i.e., 1 for 'pass', 0 for 'fail' in student's final exam. For LR, the inference function for prediction is computed as follows:

$$P(y = 1|\mathbf{x}) = \frac{\exp(\theta_L^T \mathbf{x})}{1 + \exp(\theta_L^T \mathbf{x})} \quad (6)$$

θ_L denotes the learned parameters of LR.

For SVM, this paper adopts the linear kernel and the inference function for prediction is computed as follows:

$$f(\mathbf{x}) = \text{sign}(\theta_S^T \mathbf{x}) \quad (7)$$

θ_S denotes the learned parameters of SVM. If the result $f(\mathbf{x})$ of is positive, then the model output is 'pass', otherwise 'fail'. The reason for using LR and SVM for the predictive model is twofold: (1) The sign of the coefficient θ_i reflects the positive or negative impact related to the i -th feature; (2) The magnitude of θ_i indicates the importance of the i -th feature on the student's learning performance.

In the SVM linear inference function, if the coefficient θ_i is positive, then a small positive change in the feature x_i will lead the value of inference function to rise. In other words, the factor corresponding to positive coefficient affects the input feature vector (e.g., student profiles) to belong to positive label (e.g., pass in the final exam). On the other hand, the negative factors influence the input feature vector to map to negative label (e.g., fail in the exam).

4 Experiment

4.1 Experimental Settings

The dataset (i.e., Student Performance) used in this study consists of student data from two Portuguese secondary schools, with students' ages ranging from 15 to 20 (Cortez and Silva, 2008). There are three grades in the dataset, namely, G1, G2 and G3, but the final year grade G3 is used for the assessment in this experiment. We converted the final year grade G3 (a numeric value ranging from 0 to 20) into a binary outcome. Particularly, the students are labeled as 'pass' when their G3 scores are larger or equal to 10, otherwise they are labeled as 'fail'. We conduct the experiment on two courses (i.e., maths and language) respectively, to explore the relatedness of student features with learning outcome.

4.2 Relatedness of Student Features in Math Course

Table 1. Conditional entropy and information gain over different features in math

feature	Condition entropy	Information gain	Information gain ratio
Medu ²	0.6255	0.0081	0.0059
Fedu	0.6264	0.0072	0.0052
Mjob	0.6266	0.0070	0.0046
Fjob	0.6308	0.0028	0.0024
higher	0.6221	0.0118	0.0571
failures	0.5749	0.0587	0.0895
studytime	0.6290	0.0046	0.0044
traveltime	0.6322	0.0014	0.0016
paid	0.6292	0.0044	0.0063
activity	0.6335	0.0001	0.0001
goout	0.6140	0.0196	0.0178
romantic	0.6287	0.0049	0.0076
Dalc	0.6293	0.0043	0.0057
Walc	0.6317	0.0019	0.0018

To answer RQ1, we selected 14 features from student profiles for analysis in this experiment and the investigation was conducted on two courses respectively. Specifically,

² Medu denotes the feature of occupation status the student's mother. For the meaning of denotations for other features can refer to Cortez, P. (2014). *Student Performance* UCI Machine Learning Repository.

for each feature, the conditional entropy, information gain and information gain ratio are computed according to Eq. 3, Eq. 4 and Eq. 5.

For the exploration in math course, Table 1 shows the relatedness between various student profile features and their learning performance from the entropy information perspective. It can be found that failure records in the past, personal desire to take higher education, and the frequency of going out with friends are top-3 features that are highly relevant to the learning achievement in the math course. According to the feature of failures, the pass rate of students with more than two failure records in the past math exams was much lower (20%) compared to the students who did not fail the test before (75%). It suggests that the number of failures is a significant and negative impact on student's learning outcome since the more failure records the student has, the higher chance he will probably fail in the next time.

According to the feature of taking higher education, those students who did not want to pursue a college degree achieved up to 65% fail rate in the math exam. Conversely, the students who were motivated to go to college performed significantly better with about 70% pass rate. (Aldowah et al. 2020) also have the similar findings that motivation is a fundamental factor influencing students' learning performance in MOOCs. According to the feature of going out, students who excessively socialize with friends in terms of more frequently going out with friends tend to perform much worse compared to those who had a more balanced approach in socializing, e.g., resulting in dropping about 12% pass rate in the math course.

4.3 Relatedness of Student Features in Language Course

For the exploration in language course, the relatedness in the context of entropy information between student features and learning achievement is presented in Table 2. Similarly, it can be found that the number of failure records in the past, the motivation to have higher education and study time for learning language are top-3 relevant features for the performance in language course. For the feature of previous failures, the pass rate of students without any failure records was up to 90% and the fail rate was less than 10% in language exam. From the results in Table 1 and Table 2, the past performance in terms of failing records is considered as the most important indicator for learning achievement in both math and language courses.

For the feature of willingness to get higher education, the pass rate of students with a desire to attend college is up to 90%. However, the students without motivation of obtaining college degree still had 52% pass rate in language course compared to only 35% pass rate in math course. This implies that a student's proficiency in math has a greater impact on the decision to pursue higher education than the proficiency in language skills. According to the feature of weekly study time, the pass rate from the students spending more than five hours per week in learning language achieved up to 93%. In contrast, the fail rate of students who spent less than two hours each week was five-fold increase compared to those hardworking students. It is inline with the result in online courses that increasing total interaction days with virtual learning platforms helps students pass their exams (Liang et al. 2023).

Besides the top-3 features mentioned above, the features from the family background (e.g., parent's education and occupation) are also included for measuring their relatedness

Table 2. Conditional entropy and information gain over different features in language

feature	Condition entropy	Information gain	Information gain ratio
Medu	0.4161	0.0136	0.0099
Fedu	0.4141	0.0156	0.0115
Mjob	0.4213	0.0084	0.0058
Fjob	0.4274	0.0023	0.0021
higher	0.3934	0.0363	0.1073
failures	0.3676	0.0621	0.1186
studytime	0.4145	0.0152	0.0146
traveltime	0.4278	0.0019	0.0021
paid	0.4284	0.0013	0.0059
activity	0.4286	0.0011	0.0016
goout	0.4250	0.0047	0.0043
romantic	0.4265	0.0032	0.0049
Dalc	0.4221	0.0076	0.0101
Walc	0.4246	0.0051	0.0049

with learning achievements in both math and language courses. In particular, we can find that the information gain and information gain ratio from the parents' education level are consistently higher than that from the parents' occupation status as shown in *MERGEFORMAT Table 1 and Table 2. It implies that the education background of parents is a more important factor influencing their children's learning achievement in secondary school compared to the parents' professional status. Additionally, the metrics (IG and IGR) of the feature that whether the student attends extracurricular activities is the lowest among all features as shown in Table 1 and Table 2. This implies that attending extracurricular activities has little impact on student's academic learning performance in the high school.

4.4 Performance of Predictive Model for Learning Achievement

In this experiment, we used the Sklearn (Pedregosa et al. 2011) Python Library to build the predictive model for estimating student learning performance. First of all, the dataset was randomly split into training and test sets, with 30% of the data used for testing in this experiment. The effectiveness of the model was assessed according to its output and the holdout ground-truth in the test set, and the evaluation metrics in this experiment contain prediction accuracy and F1 score.

The accuracy and F1 score for the two predictive models are presented in Table 3. The SVM predictive model achieved an accuracy of 86% and an F1 score of 0.80, whereas the LR model obtained 77% accuracy and 0.68 F1 score. The SVM predictive model outperformed LR by 11.7% in accuracy and 17.6% in F1 when predicting the

Table 3. Predictive model performance

Model	Accuracy	F1
LR	0.77	0.68
SVM	0.86	0.80

academic achievement of secondary school students. The reason why SVM performs better than LR in student learning outcome prediction is that SVM performs less sensitive to outliers compared to LR. Besides, the nature of maximization of the margin between classes helps the SVM predictive model in reducing overfitting and generalizing well to unseen data.

As the coefficients of the predictive model can reflect their impacts on classification discussed in Sect. 3, we also explored which factors could influence student learning performance when the SVM classifier was doing inference. The results of factors and their impacts are presented in Table 4. The results are consistent with our previous findings in RQ1. In particular, the father's educational and occupational background were the main attributes that affected the student's success in passing the course. For individual factors, the desire to enter college, study time, and participation in extra paid classes after school, are significant predictors of success according to the classifier.

Table 4. Factors in predictive model on student learning performance

Factors helping students succeed	Factors leading students to fail
father's education	past course failures
wants to take higher education	with a romantic relationship
study time	workday alcohol consumption
attend extra paid classes within subject	going out with friends
father's job	weekend alcohol consumption

It is worth noting that there is a strong correlation between the parents' professional background (e.g., education and job) and the student's personal motivation (e.g., desire for higher education and attending extra paid classes). It might be due to the reason that parents with promising careers are often able to provide financial support for their children which enabled students to take extra classes during secondary school study. They might also encourage students to pursue higher education thereby enhancing learning motivation.

On the other hand, the school-related feature, namely failure records, is a significant factor in predicting failure. Having a large number of failures indicates a lack of sufficient exam preparation. Additionally, spending excessive time on study-irrelevant matters, such as starting dating at an early age, frequently going out with friends, and drinking excessively during workdays, increases the likelihood of failure.

5 Conclusions

In this study, we mainly conduct the exploration from the public dataset to understand how student profile features are related to students' learning achievements in the secondary school within two different courses. The relatedness measurement based on the information entropy method is used to identify relevant features. In addition, to further assist tutors in quickly identifying potential at-risk students, based on the student profile features, we build a classifier trained by an explainable machine learning framework to predict the academic learning outcome. We found that the SVM-based model achieves up to 86% accuracy as well as 0.80 F1 score in predicting students' performances in their final grades. Finally, this paper examines the factors with important impact on secondary school students' academic performance. The results reveal that parents' education level and professional job, as well as student's personal motivation, are significant predictors of success. Conversely, having a history of failure records and engaging in study-irrelevant matters, increase the likelihood of failure.

Acknowledgement. This research has been supported by the IICA Project entitled "Developing language teachers' technological pedagogical content knowledge and enhancing students' language learning in virtual learning environments" (102707), the Direct Grant (DR23B2), and the Faculty Research Grants (DB23A3 and DB23B2) of Lingnan University, Hong Kong.

References

- Aldowah, H., Al-Samarrate, H., Alzahrani, A.I., Alalwan, N.: Factors affecting student dropout in MOOCs: a cause and effect decision-making model. *J. Comput. High. Educ.* **32**, 429–454 (2020)
- Aljaraideh, Y.: Massive Open Online Learning (MOOC) benefits and challenges: A case study in Jordanian context. *Int. J. Instr.* **12**(4), 65–78 (2019)
- Chen, X., Cheng, G., Xie, H., Chen, G., Zou, D.: Understanding MOOC reviews: text mining using structural topic model. *Hum.-Centric Intell. Syst.* **1**(3–4), 55–56 (2021)
- Chen, X., Zou, D., Xie, H., Cheng, G.: What are MOOCs learners' concerns? Text analysis of reviews for computer science courses. In: Database Systems for Advanced Applications. DASFAA 2020 International Workshops: BDMS, SeCoP, BDQM, GDMA, and AIDE, Jeju, South Korea, 24–27 September 2020, Proceedings 25 (2020)
- Cortes, C., Vapnik, V.: Support-vector networks. *Machine learning* **20**, 273–297 (1995)
- Cortez, P.: Student Performance UCI Machine Learning Repository (2014)
- Cortez, P., Silva, A.M.G.: Using data mining to predict secondary school student performance (2008)
- Farooq, M.S., Chaudhry, A.H., Shafiq, M., Berhanu, G.: Factors affecting students' quality of academic performance: a case of secondary school level. *J. Quality Technol. Manag.* **7**(2), 1–14 (2011)
- Feng, W., Tang, J., Liu, T.X.: Understanding dropouts in MOOCs. In: Proceedings of the AAAI Conference on Artificial Intelligence (2019)
- He, J., Bailey, J., Rubinstein, B., Zhang, R.: Identifying at-risk students in massive open online courses. In: Proceedings of the AAAI Conference on Artificial Intelligence (2015)
- Hlosta, M., Zdrahal, Z., Zendulka, J.: Ouroboros: early identification of at-risk students without models based on legacy data. In: Proceedings of the Seventh International Learning Analytics & Knowledge Conference (2017)

- Jayaprakash, S.M., Moody, E.W., Lauría, E.J., Regan, J.R., Baron, J.D.: Early alert of academically at-risk students: an open source analytics initiative. *J. Learn. Analyt.* **1**(1), 6–47 (2014)
- Kuzilek, J., Hlosta, M., Herrmannova, D., Zdrahal, Z., Vaclavek, J., Wolff, A.: OU Analyse: analysing at-risk students at the Open University. *Learn. Analyt. Rev.* (2015)
- Liang, Y., Zou, D., Wang, F.L., Xie, H., Cheung, S.K.: Investigating demographics and behavioral engagement associated with online learning performance. In: *International Conference on Blended Learning* (2023)
- Liu, C., Zou, D., Chen, X., Xie, H., Chan, W.H.: A bibliometric review on latent topics and trends of the empirical MOOC literature (2008–2019). *Asia Pac. Educ. Rev.* **22**(3), 515–534 (2021)
- MacKay, D.J.: *Information theory, inference and learning algorithms*. Cambridge University Press (2003)
- Odoh, L.C., Ugwuanyi, U.B., Odigbo, B.E., Chukwuani, N.V.: Influence of parental occupation and level of education on academic performance of accounting students in Nigeria. *Res. Humanit. Soc. Sci.* **7**(10), 21–27 (2017)
- Pedregosa, F., et al.: Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011)
- Prencak, B., Velardi, P., Stilo, G., Distanto, D., Faralli, S.: A survey of machine learning approaches for student dropout prediction in online courses. *ACM Comput. Surv. (CSUR)* **53**(3), 1–34 (2020)
- Quinlan, J.R.: Induction of decision trees. *Mach. Learn.* **1**, 81–106 (1986)
- Rabin, E., Kalman, Y.M., Kalz, M.: An empirical investigation of the antecedents of learner-centered outcome measures in MOOCs. *Int. J. Educ. Technol. High. Educ.* **16**(1), 1–20 (2019)
- Shannon, C.E.: A mathematical theory of communication. *Bell Syst. Tech. J.* **27**(3), 379–423 (1948)
- Tabassum, R., Akhter, N.: Effect of demographic factors on academic performance of university students. *J. Res. Reflect. Educ. (JRRE)* **14**(1) (2020)
- Tang, J.K., Xie, H., Wong, T.-L.: A big data framework for early identification of dropout students in MOOC. In: *Technology in Education. Technology-Mediated Proactive Learning: Second International Conference, ICTE 2015, Hong Kong, China, 2–4 July 2015, Revised Selected Papers 2* (2015)
- Tao, X., et al.: Towards an understanding of the engagement and emotional behaviour of MOOC students using sentiment and semantic features. *Comput. Educ. Artif. Intell.* **4**, 100116 (2023)
- Tolles, J., Meurer, W.J.: Logistic regression: relating patient characteristics to outcomes. *JAMA* **316**(5), 533–534 (2016)
- Wang, J., Xie, H., Au, O.T.S., Zou, D., Wang, F.L.: Attention-based CNN for personalized course recommendations for MOOC learners. In: *2020 International Symposium on Educational Technology (ISET)*(2020).
- Wang, J., Xie, H., Wang, F.L., Lee, L.-K., Au, O.T.S.: Top-N personalized recommendation with graph neural networks in MOOCs. *Comput. Educ. Artif. Intell.* **2**, 100010 (2021)
- Wolff, A., Zdrahal, Z., Herrmannova, D., Kuzilek, J., Hlosta, M.: Developing predictive models for early detection of at-risk students on distance learning modules (2014)
- Xie, H., Zou, D., Lau, R.Y., Wang, F.L., Wong, T.-L.: Generating incidental word-learning tasks via topic-based and load-based profiles. *IEEE Multimedia* **23**(1), 60–70 (2015)
- Xie, H., Zou, D., Wang, F.L., Wong, T.-L., Rao, Y., Wang, S.H.: Discover learning path for group users: A profile-based approach. *Neurocomputing* **254**, 59–70 (2017)
- Xie, H., Zou, D., Zhang, R., Wang, M., Kwan, R.: Personalized word learning for university students: a profile-based method for e-learning systems. *J. Comput. High. Educ.* **31**, 273–289 (2019)

- Zhang, R., Zou, D., Cheng, G., Xie, H., Wang, F.L.: Supportiveness of language MOOCs for self-regulated learning: a review of commercial language MOOCs on the market. *Inter. J. Mobile Learn. Organis.* **16**(3), 323–348 (2022)
- Zou, D., Xie, H.: Personalized word-learning based on technique feature analysis and learning analytics. *J. Educ. Technol. Soc.* **21**(2), 233–244 (2018)

Author Index

A

Aristovnik, Aleksander 254

B

Borkovcová, Anna 340

Brezovar, Nejc 254

C

Cao, Caiqin 218

Černá, Miloslava 340

Chan, Hon Tung 119, 289

Chan, Pak Yuen Patrick 299

Chan, Yan-Wai 199

Chen, Chih-Ping 173

Chen, Wenyi 94

Cheng, Gary 129

Cheung, Simon K. S. 199, 242, 340

Chui, Kwok Tai 242

D

Dong, Qian 82

F

Faikhamta, Chatree 13

G

Guo, Qing 35

H

Hoo, Yee Hui 58

Huang, Xiaoran 109

Huang, Xinyi 349

J

Janmontree, Jettarat 325

Jia, Kexin 218

Jomtarak, Rangsana 13

K

Kan, Wen Huey 58

Keržič, Damijana 254

Keung, Jacky 299

Kuntasup, Marisa 264

Kwan, Christopher Chung Lim 275

Kwok, Lam-For 20

L

Lai, YuXuan 109

Lam, Sum 20

Lee, Lap-Kei 231, 242

Lertdechapat, Kornkanok 13

Li, Caoyin 68

Li, Kam Cheong 119, 208, 289

Li, Shao-Fu 231

Liang, Yicong 349

Liu, Mengjin 208

Lu, Yi 157

Luk, Louise 199

M

Miao, Rong 82

N

Naidu, Som 3

Natsupakpong, Suriya 47

Ng, Kwan-Keung 199, 231

P

Paoprasert, Naraphorn 315

Peng, Zhiyang 157

Pisuchpen, Supisara 186

Pongcharoen, Pupong 264

Prasoplarb, Tharueseana 13

Q

Qin, Wei 35

Qiu, Yi 68, 157

R

Ravšelj, Dejan 254
Ruksorn, Pongthorn 315

S

Shi, Yinghui 218
Singhtaun, Chansiri 186
Suksawang, Suwitchaphon 47
Sun, Yilin 139
Suwannachote, Laddawan 186

T

Thepphakorn, Thatchai 264

U

Umek, Lan 254

W

Wang, Fu Lee 242, 349
Wang, Hung-Hsiang 173
Wang, Yidan 109
Watanabe, Woramol C. 325
Wei, Yitong 218
Weng, Jie 139

Witchakul, Suwitchaporn 315

Wong, Billy T. M. 119, 208, 289
Wong, Leung Pun 242

X

Xia, Wan 68
Xie, Haoran 349
Xie, Youru 68
Xu, Shun 35
Xu, Xiaoshu 139

Y

Yang, Harrison Hao 35, 218
Yang, Zhen 299
Yap, Li Lian 58
Yi, Chen 20
Yu, Qijing 109

Z

Zadek, Hartmut 325
Zhang, Yunfeng 139
Zhu, Sha 35
Zou, Di 129, 349