香港教育大學
The Education University of Hong Kong

Artificial Intelligence and Digital Competency Education Centre
人工智能及數碼能力教育中心

Themed-Based International Conference Series (TBICS)

Co-organiser:

MetaACES 2025

The 3rd International Conference on Metaverse and AI Companions in Education and Society

Conference Proceedings

18 - 20 June 2025

Supporting Organisations:

Education Bureau
The Government of the Hong Kong Special Administrative Region of the People's Republic of China

AiTLE

香港電腦教育學會
The Hong Kong Association for Computer Education

Proceedings of the 3rd International Conference on

Metaverse and Artificial Companions in Education and Society

18-20 June 2025

Hong Kong

Organized by

The Education University of Hong Kong

Co-organized by

Southern University of Science and Technology

Publication of The Education University of Hong Kong

10 Lo Ping Road, Tai Po, New Territories, Hong Kong SAR

ISBN 978-988-70582-0-5
ISSN 3104-9710

MetaACES 2025

Editors

Prof. Siu Cheung KONG
The Education University of Hong Kong, Hong Kong SAR

Prof. Yu-Ju LAN
National Taiwan Normal University, Taiwan

Prof. Jianhua ZHAO
Southern University of Science and Technology, China

Prof. Yanjie SONG
The Education University of Hong Kong, Hong Kong SAR

Prof. Tanja MITROVIC
University of Canterbury, New Zealand

Preface

International Conference on Metaverse and Artificial Companions in Education and Society (MetaACES) is one of the Asia-Pacific Society for Computers in Education (APSCE) Theme-based International Conference Series. MetaACES 2025 is the third international conference, organized by The Education University of Hong Kong (EdUHK) and co-organized by Southern University of Science and Technology (SUSTech).

Metaverse and AI companions in education and society cover a wide variety of research topics. They include companionship of human and virtual entities involving applications in hardware and software forms using AI techniques. AI techniques can be in the form of multimodal processing and analytics in languages, visions, speeches, and other modalities of interaction. We call for submissions from researchers doing research in these hot topic areas. We are particularly interested in research that uses deep learning methods to support interactions between humans and AI companions in a variety of real-world educational and societal applications. These topics also include assistive robotics in addition to formal and informal learning, training, workforce, and other forms of cognitive and emotional support.

MetaACES 2025 is held on 18-20 June 2025. Days 1 and 2 of the conference are held at EdUHK's Tai Po Campus, while Day 3 is held at SUSTech's Campus in Shenzhen. The conference this year includes keynote speeches, a teacher forum and paper presentations.

Conference Organization:

Conference Chair

Siu Cheung KONG	The Education University of Hong Kong, Hong Kong SAR

Program Chair

Yu-Ju LAN	National Taiwan Normal University, Taiwan

Program Co-Chair

Jianhua ZHAO	Southern University of Science and Technology, China

Track Chairs

Yanjie SONG	The Education University of Hong Kong, Hong Kong SAR
Tanja MITROVIC	University of Canterbury, New Zealand

Senior Program Committee Members *(Surnames in alphabetical order)*

Ching Sing CHAI	Chinese University of Hong Kong, Hong Kong SAR
Maiga CHANG	Athabasca University, Canada
Gwo-Dong CHEN	National Central University, Taiwan
Cristina CONATI	University of British Columbia, Canada
Claude FRASSON	University of Montreal, Canada
Ronghuai HUANG	Beijing Normal University, China
Gwo-Jen HWANG	National Taiwan University of Science and Technology, Taiwan
Lewis JOHNSON	Alelo, United States
KINSHUK	University of North Texas, United States
Noboru MATSUDA	North Carolina State University, United States
Shengquan YU	Beijing Normal University, China

Program Committee Members *(Surnames in alphabetical order)*

Kaushal kumar BHAGAT	Indian Institute of Technology Kharagpur, India
Linda BRADLEY	The University of Gothenburg, Sweden

Chia-Kai CHANG	National Central University, Taiwan
Ching-Yi CHANG	Taipei Medical University, Taiwan
Rong-Chi CHANG	Taiwan Police College, Taiwan
Chi Yuan CHEN	National Taiwan Ocean University, Taiwan
Chiu-Jung CHEN	National Chiayi University, Taiwan
I-Chun CHEN	Chinese Culture University, Taiwan
Yi-Jing CHEN	National Chin-Yi University of Technology, Taiwan
Pei-Yu CHENG	Tamkang University, Taiwan
Fu-Ling CHUNG	Tamkang University, Taiwan
Xun GE	University of North Texas, United States
Chung-Hao HUANG	Chung Yuan Christian University, Taiwan
Tai-Yi HUANG	University of North Texas, United States
Hui-Chun HUNG	National Central University, Taiwan
Regina KAPLAN-RAKOWSKI	University of North Texas, United States
Chiu-Lin LAI	National Taipei University of Education, Taiwan
Victor LAW	University of New Mexico, United States
Sangmin LEE	Kyung Hee University, Korea
Hao-Chiang Koong LIN	National University of Tainan, Taiwan
Vivien LIN	National Changhua University of Education, Taiwan
Pei-Lin LIU	National Chiayi University, Taiwan
Boning LYU	Xiamen University, China
Zhongling PI	Shaanxi Normal University, China
Grace QI	Massey University, New Zealand
Katrin SAKS	University of Tartu, Estonia
Rustam SHADIEV	Zhejiang University, China
Kai-Yu TANG	National Chung Hsing University, Taiwan
Chia-En TENG	Tunghai University, Taiwan
Judy C. R. TSENG	Chung Hua University, Taiwan
Sheng-Shiang TSENG	National Taipei University of Business, Taiwan
Yun-Fang TU	National Taiwan University of Science and Technology, Taiwan

Airong WANG	Xi'an Jiaotong-Liverpool University, China
Li-Jen WANG	National Central University, Taiwan
Yaling WANG	National Taiwan Normal University, Taiwan
Wei WEI	Macao Polytechnic University, Macao
Jiun-Yu WU	Southern Methodist University, United States
Junjie Gavin WU	Macao Polytechnic University, Macao
Sheng-Yi WU	National Tsing Hua University, Taiwan
Wen-Chi Vivian WU	Asia University, Taiwan
Fu Yu YANG	Chang Gung University of Science and Technology, Taiwan
Hui Chin YEH	National Yunlin University of Science and Technology, Taiwan
Xuesong ZHAI	Zhejiang University, China
Shu ZHAO	Shaanxi Normal University, China

Local Organizing Committee Secretariat

Chair

Prof. Siu Cheung KONG	The Education University of Hong Kong, Hong Kong SAR

Member

Ho Fung Peter LAM	The Education University of Hong Kong, Hong Kong SAR
Mei Ki Vera CHAN	The Education University of Hong Kong, Hong Kong SAR
Yiling Elaine CHEN	The Education University of Hong Kong, Hong Kong SAR
Wan Yee TING	The Education University of Hong Kong, Hong Kong SAR
Luwei Vicky YE	The Education University of Hong Kong, Hong Kong SAR

Conference Theme:

MetaACES 2025 focuses on the themes related to education and society. The main themes of MetaACES 2025 include but are not limited to the following (in alphabetical order):

Track 1: AI and Artificial Companions in Education

- Artificial Companion in Education
- Artificial Intelligence (AI)
- Automated Feedback
- Behaviour and/or Interaction Modeling, Detection and Visualization
- Big Data Analyzed and Processed by Computers
- Chatbot
- Computational Models of Knowledge and Expertise
- Computer Supported Discussion Analysis and Assessment
- Educational Robots and Toys
- Enhancing Grading, Scoring and Feedback
- Game Analytics
- Intelligent Agents
- Intelligent Tutors and Mentors
- Learning Companion Robots (Robotic Learning Companions)
- Learning Analytics in Educational Games
- Learning Companions
- Natural Language Processing supported Tools, Systems, Applications, Mobile Apps, and Chatbots
- Roles of Artificial Companions in Metaverse
- Role Playing Games for Learning
- Security and Privacy Issues
- Sentiment Analysis
- Simulation and Training (Skill, Competence, Vocational Learning)
- Speech Recognition and Synthesis

- Stealth Assessment
- Unstructured and Semi-structured Data for Computer to Read and Learn

Track 2: Metaverse in Education

- Assessment in Games and Virtual Worlds
- Authentic Environments and Worlds
- Avatars or Player Characters for Learning
- Educational Applications of Metaverses
- Internet of Things (IoT), Internet of Everything (IoE), and/or Sensors
- Metaverse in Education
- Non-Player Characters for Learning
- Virtual and Augmented Learning Environments
- Virtual Characters in Learning and Life
- Virtual Companions in Learning and Life
- VR, AR and Simulation Technology

Track 3: Social Issues

- Artificial Companion in Society
- Bridging Informal and Formal Learning Outcome
- Emotion (Affective State) Modeling, Recognition and Detection
- Emotive Agents
- Human Computer Interaction (HCI)
- Human Robot Interaction (HRI)
- Languages, Thinking Skills, Meta-cognitive Skills, Cognitive Skills, and STE(A)M
- Metaverse in Society
- Motivational and Affective Factors on Learning with Technology
- Personal Learning Environments (PLE)
- Social Network Analysis (SNA)
- User Experience (UX) Evaluation
- Virtual Animal Learning Companions

Teacher Forum

The conference received a total of 26 submissions (11 full papers, 4 short papers and 11 poster papers) by 62 authors from 6 countries/regions (see Table 1).

Table 1: Distribution of Accepted Papers for MetaACES 2025

Country / Region	No. of Authors	Country / Region	No. of Authors
Hong Kong SAR	24	**Croatia**	5
China	16	**New Zealand**	1
Taiwan	15	**Poland**	1
		Total	**62**

The International Programme Committee (IPC) is formed by 55 Members and 5 Co-chairs worldwide. Each paper with author identification anonymous was reviewed by at least three IPC Members. Related sub-theme Chairs then conducted meta-reviews and made recommendation on the acceptance of papers based on IPC Members' reviews. With the comprehensive review process, 23 accepted papers are presented (9 full papers, 4 short papers and 10 poster papers) (see Table 2) at the conference.

Table 2: Paper Presented at MetaACES 2025

Track	Full Paper	Short Paper	Poster Paper	Total
- Track 1: AI and Artificial Companions in Education	3	3	2	8
- Track 2: Metaverse in Education	5	0	1	6
- Track 3: Social Issues	1	1	0	2
- Teacher Forum	0	0	7	7
Total	**9**	**4**	**10**	**23**

The conference comprises keynote speeches by internationally renowned scholars; a teacher forum, as well as academic and poster paper presentations.

Academic and Poster Paper Presentations

There are 7 sessions of academic and poster paper presentations with 16 papers (9 full papers, 4 short papers and 3 poster papers) in the conference. Worldwide scholars present and exchange the latest research ideas and findings, which highlight the importance and pathways of metaverse and artificial companions in education and society sectors.

Teacher Forum

There are 2 sessions of teacher paper presentations with 7 papers in the conference. K-12 teachers share best practices and key challenges in implementing metaverse and artificial intelligence education in various countries/regions.

On behalf of the Conference Organizing Committee, we would like to express our gratitude towards all speakers as well as paper presenters for their contribution to the success of MetaACES 2025.

We sincerely hope everyone enjoys and gets inspired from MetaACES 2025.

Prof. Siu Cheung KONG
The Education University of Hong Kong, Hong Kong SAR
Conference Chair of MetaACES 2025

Prof. Yu-Ju LAN
National Taiwan Normal University, Taiwan
Program Chair of MetaACES 2025

Prof. Jianhua ZHAO
Southern University of Science and Technology, China
Program Co-Chair of MetaACES 2025

Prof. Yanjie SONG
The Education University of Hong Kong, Hong Kong SAR
Track Chair of MetaACES 2025

Prof. Tanja MITROVIC
University of Canterbury, New Zealand
Track Chair of MetaACES2025

Table of Contents

Full Paper

Short Paper

Poster Paper

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Challenges and Strategies of ChatGPT in Foreign Language Teaching: A Literature Review Based on Articles Published in the Past Three Years

Wen SHEN[1], Chunlin LEI[2*], Meimei WANG
[1,2,3]Shanghai University of International Business and Economics, Shanghai
[1]shenwen0114@163.com，[2]leichunlin@suibe.edu.cn

Abstract: *It is indisputable that artificial intelligence（AI）, represented by ChatGPT, is exerting an increasingly significant influence on foreign language teaching and learning. However, while empowering foreign language education, intelligent technologies also bring about a wide range of challenges. Based on 70 research papers on the application of ChatGPT published from 2023 to 2025, this paper focuses on and reviews the problems and drawbacks of using artificial intelligence in foreign language teaching and learning. Corresponding measures and strategies are also summarized. The significance of this paper lies in encouraging all stakeholders to proactively prepare, leverage strengths and mitigate weaknesses to better utilize artificial intelligence for learning.*

.

Keywords: ChatGPT, Foreign language teaching, Challenges, Strategies

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

ChatGPT 在外语教学中的挑战与应对 – 一项基于近三年文献的研究

沈文 [1]，雷春林 [2*]，汪美美 [3]
[1,2,3] 上海对外经贸大学，上海
[1]shenwen0114@163.com，[2]leichunlin@suibe.edu.cn

【摘要】 毋庸置疑，以 ChatGPT 为代表的人工智能在外语教学与学习中发挥着越来越大的作用。但是，智能技术赋能外语教育的同时，其挑战也广泛存在。本研究基于 2023-2025 年发表的 70 篇有关 ChatGPT 应用的研究文献，聚焦和梳理了外语教学与学习中使用人工智能存在的问题与不足，同时，也总结了一些应对的措施和策略。本文的意义在于，促使所有利益攸关方未雨绸缪，扬长避短，更好地利用人工智能促进学习。

【关键词】 ChatGPT；外语教学；挑战；策略

1. 导言

ChatGPT 问世以来，在辅助语言学习方面展示了很多优势，比如，ChatGPT 能够生成语言学习文本、提供互动与反馈机会、提供知识以及翻译、拼音及解释等帮助功能(蔡薇，2023)。尽管 ChatGPT 赋能外语学习作用明显，也有专家就其应用局限发出了警示，但是人们对其潜藏的负面问题，尚未形成全面、深入的认知，在应对方法上，也缺乏行之有效的策略。因此，本研究将聚焦于 ChatGPT 在外语教学与学习场景下的局限性，深入探讨以下两大研究问题：(1) ChatGPT 在外语教学中的主要挑战有哪些？(2) 如何应对这些挑战？

回答上述两个问题，可以为我们更好地利用人工智能促进学习夯实基础。

2. 研究方法

本研究选择 Web of Science 和知网两大数据库，采用关键词文献检索策略，在 Web of Science 输入“ChatGPT and foreign language learning”、“ChatGPT and foreign language teaching”，检索到 135 条结果；在知网输入“人工智能与外语学习”、“人工智能与外语教学”，限定来源为北大核心和 CSSCI，得到 32 条结果。经去重、标题摘要初筛与全文筛选，在 Web of Science 和知网最终分别获取 54 篇、16 篇有效文献，共计 70 篇 。研究文献的筛选遵循严谨的纳入与排除标准，具体内容如下（表 1）：

表 1　纳入和排除标准

标准	纳入标准	排除标准
重复与否	非重复	重复
文章主题	ChatGPT 和外语教学都涉及的文章	与外语教学无关的文章
文章内容	学习的外语为英语的文章	学习的外语为非英语的文章
文章类型	学术文章	非学术文章
文章语言	英语、汉语	英语和汉语以外的文章

3. 分析与结果

3.1. 人工智能用于外语教学的挑战

笔者将筛选的 70 篇文献进行了标注，然后通篇阅读，在阅读的过程中，对文中涉及的 ChatGPT 或者人工智能在外语教学中的不足、缺陷或者关切等进行重点抓取和标记，同时运

用 AntConc 文本分析工具，一共统计出 8 类高频的挑战或关切（表 2）。这些挑战包括，信息准确性问题、伦理问题、学术诚信问题、阻碍批判性思维、损害创造力、技术问题、文化背景不足以及缺乏情感交流等。

表 2　人工智能在外语运用中的挑战（频度统计）

	Type	Rank	Freq	Range
1	信息准确性问题	1	57	1
2	伦理问题	2	42	1
3	学术诚信问题	3	42	1
4	阻碍批判性思维	4	25	1
5	损害创造力	5	16	1
6	技术问题	6	12	1
7	文化背景不足	7	11	1
8	缺乏情感交流	7	11	1

从频度上来看，人们最关心或者诟病最多的是信息准确性问题。这是因为，ChatGPT 可能会提供不准确的回答，也不能提供准确的引用和参考文献，因此它不是很值得信任的工具（Mohamed & Amr M., 2023）。ChatGPT 有时会很正经地说一些废话，提供不实信息，Guo (2024) 提到，ChatGPT 会给出看似合理但实际错误的信息，这会让人们质疑其可信度。胡壮麟（2023）也认为，它的训练数据来源于各种语料库和互联网文本，它的反馈可能包含一些不准确或不合适的信息。

伦理问题也是人工智能用户非常担心的问题。首先是数据隐私与安全的问题。郝磊、温志强、王妃等(2023)指出，师生与 ChatGPT 互动的数据有可能被保存下来，如果这些数据被泄露，可能会严重威胁师生的数据隐私和安全。此外，还有偏见问题。D’Agostino 在 Kim et al.(2024)的文章中表示，语言偏见是一个严重的问题，因为人工智能系统更偏向英语语言，这样非英语母语者在信息获取与知识提升方面就会落后于英语母语者。Almanea, M.（2024）的研究表明，ChatGPT 似乎带有“白人、西方”视角，这或许会使它给出的回复存在文化偏见。

学术诚信问题是和伦理问题并列为第二位的挑战。学术诚信问题包括作弊、剽窃和抄袭等不诚实的行为。一些研究指出，学习者经常用 ChatGPT 做作业，这可能会导致抄袭，从而引发学术不端和诚信的问题（Moon, D. , 2024）。美国网络调查公司对 1223 名美国大学生开展的一项调查显示，30%的学生曾用 ChatGPT 完成作业（杨连瑞，2024）。就连 Chomsky 也指出，ChatGPT 本质上属于高科技剽窃，是一种更具隐蔽性的抄袭行径（胡加圣和戚亚娟，2023）。

阻碍批判性思维也是使用人工智能过程中的隐忧。例如，初、中级学习者往往更易过度依赖 ChatGPT，他们常常不假思索地接纳其反馈，从而抑制批判性思维与自主学习能力的发展（Moon, D.,2024）。“信息茧房”是指人们只重点关注自己感兴趣的领域，难以接触到多元的思想，容易形成片面、偏执的认知。在此现象下，若盲目接纳 ChatGPT 输出的内容，会弱化其批判性思维能力（杨连瑞，2024）。

损害创造力也是人们关注的缺点之一。ChatGPT 既有积极的一面，也有消极的一面。若能合理运用，它将成为提升效率的得力助手；否则，便可能助长惰性，成为抑制创造力的枷锁（陈新仁，2024）。如果全盘套用 ChatGPT 给的答案，原本活跃的创造力思维将逐渐僵化。正如任伟、刘远博和解月(2024）所说，如果学生在修改作文时过度依赖 ChatGPT 而非独立思考，长此以往，会其写作与创新能力的提升有害无利。

技术问题也是不容忽视的挑战之一。在使用 ChatGPT 的过程中，会存在一些技术上的问题，影响用户的使用体验，比如，响应进程缓慢，在解析某些语言结构时会陷入困境，存在易被黑客攻击的风险，并且运行时需要稳定、高质量的网络连接支持等问题（Mohamed & Amr M., 2023）。网络连接不稳定会导致在线课堂中教师示例操作卡顿，这会打断学生的学习思路，破坏教学节奏的连贯性，消磨学生的学习热情。

文化背景不足也是人工智能的局限之一。例如，张震宇和洪化清(2023)认为，ChatGPT 的关注重点倾向于语言形式，却对文化内涵有所忽视，因此有可能输出违背跨文化交际规则的信息，容易误导学生学习者，从而降低他们的跨文化交际能力。胡壮麟(2023）也提到，ChatGPT 在翻译特定文化背景、词汇及习惯用语时，难以全面权衡原文与译文在语言、文化层面的差异，其翻译质量往往逊于人工翻译。

缺乏情感交流也是高频提及的挑战之一。尽管人工智能工具能提供即时反馈，但却无法营造传统语言课程所具备的人际互动氛围，这会使学生难以在情感和文化层面与所学语言建立起紧密联系（Mohamed & Amr M., 2023）。按照秦颖（2023）的看法，学习语言需要实践，人机对话难以模仿真实情境下师生间的情感交流互动，如果缺乏情感交流，学习者只能停留在生硬的词句学习上，难以内化知识，这会阻碍构建地道语言思维，导致交流障碍。

3.2. 应对措施与建议

笔者在现有文献中标记人工智能使用过程中解决问题的对策、建议和展望等关键信息，利用 Excel 表格进行摘录，然后使用主题分析法，进行归类，从而总结出针对各种挑战的建议。

（1）求证信息正误，查阅权威资料

为应对信息准确与否的问题，首先，教师要引导学生养成“质疑”的习惯。批判性地评估 ChatGPT 输出的信息，而非盲目接受其提供的所有内容。Hınız & Gökhan（2024 ）表示，外语教师应聚焦数字素养、媒体素养及批判性思维等高阶思维与学习技能的教学，从而提高鉴别信息质量的能力。其次，可通过对比 ChatGPT 的回复与其他可靠来源（如学术文献、权威词典）的信息，培养信息鉴别准确与否的能力。例如 Kim et al.（2024)指出，学生可查阅教科书、学术期刊或更具权威性的网站来验证信息的准确性。

（2）确立明晰规范，深化伦理教育

为应对伦理方面的挑战，学校应确立明晰的伦理规范，制定隐私保护相关的政策，防止未经授权的数据被用于传播虚假信息或进行网络欺凌。Tram et al.（2024）指出，构建完善的数据隐私保护机制，能够有效增进用户间的信任与信心，这是实现广泛应用的关键所在。此外，重视对学生进行伦理道德教育，帮助学生理解人工智能技术的潜在伦理风险。李佐文（2024)认为，外语教学，不仅是语言知识的传授，更是一个提高学生人文素养和道德品质的过程。外语教师应在教学中融入道德与人文教育，增强跨文化交流能力与社会责任感。

（3）明确诚信边界，加强道德引导

面对学术诚信的担忧，学校和教育机构应尽快制定明确的政策，规范人工智能的使用边界（如禁止用于考试或作业）。例如，周明阳（2023）在胡壮麟（2023）的文章中提及，国外众多名校纷纷出台禁令，严禁学生在校内借助类似 ChatGPT 的人工智能工具完成教学任务以及参加考试。同时，加强学术诚信教育，帮助学生理解学术道德的重要性，引导学生通过自身努力来获取知识，形成诚信为本的学术风气。依据 Moon, D.(2024)的见解，加强道德教育，对防范剽窃行为而言极为必要。通过这类教育，能引导学习者以负责任的态度，合理运用 ChatGPT 生成的内容，杜绝将其成果据为己有等不当行径。

（4）培养批判思维，引导反思实践

为克服阻碍批判性思维的问题，教师可鼓励学生针对 ChatGPT 给出的回答，深度分析背后的逻辑。比如，当 ChatGPT 提供某个观点时，引导学生追问支撑该观点的论据或理论基础，促使学生主动分析内容，培养批判性思维。根据焦建利和陈婷（2023）的研究，教师可以鼓励学生借助 ChatGPT 生成答案，但要对这些答案进行深度分析和合理改进，并分享改进的思路。通过这一方式，学生的批判性思维能力和协作能力就能得到提升。另外，教师可引导学生积极参与反思实践。在课堂或小组学习中，定期回顾使用人工智能的过往经历，例如是否存在不假思索接受其观点，未从其他角度深入分析的情况。通过持续反思，提升批判性思维，有效削弱人工智能对思维可能产生的负面影响。

（5）激发创造活力，布置创意任务

为应对损害创造力的缺点，教师可以鼓励学生将 ChatGPT 生成的内容作为灵感来源，而不是最终成果，思考内容背后的逻辑是否存在局限，并通过自己的创意进行改编或扩展，从而打破思维定式，唤醒创造活力。同样，教师也可以布置创意任务（如写作、演讲、角色扮演），要求学生结合人工智能的输出与自己的原创想法。比如在角色扮演中，学生可利用 ChatGPT 能提供的剧情框架与角色设定，但要自己增添性格细节，深挖角色内心，结合自身经历与想象，赋予角色独一无二的性格，这一过程可以极大地锻炼创造力。

（6）排除技术故障，提升技术素养

为处理技术有关的问题，学校应提供技术支持保障，比如构建专属的高速校内网络通道，确保在使用 ChatGPT 这类在线工具时，网络稳定流畅；安排专业的信息技术人员定期对网络进行监测与维护，及时处理潜在的网络故障隐患；还可与技术平台沟通，反馈使用中出现的问题，寻求优化方案。此外，教师也要提升技术素养。正如 Kovačević & Darko（2023）所说，使用 ChatGPT 时，需要一定的计算机知识，包括了解基本的计算机操作，熟悉软件的安装及使用方法，懂得浏览互联网和运用各类网络工具。李佐文（2024)也表示，外语教师应提升运用 ChatGPT 等大模型技术的能力，学会借助其高效分析、解决问题，利用大数据与 5G 打造智慧课堂，开启外语学习新模式。

（7）补充多元文化，组织文化活动

面对 ChatGPT 文化背景不足的问题，教师要补充多元文化资料，收集不同国家的历史故事、民俗风情、文学艺术作品等资料，与 ChatGPT 提供的内容相结合。张震宇和洪化清（2023）在于银磊和饶辉（2023）的文章中都提到，大学外语教师可以利用 ChatGPT 等人工智能产品的优势，倡导深化文化、历史等人文知识与跨学科知识的学习与探究。另外，可以组织文化活动来提升学生的文化素养，比如，举办"文化沙龙"，邀请外教或留学生分享真实的文化体验，弥补人工智能文化背景不足的问题。

（8）加强人际互动，增进情感交流

由于缺乏情感交流，ChatGPT 无法给予学生情感支持与鼓励，这使得学生在学习过程中，很难与学习内容产生深度的情感共鸣。为应对这一弊端，教师需要充分发挥主导作用。即便聊天机器人等人工智能技术愈发成熟，教师在助力学生发展社交、情感、认知和自主能力方面，依旧不可或缺(Hınız & Gökhan，2014)。依照胡壮麟(2023)的观点，教师能了解学生的学习情况及需求，据此开展针对性教学，还会给予学生情感支持与鼓励，助力他们攻克学习难关。因此，教师要增进与学生的情感互动，在课堂上及时给予学生肯定的微笑、鼓励的话语，结合自身或学生的生活实例讲解外语知识，让学生在充满人文关怀的环境中学习外语。另外，教师还可以开展小组讨论、情景对话等活动，让学生在面对面的情境中增进情感交流。

4. 结论

本研究通过对最新的有关人工智能在外语教学中应用的文献分析，找出了以 ChatGPT 为代表的人工智能在外语教学应用中的八个最主要的局限和关切，并根据文献，归纳了一系列应对的策略，本文的意义在于帮助教师、学生和学校等各个利益攸关方更好地认识到人工智能在带给我们无限机遇的同时，也带来了巨大的挑战。教育界一定要未雨绸缪，有针对性地加强人工智能教育，提升人工智能素养，让人工智能更好地为人类的学习服务。当然，本研究还存在诸多不足，例如挑战和建议都是基于前人研究成果，而且可能没有把最新成果纳入。同时，在分析中，也只是依赖高频词和主题分析，存在一些主观判断性的结论。未来研究可进一步扩大文献检索范围，跟踪最新研究动态，用交叉编码检验等形式提高信度，从而确保研究的全面性和时效性，为人工智能在外语教育中的实践提供更具前瞻性的方向和依据。

注：本论文是上海对外经贸大学国际商务外语学院 2025AI+数智课程建设项目的部分成果。

参考文献

蔡薇（2023）。ChatGPT 环境下的汉语学习与教学。**语言教学与研究**，(4)，13-23。

陈新仁(2024)。ChatGPT 与二语语用教学。**外语教学理论与实践**，(04)，16-29。

胡加圣和戚亚娟(2023)。ChatGPT 时代的中国外语教育：求变与应变。**外语电化教学**， (01)，3-6。

郝磊、温志强、王妃等(2023)。ChatGPT 类人工智能催生的多领域变革与挑战（笔谈）。天津师范大学学报(社会科学版)，(04)，8-23.

胡壮麟（2023）。ChatGPT 谈外语教学。**中国外语**，**20**(3)，1, 12-15。

焦建利和陈婷(2023)。大型语言模型赋能英语教学：四个场景。**外语电化教学**，(02)，12-17。

李佐文(2024)。ChatGPT 赋能外语教学：场景与策略。北京第二外国语学院报,46(01),109-118。

秦颖 (2023)。人机共生场景下的外语教学方法探索——以 ChatGPT 为例。**外语电化教学**，**2**， 24-29。

任伟、刘远博和解月(2024)。英语写作教学中 ChatGPT 与教师反馈的对比研究。**外语教学理论与实践**，(04)，30-38。

杨连瑞(2024)。Chatgpt 大语言模型背景下的二语习得。**现代外语**，**47**(4)， 578-584。

于银磊和饶辉(2023)。智能时代大学外语教师的角色危机与身份重塑。**外语电化教学**，（03），79-85。

张震宇和洪化清(2023)。ChatGPT 支持的外语教学：赋能、问题与策略。**外语界**，(02)，38-44。

Almanea, M. (2024). Instructors' and learners' perspectives on using ChatGPT in English as a foreign language courses and its effect on academic integrity. *Computer Assisted Language Learning*, 1-26.

Bai, L., & Wei, Y. (2024). Exploring EFL Learners' Integration and Perceptions of ChatGPT's Text Revisions: A Three-Stage Writing Task Study. *IEEE Transactions on Learning Technologies*.

Dizon, G. (2024). ChatGPT as a tool for self-directed foreign language learning. *Innovation in Language Learning and Teaching*, 1-17.

Guo, X. (2024). Facilitator or thinking inhibitor: understanding the role of ChatGPT-generated written corrective feedback in language learning. *Interactive Learning Environments*, 1-19.

Hınız, G. (2024). A year of generative AI in English language teaching and learning-A case study. *Journal of Research on Technology in Education*, 1-21.

Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies,* 1-24.

Kim, H. S., & Kim, N. Y. (2024). Exploring the Effects of ChatGPT on Video-Making Projects in an EFL Course. *Asia-pacific Journal of Convergent Research Interchange (APJCRI)*, 737-750.

Klimova, B., Pikhart, M., & Al-Obaydi, L. H. (2024). Exploring the potential of ChatGPT for foreign language education at the university level. *Frontiers in Psychology,* 15, 1269319.

Kovačević, D. (2023, March). Use of ChatGPT in ESP teaching process. *In 2023 22nd International Symposium Infoteh-Jahorina (INFOTEH)* (pp. 1-5). IEEE.

Kucuk, T. (2024). ChatGPT integrated grammar teaching and learning in EFL classes: A Study on Tishk International University Students in Erbil, Iraq. *Arab World English Journal (AWEJ)*, 100-111.

Liu, X. (2024). Navigating uncharted waters: Teachers' perceptions of and reactions to AI-induced challenges to assessment. *The Asia-Pacific Education Researcher*, 1-12.

Mohamed, A. M. (2024). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: perceptions of EFL Faculty Members. *Education and Information Technologies*, 29(3), 3195-3217.

Moon, D. (2024). Enhancing Cyber Foreign Language Writing Education with ChatGPT. *International Journal of Internet, Broadcasting and Communication*, 16(4), 475–481. https://doi.org/10.7236/IJIBC.2024.16.4.475

Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. Cogent Education, 11(1), 2410101.

Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigation in Health, Psychology and Education*, 13(9), 1937-1960.

Teng, M. F. (2025). Metacognitive Awareness and EFL Learners' Perceptions and Experiences in Utilising ChatGPT for Writing Feedback. European Journal of Education, 60(1), e12811.

Tram, N. H. M., Nguyen, T. T., & Tran, C. D. (2024). ChatGPT as a tool for self-learning English among EFL learners: A multi-methods study. *System*, 103528.

Wang, Y. (2024). Cognitive and sociocultural dynamics of self-regulated use of machine translation and generative AI tools in academic EFL writing. *System*, 126, 103505.

Woo, D. J., Wang, D., Guo, K., & Susanto, H. (2024). Teaching EFL students to write with ChatGPT: Students' motivation to learn, cognitive load, and satisfaction with the learning process. *Education and information technologies*, 1-28.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

ChatGPT Addiction, Self-Efficacy, and Continuous Intention: Differences in Gender, Academic Majors, and Purposes of Use

Pin Hui JIANG[1*], Jon Chao HONG[2]
[1, 2] National Taiwan Normal University, Taiwan
pinhuijiang1117@gmail.com*, tcdahong@gmail.com

Abstract: *With the increasing adoption of AI tools like ChatGPT and DeepSeek, questions have been raised about their impact on cognitive and affective learning outcomes. This study explores the role of ChatGPT in students' use of AI for English grammar learning, investigating differences in variables such as gender, academic discipline, and purpose of use, based on Trait Activation Theory. The study utilizes t-tests and ANOVA to analyze ChatGPT usage across three dimensions: ChatGPT addiction, ChatGPT self-efficacy, and continuous intention, comparing variables like gender differences, academic major, and primary use of the platform. Significant gender differences were found in technological innovation, ChatGPT addiction, and ChatGPT self-efficacy, with females outperforming males across all dimensions. Academic major also showed significant differences, with science majors scoring higher than liberal arts majors. In comparing primary platform use, significant differences were noted in ChatGPT addiction and continuous intention. Scheffé's post hoc analysis indicates that time passing (TP) scores higher than answer seeking (AS) and creativity (C), with the order TP > AS > C. For sustained engagement, significant differences were observed between academic writing (AW) and time passing (TP), with scores ranked as TP > AW > C.*

Keywords: ChatGPT, self-efficacy, addiction, continuous intention, chatbot-based learning

1. Introduction

Learning through chatbots has been widely adopted and proven effective in enhancing knowledge and skills while fostering learners' engagement (Hwang & Chang, 2023). Zhang et al. (2023) explored chatbot-based learning's impact on intrinsic motivation, self-efficacy, knowledge development, and learner engagement. Chatbot-based learning demonstrates greater potential with the integration of generative AI, which learns from data, generates meaningful content, and personalizes it based on specific needs. ChatGPT, a leading example of generative AI, excels in natural language processing, including translation, question-answering, and cloze tasks (Brown et al., 2020; Zhang et al., 2025). For example, Tram et al. (2024) noted that English as a Foreign Language learners used ChatGPT for primary and secondary learning, supporting reading, writing, vocabulary, and grammar. Moreover, Annamalai et al. (2024) indicated that autonomy and relatedness predicted English as a Foreign Language students' competence that motivate them for continuous intention, with autonomy being the most important and highest-performing factor. Despite valuing ChatGPT, most learners face challenges using it, highlighting the need to explore adoption factors. Morris et al. (2023) indicated that social media addiction can result in negative emotional and behavioral consequences due to extended and excessive engagement. However, few studies have examined ChatGPT addiction across different age groups, genders, majors, and usage purposes. Therefore, this study explored the factors that may influence individuals' ChatGPT self-efficacy (ChatGPT-SE), ChatGPT addiction, and continuous intention (ChatGPT-CI) in the context of English grammar learning.

1.1 The purpose of this study

Bandura's socio-cognitive theory conceptualizes self-efficacy as an individual's confidence in their ability to perform actions required to manage future situations (Bandura, 1977). AI self-efficacy refers to individuals' overall confidence in their capability to utilize and engage with AI (Wang & Chuang, 2024). Within the educational context, ChatGPT SE is described as students' belief in their competence to utilize, engage with, comprehend, and acquire knowledge about AI technologies and applications. Additionally, it explores whether individuals exhibit addiction to ChatGPT and intentions for continuous use under different usage purposes. Building on this, the study aims to investigate how factors such as gender, usage purposes, and academic majors influence ChatGPT-SE, addiction, and CI. By analyzing these variables, the research seeks to uncover patterns and differences in how diverse demographic and contextual factors shape users' interactions with AI technologies. Furthermore, the findings will provide valuable insights into the educational implications of these influences, highlighting how tailored strategies can be developed to enhance AI integration in education, promote equitable access, and address the unique needs of different student groups.

1.2 Hypotheses

Trait Activation Theory (Tett & Burnett, 2003) was applied to explain how psychological and personal traits shape individuals' behavior across different contextual cues (Bisht & Mahajan, 2021). It was identified as a suitable theoretical framework for analyzing the impact of self-construal on the intention to use technology tools. Building on TAT, this study aimed to investigate the differences across gender, academic majors, and usage purposes. To explore their difference, three hypotheses were proposed as follows:

H1: There are significant difference in ChatGPT-addiction, SE, and CI across different gender.

H2: There are significant differences in ChatGPT-addiction, SE, and CI across different academic majors.

H3: There are significant differences in ChatGPT-addiction, SE, and CI across different usage of purposes.

2. Method

2.1. Participants and Procedure

A total of 213 responses were collected over a four-week period. After excluding invalid submissions, including incomplete responses, 183 valid returns were used for statistical analysis. Among the 183 valid responses, 101 (55.2%) were from male participants and 82 (44.8%) from female participants. In terms of educational level, 126 respondents (68.9%) were undergraduate students, while 57 (31.13 %) were graduate students. Regarding age, 126 participants (68.9%) were between 18 and 23 years old, and 57 (31.1%) were over 23 years old. Concerning academic majors, 80 participants (43.71%) were enrolled in science, whereas 103 (56.3%) were from liberal disciplines. Regarding the ChatGPT versions used, 28 participants (15.3%) had experience with ChatGPT-4o, while the remaining 155 (84.7%) had used ChatGPT-4.

2.2. Development of the Questionnaires

The questionnaire was modified and translated from previous studies and employed a 5-point Likert scale, where “strongly disagree” counted as 1 point, “neutral” counted as 3 points, and “strongly agree”. After its development, five domain experts evaluated its face validity. Following this, reliability and validity tests were performed to confirm its accuracy.

ChatGPT-addiction: Digital addiction refers to a form of dependency arising from the excessive and problematic use of modern technological devices, a topic that has been extensively debated since 2013 (Ime et al., 2024). To explore ChatGPT addiction, the scale was developed based on the digital addiction scale for children created by Hawi et al. (2019), which assesses various behaviors exhibited by children when using digital devices. The reliability of the scale was evaluated, with a Cronbach’s α of 0.936.

ChatGPT-SE:

Digital Self-Efficacy, referring to an individual's confidence in their capability to effectively utilize digital tools, has gained growing significance (Ulfert-Blank & Schmidt, 2022). In this study, we developed four items to measure ChatGPT-SE, for example, "When I encounter issues using ChatGPT, I feel that I can resolve them" and "When I come across new issues with using ChatGPT, I am confident I can find a way to resolve them."

ChatGPT-CI:

The ongoing intention to use digital systems is essential for maintaining long-term educational benefits (Henríquez & Hilliger, 2024). This study adopted the continued intention to use facial recognition systems index (Hong et al., 2022), which demonstrated high reliability with a Cronbach's α of 0.961. To explore the influence of ChatGPT-CI in this study. Four items were included to examine ChatGPT-CI.

2.3. Data Analysis

In this study, Confirmatory Factor Analysis (CFA) was applied, using IBM SPSS AMOS 20 statistical software to assess the reliability and validity of the research instruments. Considering the Hawthorne effect, a control group was not included for comparison. Therefore, a single-group quasi-experimental design was adopted to test the research hypotheses (Jones, 1992). For data analysis, a one-way ANOVA was conducted using SPSS version 23 to compare the differences among the three groups.

3. Result

3.1. Construct Reliability and Validity

The study initially evaluated the reliability of the statistical data through Composite Reliability (CR) and Cronbach's α coefficient. Higher CR values indicate stronger correlations between the measurement items. According to the guidelines provided by Hair et al. (2019), α values above 0.6 are provisionally acceptable, with adjustments to wording or the addition of items permitted if necessary. In this study, all Cronbach's α values exceeded 0.86, and CR values were greater than 0.91. Secondly, convergent validity was examined by verifying whether the Average Variance Extracted (AVE) values were above 0.5, following Hair et al. (2024) recommendations. Additionally, the factor loadings of all items within three constructs were found to be significant and exceeded 0.71. Both criteria were met, indicating acceptable convergent validity, as shown in Table 1.

Table 1. Reliability and validity analysis.

Construct	*M*	*SD*	Cronbach's α	CR	FL	AVE
ChatGPT-addiction	2.02	.96	.91	.94	.71	.79
ChatGPT-SE	3.50	.80	.86	.91	.78	.71
ChatGPT-CI	3.43	.87	.86	.91	.80	.71

3.2. Analysis of Variance

This study applied t-tests and one-way ANOVA to analyze the differences in background variables across various dimensions. The results are presented in Tables 2 to 4. As shown in Table 2, gender differences were significant in ChatGPT addiction (t =2.11*, $p < .05$) and ChatGPT self-efficacy (t =2.07*, $p < .05$), with female participants scoring higher than males. Table 3 shows that there were significant differences in major fields of study across all dimensions (t =3.83***, $p < .001$), with science majors scoring higher than liberal majors. Lastly, as presented in Table 4, the primary purpose of using ChatGPT showed significant differences in ChatGPT addiction (F =3.99**, $p < .01$). Post hoc analysis using Scheffe's test indicated significant differences between "time passing" and "answer seeking" ($p < .05$) as well as "time passing" and "creativity" ($p < .01$), with the mean scores showing that "time passing" was higher than "answer seeking", which was higher than "creativity". Regarding the intention to continue using ChatGPT, significant differences were also observed (F =5.84***, p < .001). Post hoc analysis revealed that "creativity" significantly differed from "academic writing" (p < .01) and "time passing" (p < .01), with the mean scores indicating that "time passing" was the highest, followed by "academic writing," and then "creativity".

Table 2. Gender differences across constructs.

Construct	Gender	N	M	SD	*t*	*p*
ChatGPT-addiction	male	101	1.88	.91	2.11	.04*
	female	82	2.20	1.08		
ChatGPT-SE	male	101	3.39	.85	2.07	.04*
	female	82	3.63	.72		
ChatGPT-CI	male	101	3.33	.94	1.78	.08
	female	82	3.56	.77		

Table 3. Science and liberal major differences across constructs.

Construct	Major	*N*	*M*	*SD*	*t*	*p*
ChatGPT-addiction	science	104	1.79	.81	3.60	.00***
	liberal	79	2.33	1.14		
ChatGPT-SE	science	104	3.31	.77	3.83	.00***
	liberal	79	3.75	.78		
ChatGPT-CI	science	104	3.18	.84	4.85	.00***
	liberal	79	3.77	.79		

Table 4. Differences across constructs by primary ChatGPT usage purpose.

Construct	usage purpose	*N*	*M*	*SD*	*SE*	F	*p*
ChatGPT-addiction	answer seeking	113	1.99	.98	.09	3.99	.01**
	academic writing	32	2.08	1.07	.19		
	time passing	12	2.88	1.05	.30		
	creativity	26	1.71	.79	.15		
ChatGPT-SE	answer seeking	113	3.43	.84	.08	1.01	.39
	academic writing	32	3.56	.69	.12		
	time passing	12	3.81	.43	.12		
	creativity	26	3.56	.87	.17		
ChatGPT-CI	answer seeking	113	3.41	.87	.08	5.84	.00***
	academic writing	32	3.69	.68	.12		
	time passing	12	4.00	.67	.19		
	creativity	26	2.94	.90	.18		

4. Discussion

The study investigated ChatGPT addiction across genders, academic majors, and usage purposes. It further explored the traits that may impact individuals' ChatGPT-SE and ChatGPT-CI. This study applied Trait Activation Theory, gathered 183 valid responses, and utilized *t*-test and ANOVA to conduct an empirical investigation. Four hypotheses were verified as follows.

Andreassen et al. (2016) define internet addiction as the excessive use of social media, where an individual devotes so much time and energy to it that it impacts their daily activities. The results showed that the average of female's ChatGPT-addiction was M = 2.20, SD = 1.08, ChatGPT-SE ChatGPT-addiction was M = 3.63, SD = 0.72. These averages indicated that there were gender significant differences across addiction and self-efficacy, indicating that male had lower levels. This finding is consistent with the study conducted by Kor and Shoshani (2023), which indicated that females spent more hours on the internet. Additionally, Wang and Zhang (2023) discovered that females had higher averages in sense of efficacy, learning quality, and perceived learning achievement. Considering ChatGPT as a form of social media and analyzing gender differences in ChatGPT addiction and SE, this study found support for H1. By examining the differences between science and liberal majors in three constructs, the study found significant differences in ChatGPT addiction, SE, and CI across different majors. That is, Valeri et al. (2025) revealed that ChatGPT is widely used by participants across science subjects, with notable adoption in biology, primarily serving as a tool to aid in the comprehension of concepts. However, there were significant differences in self-efficacy, continuous intention and addiction between science and liberal majors, with science majors showing significantly lower averages in both dimensions. Future research can further explore this issue, providing suggestions for researchers and educators in STEM teaching. Lastly, Moravec et al. (2024) explored how the use of ChatGPT can reveal the impact of deep learning on individuals' intentions to use this tool for different purposes. For instance, the primary purpose for using ChatGPT was for passing time, followed by work. This finding was consistent with the present study. However, the study also explored various aspects of ChatGPT addiction and CI. Significant differences were found across the three constructs for different usage purposes. The results showed that using ChatGPT for passing time and academic writing had higher average scores compared to other purposes, indicating that this group exhibited higher levels of addiction and a stronger intention to continue using ChatGPT.

The primary aim of this study was to examine the role of ChatGPT in students' use of AI for learning English grammar, focusing on the differences in variables like gender, academic discipline, and usage purposes. In future work, based on the present research showing that different purposes for using ChatGPT are associated with varying levels of addiction, self-efficacy, and continuous usage, it may be suggested that researchers further explore the underlying factors influencing these differences, such as individual learning styles, motivation, and contextual variables. Additionally, teachers may consider integrating ChatGPT strategically into their instructional practices to enhance students' learning outcomes while promoting healthy usage habits and fostering digital literacy skills.

References

Andreassen, C. S., Billieux, J., Griffiths, M.D., Kuss, D.J., Demetrovics, Z., Mazzoni, E., & Pallesen, S. (2016). The relationship between addictive use of social media and video games and symptoms of psychiatric disorders: a large-scale cross-sectional study. *Psychology of Addictive Behaviors, 30* (2), 252-262. https://doi.org/10.1037/adb0000160

Annamalai, N., Bervell, B., Mireku, D. O., & Andoh, R. P. K. (2024). Artificial intelligence in higher education: Modelling students' motivation for continuous use of ChatGPT based on a modified self-determination theory. *Computers and Education: Artificial Intelligence*, 100346. https://doi.org/10.1016/j.caeai.2024.100346

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychol. Rev.*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bisht, N. S., & Mahajan, A. (2021). Shared stressors and core self-evaluations: A trait activation perspective on employee performance. Journal of Business Research, 131, 103-111. https://doi.org/10.1016/j.jbusres.2021.03.053.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., … Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901. https://proceedings.neurips.cc/paper/2020/ha sh/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, *18*(1), 39-50. https://doi.org/10.1177/002224378101800104

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS SEM. *European business review*, *31*(1), 2-24. https://doi.org/10.1108/EBR-11-2018-0203

Hawi, N. S., Samaha, M., & Griffiths, M. D. (2019). The digital addiction scale for children: Development and validation. Cyberpsychology, Behavior and Social Networking, 22(12), 771–778. https://doi.org/10.1089/cyber.2019.0132

Henríquez, V., & Hilliger, I. (2024). Blended learning in rural K－12 education: Stakeholder dynamics and recommendations. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12963

Hwang, G. J., & Chang, C. Y. (2023). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments, 31(7)*, 4099–4112. https:// doi.org/10.1080/10494820.2021.1952615

Ime, Y., Akyıl, Y., & Caglar, A. (2024). The examination of the relationships among digital addiction, loneliness, shyness, and social anxiety in adolescents. *Anales de Psicología/Annals of Psychology*, *40*(2), 236-241. https://doi.org/10.6018/analesps.579801

Jones, S. R. (1992). Was there a Hawthorne effect? *American Journal of sociology*, *98*(3), 451-468. https://doi.org/10.1086/230046

Kor, A., & Shoshani, A. (2023). Moderating the impact of the COVID-19 pandemic on children's and adolescents' substance use, digital media use, and mental health: A randomized positive psychology addiction prevention program. *Addictive Behaviors*, *141*, 107660. https://doi.org/10.1016/j.addbeh.2023.107660

Moravec, V., Hynek, N., Gavurova, B., & Rigelsky, M. (2024). Who uses it and for what purpose? The role of digital literacy in ChatGPT adoption and utilisation. *Journal of Innovation & Knowledge*, *9*(4), 100602. https://doi.org/10.1016/j.jik.2024.100602

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology, 88,* 500–517. http://dx.doi.org/ 10.1037/0021-9010.88.3.500

Tram, N. H. M., Nguyen, T. T., & Tran, C. D. (2024). ChatGPT as a tool for self-learning English among EFL learners: A multi-methods study. *System*, 103528. https://doi.org/10.1016/j.system.2024.103528

Ulfert-Blank, A.-S., & Schmidt, I. (2022). Assessing digital self-efficacy: Review and scale development. Computers & Education, https://doi.org/10.1016/j.compedu.2022.104626

Valeri, F., Nilsson, P., & Cederqvist, A. M. (2025). Exploring students' experience of ChatGPT in STEM education. *Computers and Education: Artificial Intelligence*, *8*, 100360. https://doi.org/10.1016/j.caeai.2024.100360

Wang, M., & Zhang, L. J. (2023). Understanding teachers' online professional learning: A "community of inquiry" perspective on the role of Chinese middle school teachers' sense of self-efficacy, and online learning achievement. *Heliyon*, *9*(6). https://doi.org/10.1016/j.heliyon.2023.e16932

Zhang, R., Zou, D., & Cheng, G. (2023). A review of chatbot-assisted learning: Pedagogical approaches, implementations, factors leading to effectiveness, theories, and future directions. *Interactive Learning Environments,* 1–29. https://doi.org/ 10.1080/10494820.2023.2202704

Zhang, R., Zou, D., Cheng, G., & Xie, H. (2025). Flow in ChatGPT-based logic learning and its influences on logic and self-efficacy in English argumentative writing. *Computers in Human Behavior*, *162*, 108457. https://doi.org/10.1016/j.chb.2024.108457

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Designing Architecture and Application Interfaces for Educational Robotics Based on Advanced Hardware Components

Bartol BORAS[1*], Antun DROBNJAK[2], Lovro JAKIC[3], Ivan TERZIC[4], Ivica BOTICKI[5]
[1,2,3,4,5]Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
bartol.boras@fer.hr, antun.drobnjak@fer.hr, lovro.jakic@fer.hr, ivan.terzic@fer.hr, ivica.boticki@fer.hr

Abstract: *This paper describes the design and implementation considerations of educational robots based on novel high-performance components. The paper lays out the architecture of a robotics system together with its interfaces towards the specific applicative educational solutions with an overall aim of achieving seamless learning scenarios in educational robotics. The applicative scenarios illustrated as part of this paper provide seamless opportunities for students with various educational backgrounds, levels of knowledge and affinities. The modular architecture of the system, coupled with the high-power hardware components and well-designed application interfaces, open opportunities for further developments in the field of educational robotics coupled with the metaverse technologies.*

Keywords: educational robotics, metaverse, system architecture, application interfaces

1. Introduction

In the past decade, the field of robotics has gained great traction and popularity in education. Robotics in education has a long history, though its widespread adaption began in the 2000s. The technology proved valuable not only for teaching technical subjects but also for language and science education. Educational robots serve diverse roles in learning, ranging from instructional tools to interactive learning partners. At the beginning educational robots used in classroom settings involved primarily physical robots. For instance, the HERO series in the 1980s (López-Belmonte et al., 2021) and LEGO Mindstorms introduced in 1998 which (Christopoulos et al., 2022) were designed for hands-on learning experiences within educational environments. However, the integration of virtual reality (VR) into education, accelerated by the COVID-19 pandemic, has transformed this field. The pandemic necessitated remote learning solutions, promoting educators to adopt VR technologies. This shift driven by immersive technologies and seamless learning principles has expanded educational robotics beyond traditional classrooms, enabling students to interact with virtual robots and simulations, thereby broadening and enriching learning experiences

Educational robotics is constantly evolving in terms of technology and educational approaches, where novel hardware developments open significant opportunities for development or innovative educational solutions. This is especially the case where the integration of novel artificial intelligence technology is concerned. Advanced computer vision technology and language models are now becoming standard building blocks of an educational robotics solution, which opens the possibilities for the design and implementation of learning activities which are more interactive and based on novel pedagogies.

In this paper we propose the architecture of an educational robotics system based on novel technological advancements and propose an architecture based on ROS2 components and interfaces. Such a modular design allows for a unified approach towards both physical and virtual educational robots, where the applications and services are designed as system-type agnostic and serve to drive both virtual and physical educational robotics solutions.

The paper is organized as follows: after the introductory chapter, an account of theoretical approaches underpinning educational robotics is given. This is followed by the description of the system architecture with both its internal structure and external application-oriented interfaces. The paper concludes with an overview of future developments.

2. Theoretical Background: Educational Robotics in the Metaverse

Educational robotics has witnessed significant advancement in the past decade. The most commonly used robots fall into three main categories: wheeled robots, robotics kits, and social robots. Wheeled robots use wheels for navigation and come in several forms: two- or four-wheeled robots (Krūmiņš et al., 2024), car-like robots, line-following robots, and omnidirectional wheeled robots (Escobar et al., 2020). Robotics kits serve as educational tools for teaching programming, electronics, and technology through hands-on construction. These kits include controllers, motors, sensors, and building components that allow students to create various robots—from line-following and wheeled robots (Joventino et al., 2023) to systems using physical blocks, tiles, and tangible objects (Hsu et al., 2024), as well as modular designs (Lopez-Rodriguez & Cuesta, 2020). Social robots, designed for human and robot interaction, serve multiple roles in education: music companions, pet robots, tutors, guides, and language learning aids (Hakim et al., 2024; Khairy et al., 2021; Sui et al., 2020). Additional categories include robotic arms (Da Silva et al., 2023), toy robots (Rodríguez Corral et al., 2019), humanoid robots (Tutul et al., 2024), service robots (Mahmud et al., 2021), and autonomous robots (Naya-Varela et al., 2023), among others. Augmented and virtual reality offers a more realistic and immersive experience. The major subjects and educational activities involving educational robotics are robotics, STEM education, and engineering. Robotics education takes place in both formal and informal settings (Dahal et al., 2023), encompassing basic concepts, sensors, remote labs, competitions, and IoT applications. The field has grown to include home automation, simulation, and autonomous systems (Rukangu et al., 2021).

Of special importance are the sim-to-real approaches as virtual and simulated environments encourage students to work harder in real settings by equalizing their class level (Joventino et al., 2023). Due to hardware limitations, virtual laboratories offer an effective alternative. Simulations help equalize student skills early on, while combining virtual and physical robots reduces cognitive load (Zhong et al., 2023). These technologies provide new opportunities for immersive and personalized learning, offering advanced context awareness and seamless transitions between real and virtual worlds.

These technological advances open opportunities for seamless learning and the continuous flow of learning experiences across different environments, devices, and time. It allows learners to access personalized resources anytime and anywhere, promoting autonomy and engagement (Kerawalla et al., 2007). The educational metaverse takes this concept further by creating immersive digital environments where learners can interact through virtual identities, augmented reality, and wearable devices. Combining elements of lifelogging, mirror worlds and virtual words (Samala et al., 2023), the metaverse enables real-time collaboration without the constraints of physical location (Zhang et al., 2022). This aligns with seamless learning principles by enabling continuous learning across different contexts (Sharples et al., 2012) and devices (Wong, 2012).

3. Systems Design and Architecture

To facilitate easier development and deployment of educational robotics applications and to facilitate the sim-to-real approach to educational robotics, the modular robotics system based on ROS2 components is proposed. Both physical and virtual robots utilize ROS2 as a middleware to improve interoperability and share the same underlying architecture, as illustrated in Figure 1. The communication system is designed as a ROS2 solution and is structured around a distributed network of nodes. A node functions as a fundamental computational unit, representing a single process responsible for executing specialized tasks such as sensor data processing, motion control, or AI inference (*ROS 2 Documentation: Nodes*, n.d.). Multiple nodes operate concurrently and typically communicate by exchanging messages on topics through a publish-subscribe mechanism (*ROS 2 Documentation: Topics,* n.d.).

While certain architectural components, such as sensors, actuators, and layers that directly interact with them, are inherently specific to either virtual or physical platforms, the core logic, including the recognition and planning layers, remains fully transferable between the two platforms. This transferability is achieved by standardizing the interfaces between transferable and non-transferable components using ROS2 topics.

To support varying levels of abstraction, nodes in the recognition and planning layers are implemented as lifecycle nodes. This design allows multiple nodes to run within these layers while only activating those currently needed. For instance, simpler interactions, such as natural language processing for robot movement, could involve a set of nodes dedicated to interpreting commands in natural language. Conversely, more complex interactions may involve nodes that control the robot by receiving velocity values for each wheel, representing a more advanced form of control (Jakic et al., n.d.).

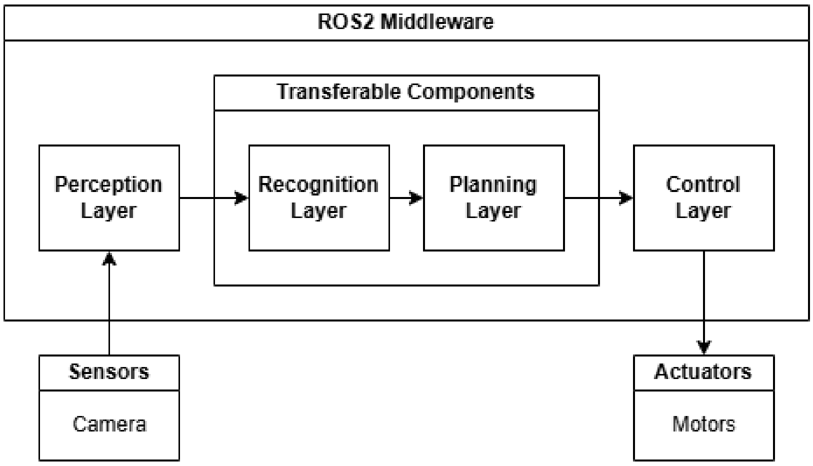


Figure 1. System architecture of virtual and physical robots

To facilitate easier programmatic robot control for the educational robotics programming applications and integration with ROS2, a web server component was designed. The server functions as an application programming interface (API) receiving HTTP or HTTPS requests to the designed routes and forwards the call to functions defined within ROS2 nodes. The nodes map these parameters into commands that are sent to the robot initiating direct robot control. In addition to its API functionality, the web server also serves standard web content.

The proposed architecture facilitates the sim-to-real approach by defining a simple workflow for processing HTTPS requests that are then mapped to low-level robot commands targeting either virtual or physical robots. The commands sent from web-applications are sent as a JSON payload to the central web service, which then forwards them to dedicated functions for processing the robot movement and other functions. In such an approach, high-level commands are mapped to specific actions, such as scaling speed and setting the direction to adjusting joint velocities, which determines robot wheel movement. Finally, these parameters are packaged into state messages, such as the message that holds data to describe the state of a set of torque-controlled joints, and published through the ROS2 publisher to the subscribed physical and virtual robots subscribes, finally resulting in the desired robot operation.

4. Applications for Learning Educational Robotics

Building on the proposed architecture, applicative interfaces for controlling both physical and virtual robots is proposed. This is demonstrated via the use of three web applications that enable students to control virtual or physical robots remotely from any location. Applications use block-based features for computational thinking educational approaches, a highly interactive controller solution for real-time robot control, and a virtual avatar with features for collaborative conversation. Figure 2 illustrates the system workflow, where the key component is a web service which receives requests from web applications, and forwards the commands to the virtual and physical solutions using the underlying ROS2 architecture described in the previous chapter. During the system use, students are able to observe both the physical and/or virtual robot operation via the continuous livestream in the Nvidia Isaac Sim environment.

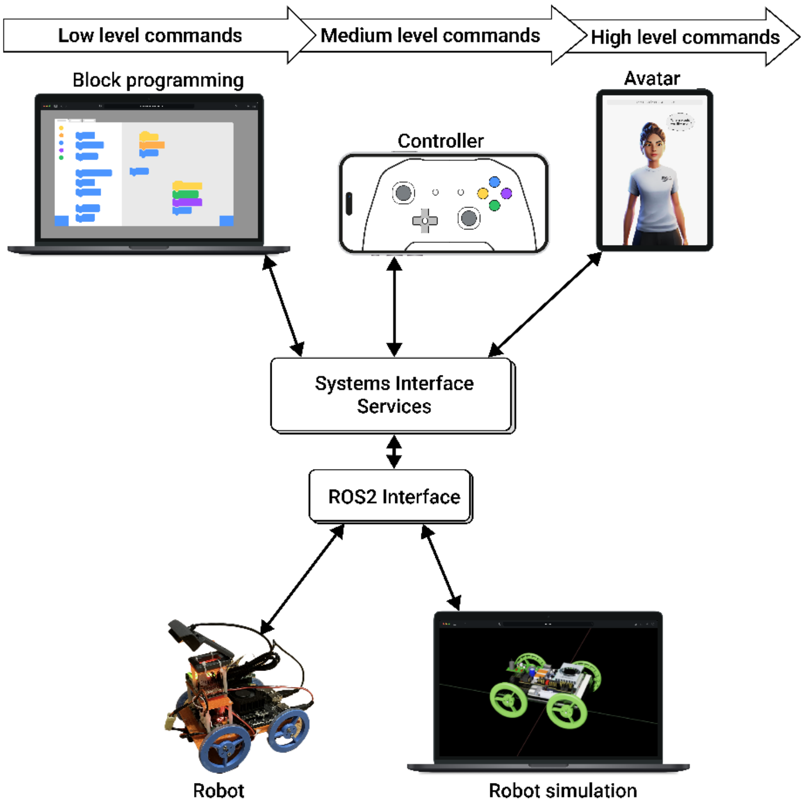


Figure 2. High Level Overview of the System Application Layer

The presented applications offer varying levels of abstraction in robot control, teaching robot operation and kinematics in a manner tailored to the age and expertise of their students. Younger learners or those with less experience utilize interfaces via natural language commands, assisted by an avatar. More experienced young learners use block-based robot control for practicing computational, strategic thinking, and problem-solving skills. More seasoned learners can still utilize block-based interfaces to build solutions which aim at better understanding of the mechanics of robot movement and to learn about the methodologies that ensure seamless integration of the system's various motion components. The approach allows for seamless learning scenarios where students who initially use high-level commands to gain an understanding of the robot's motion, subsequently switching to low-level commands to simulate and potentially enhance those movements. This approach offers students of varying experience, knowledge and age groups an environment to develop fundamental understanding of robotic systems, but also to make innovative advancements within.

5. Conclusions

This paper has explored the design and implementation of an educational robotics system that leverages advanced hardware components and a modular architecture based on ROS2. By integrating both physical and virtual robotics solutions, the proposed framework enhances interoperability and facilitates seamless learning experiences across diverse educational contexts. The seamless transition between physical and virtual robotics environments not only mitigates hardware use limitations but also fosters a more flexible and scalable approach to robotics education. By opening opportunities for real-time collaboration, adaptive learning interfaces, and advanced simulation tools, the proposed solution aligns with modern pedagogical trends including seamless learning scenarios.

Acknowledgements

This research has been funded by the European Union – NextGenerationEU, under the project "MetaRoboLearn - Seamless learning with educational robots through the metaverse" (NPOO.C3.2.R3-I1.04.0194).

References

Christopoulos, A., Coppo, G., Andolina, S., Priore, S. L., Antonelli, D., Salmas, D., Stylios, C., & Laakso, M.-J. (2022). Transformation of Robotics Education in the Era of Covid-19: Challenges and Opportunities. *IFAC-PapersOnLine*, *55*(10), 2908–2913. https://doi.org/10.1016/j.ifacol.2022.10.173

Da Silva, R. C., De Magalhães Netto, J. F., Lopes, A. M. M., De Menezes, M. F., & Menezes, R. A. (2023). ERPLab: Remote Laboratory for Teaching Robotics and Programming. *2023 IEEE Frontiers in Education Conference (FIE)*, 1–7. https://doi.org/10.1109/FIE58773.2023.10343379

Dahal, M., Kresin, L., & Rogers, C. (2023). Introductory Activities for Teaching Robotics with SmartMotors. In R. Balogh, D. Obdržálek, & E. Christoforou (Eds.), *Robotics in Education* (Vol. 747, pp. 229–241). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-38454-7_20

Escobar, L., Moyano, C., Aguirre, G., Guerra, G., Allauca, L., & Loza, D. (2020). Multi-Robot platform with features of Cyber-physical systems for education applications. *2020 IEEE ANDESCON*, 1–6. https://doi.org/10.1109/ANDESCON50619.2020.9272030

Hakim, V. G. A., Yang, S.-H., Wang, J.-H., Lin, H.-H., & Chen, G.-D. (2024). Digital Twins of Pet Robots to Prolong Interdependent Relationships and Effects on Student Learning Performance. *IEEE Transactions on Learning Technologies*, *17*, 1843–1857. https://doi.org/10.1109/TLT.2024.3416209

Hsu, T.-C., Chang, C., & Liang, Y.-S. (2024). Sequential Behavior Analysis of Interdisciplinary Activities in Computational Thinking and EFL Language Learning With Game-Based Learning. *IEEE Transactions on Learning Technologies*, *17*, 270–279. https://doi.org/10.1109/TLT.2023.3311331

Jakic, L., Boticki, I., & Sovic Krzic, A. (n.d.). Bridging Virtual and Physical Robotics: An AI-Driven Educational Platform Using NVIDIA Omniverse Isaac Sim and Jetson Orin Nano (in review). *48th ICT and Electronics Convention*, Opatija.

Joventino, C. F., Silva, R. D. A. A. E., Pereira, J. H. M., Yabarrena, J. M. S. C., & De Oliveira, A. S. (2023). A Sim-to-real Practical Approach to Teach Robotics into K-12: A Case Study of Simulators, Educational and DIY Robotics in Competition-based Learning. *Journal of Intelligent & Robotic Systems*, *107*(1), 14. https://doi.org/10.1007/s10846-022-01790-2

Kerawalla, L., O'Connor, J., Underwood, J., duBoulay, B., Holmberg, J., Luckin, R., Smith, H., & Tunley, H. (2007). Exploring the Potential of the Homework System and Tablet PCs to Support Continuity of Numeracy Practices between Home and Primary School. *Educational Media International*, *44*(4), 289–303. https://doi.org/10.1080/09523980701680904

Khairy, D., Alkhalaf, S., Areed, M. F., Amasha, M. A., & Abougalala, R. A. (2021). Emotional Intelligence Robotics to Motivate Interaction in E-Learning: An Algorithm. *International Journal of Advanced Computer Science and Applications*, *12*(6). https://doi.org/10.14569/IJACSA.2021.0120619

Krūmiņš, D., Schumann, S., Vunder, V., Põlluäär, R., Laht, K., Raudmäe, R., Aabloo, A., & Kruusamäe, K. (2024). Open Remote Web Lab for Learning Robotics and ROS With Physical and Simulated Robots in an Authentic Developer

Environment. *IEEE Transactions on Learning Technologies*, *17*, 1313–1326. https://doi.org/10.1109/TLT.2024.3381858

López-Belmonte, J., Segura-Robles, A., Moreno-Guerrero, A.-J., & Parra-González, M.-E. (2021). Robotics in Education: A Scientific Mapping of the Literature in Web of Science. *Electronics*, *10*(3), 291. https://doi.org/10.3390/electronics10030291

Lopez-Rodriguez, F. M., & Cuesta, F. (2020). An Android and Arduino Based Low-Cost Educational Robot with Applied Intelligent Control and Machine Learning. *Applied Sciences*, *11*(1), 48. https://doi.org/10.3390/app11010048

Mahmud, S., Weber, N., Higgins, P., & Kim, J.-H. (2021). An Intelligent Trash Can Robt for Early Childhood Green Education. *2021 16th International Conference on Computer Science & Education (ICCSE)*, 870–874. https://doi.org/10.1109/ICCSE51940.2021.9569360

Naya-Varela, M., Guerreiro-Santalla, S., Baamonde, T., & Bellas, F. (2023). Robobo SmartCity: An Autonomous Driving Model for Computational Intelligence Learning Through Educational Robotics. *IEEE Transactions on Learning Technologies*, *16*(4), 543–559. https://doi.org/10.1109/TLT.2023.3244604

Rodríguez Corral, J. M., Ruíz-Rube, I., Civit Balcells, A., Mota-Macías, J. M., Morgado-Estévez, A., & Dodero, J. M. (2019). A Study on the Suitability of Visual Languages for Non-Expert Robot Programmers. *IEEE Access*, *7*, 17535–17550. https://doi.org/10.1109/ACCESS.2019.2895913

ROS 2 Documentation: Nodes. (n.d.). [Computer software]. https://docs.ros.org/en/humble/Concepts/Basic/About-Nodes.html

ROS 2 Documentation: Topics. (n.d.). [Computer software]. https://docs.ros.org/en/humble/Concepts/Basic/About-Topics.html

Rukangu, A., Tuttle, A., & Johnsen, K. (2021). Virtual Reality for Remote Controlled Robotics in Engineering Education. *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 751–752. https://doi.org/10.1109/VRW52623.2021.00258

Samala, A. D., Usmeldi, Taali, Ambiyar, Bojic, L., Indarta, Y., Tsoy, D., Denden, M., Tas, N., & Dewi, I. P. (2023). Metaverse Technologies in Education: A Systematic Literature Review Using PRISMA. *International Journal of Emerging Technologies in Learning (iJET)*, *18*(05), 231–252. https://doi.org/10.3991/ijet.v18i05.35501

Sharples, M., McAndrew, P., Weller, M., Ferguson, R., Hirst, T., Mor, Y., Gaved, M., & Whitelock, D. (2012, December). *Innovating Pedagogy 2012: Open University Innovation Report 1*. The Open University. https://oro.open.ac.uk/93816/

Sui, L., Su, Y., Yi, Y., Li, Z., & Zhu, J. (2020). Intelligent Drumming Robot for Human Interaction. *2020 International Symposium on Autonomous Systems (ISAS)*, 168–173. https://doi.org/10.1109/ISAS49493.2020.9378858

Tutul, R., Buchem, I., Jakob, A., & Pinkwart, N. (2024). Enhancing Learner Motivation, Engagement, and Enjoyment Through Sound-Recognizing Humanoid Robots in Quiz-Based Educational Games. In C. Biele, J. Kacprzyk, W. Kopeć, J. Możaryn, J. W. Owsiński, A. Romanowski, & M. Sikorski (Eds.), *Digital Interaction and Machine Intelligence* (Vol. 1076, pp. 123–132). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-66594-3_13

Wong, L. (2012). A learner-centric view of mobile seamless learning. *British Journal of Educational Technology*, *43*(1). https://doi.org/10.1111/j.1467-8535.2011.01245.x

Zhang, X., Chen, Y., Hu, L., & Wang, Y. (2022). The metaverse in education: Definition, framework, features, potential applications, challenges, and future research topics. *Frontiers in Psychology*, *13*, 1016300. https://doi.org/10.3389/fpsyg.2022.1016300

Zhong, B., Zheng, J., & Zhan, Z. (2023). An exploration of combining virtual and physical robots in robotics education. *Interactive Learning Environments*, *31*(1), 370–382. https://doi.org/10.1080/10494820.2020.1786409

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Exploring Self-Regulated Learning Through an English Speaking Task in an Immersive Virtual Reality Environment: A Sequential and Cluster Analysis

Lei TAO[1], Yanjie SONG[2*], Jiachen FU[3], Kaiyi WU[4]
[1,2,3,4] The Education University of Hong Kong, Hong Kong SAR
s1147997@s.eduhk.hk, ysong@eduhk.hk, fjiachen@eduhk.hk, kwu@eduhk.hk

Abstract: *Immersive virtual reality (iVR) has shown promise for language education, particularly in developing speaking skills. However, limited research has investigated how learners engage in self-regulated learning (SRL) behaviours during immersive speaking practices. This study examined SRL behavioural sequences of 11 postgraduate students completing an English-speaking task in LearningverseVR, an immersive platform featuring interactions with AI-powered digital humans. Behavioural data analysis revealed significant variation in learners' SRL sequences, particularly regarding their length, diversity, and entropy. A subsequent cluster analysis identified three distinct learner profiles: Cluster 0, Cluster 1 and Cluster 2. Cluster 2 learners exhibited the longest and most evenly distributed SRL sequences, frequently engaging across forethought, performance, and reflection strategies; Cluster 0 learners demonstrated moderate-length sequences primarily characterised by repeated help-seeking and resource access behaviours, with limited engagement in planning and reflection; Cluster 1 learners displayed the shortest, least diverse sequences, focusing narrowly on task presentation, task completion, and attribution. These profiles highlight critical differences in how learners sequence and apply SRL strategies. The findings provide empirical insights for designing adaptive scaffolds aligned to learners' observed regulatory patterns in iVR environments.*

Keywords: immersive virtual reality, English speaking, self-regulated learning, sequential analysis, cluster analysis

1. Introduction

Immersive virtual reality (iVR) is increasingly recognized as a powerful tool in language education, especially for developing speaking skills. By simulating real-life communicative scenarios, iVR provides learners with high levels of immersion and interactivity, which are known to enhance language input, output, and oral fluency (Kaplan-Rakowski & Gruber, 2021; Slater & Sanchez-Vives, 2016). Compared to traditional environments, iVR can improve learner motivation, engagement, and long-term retention (Lee et al., 2024; Makransky & Lilleholt, 2018; Merchant et al., 2014).

However, iVR's benefits come with cognitive demands. Without teacher supervision, learners may feel overwhelmed or distracted by the sensory complexity of VR, which calls for high degrees of self-regulated learning (SRL)—learners' ability to plan, monitor, and reflect on their own performance (Wang et al., 2022). According to Zimmerman's model, SRL involves three cyclical phases: forethought, performance, and reflection (Zimmerman, 2002).

While SRL has been widely studied in traditional and online settings through self-reports and learning journals, research in VR-based environments is still emerging. Recent studies focus mainly on affective or motivational aspects (Chen et al., 2021; Gruber & Kaplan-Rakowski, 2020; Kaplan-Rakowski & Wojdynski, 2018; Yudintseva, 2024), or on how system-level feedback and scaffolding can support SRL (Shen et al., 2025). However, these studies rarely examine how learners engage in SRL step-by-step during VR tasks, by directly observing their behaviours as they happen. Behavioural data such as speech, gaze, and interaction logs offer a promising alternative to self-reports. These traces can reveal how learners regulate task engagement in real time and help identify learner profiles with different regulatory patterns (Sobocinski et al., 2024). However, little research has examined the sequential nature of SRL behaviour in VR-based language tasks—how learners move between phases and strategies during a task.

To address this gap, the current study investigates how university students regulate their learning while completing an English-speaking task in LearningverseVR, an iVR platform developed by our research group. Delivered through a headset-based system, LearningverseVR allows learners to interact with AI-powered digital humans in authentic 3D learning scenarios. By analysing learners' behavioural sequences and comparing across learners, this study aims to reveal (1) the types of SRL behaviours exhibited and (2) distinct learner profiles based on behavior patterns. Specifically, it addresses the following research questions:

- RQ1: What self-regulated learning behaviours do students exhibit during English speaking tasks in LearningverseVR?
- RQ2: What learner types can be identified, and how do these types differ across SRL phases?

2. Methodology

2.1. Participants and context

A total of 11 postgraduate students (9 female, 2 male; aged 22–27) from education- and language-related disciplines at a Hong Kong university participated in this study. All had comparable oral proficiency levels (intermediate to upper-intermediate). Each participant completed a simulated English-speaking task individually in LearningverseVR, an immersive, headset-based environment. The task scenario was set in a virtual university library, where the learner had to report a lost book and request assistance. Learners interacted with a generative AI-powered digital human trained to respond naturally and adaptively (see Figure 1).

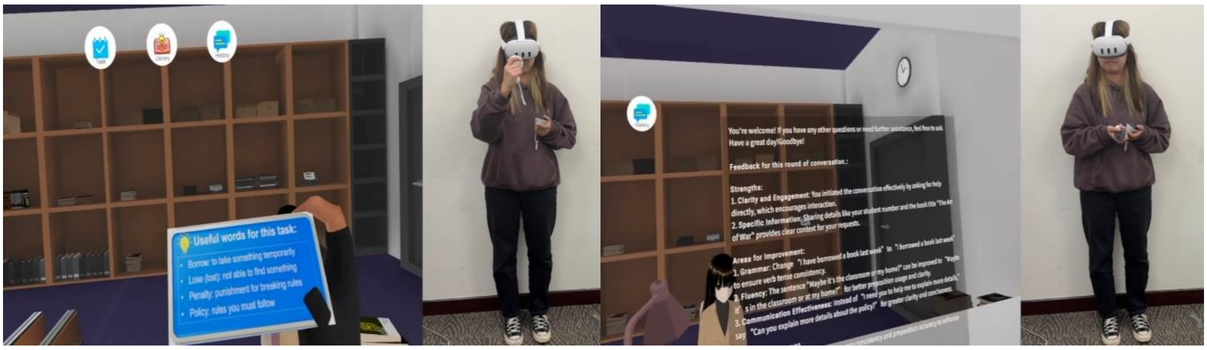


Figure 1. Learner interacting with LearningverseVR

2.2. Data collection and analysis

Data collection included logged dialogues between the learner and the digital human, and behavioural data, such as the use of vocabulary support cards for the speaking task in the iVR environment.. Content analysis (Stemler, 2015) was adopted. The entire interaction was manually segmented and coded using a predefined SRL behaviour coding scheme, covering three SRL phases: (1) Forethought (e.g., Goal setting-G, Strategic planning-SP, Task analysis-TA, Self efficacy-SE); (2) Performance (e.g., Presenting the problem-P, Requesting help-R, Accessing resources-A, Monitoring/Modifying-M, Errors-E, Finish-F); and (3) Reflection (e.g., Self evaluation-EV, Attribution-AT, Affect/Satisfaction-AF, Strategic revision-SR).

3. Results

3.1. SRL behavioural sequences

Each participant's dialogue with the digital human was coded into a sequence of self-regulated learning (SRL) behaviours. As shown in figure 2, the sequences vary in length, complexity, and phase transitions. Some learners demonstrated structured, phase-aligned patterns (e.g., [G, SP, P, R, A, F, EV, AT]), while others skipped forethought or reflection steps entirely.

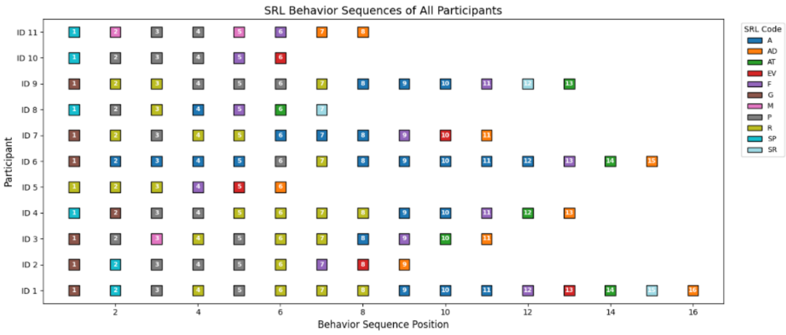


Figure 2. SRL behaviour sequences of all participants

(Each row represents one participant, each square a coded behaviour. Color-coded by SRL code.)

Table 1 summarizes the overall trends in learners' SRL behaviours sequences, including sequence length, behavioural diversity, and entropy. Learners exhibited varying levels of regulatory complexity: while some showed long and diverse sequences (e.g., ID1 with 16 actions and 10 unique codes), others demonstrated shorter, more repetitive patterns (e.g., ID5 with only 6 actions and 4 types). On average, behavioural entropy remained moderate (M = 2.47), but outliers at both ends suggest different degrees of strategic flexibility. High-entropy learners tended to distribute actions more evenly across SRL phases, whereas lower-entropy sequences reflected reliance on a narrow set of behaviours, such as repeated help-seeking or support access. These individual differences suggest distinct learner profiles in their approach to the task. In the following sections, we examine how these patterns manifest in behavioural transitions (Section 3.2) and how learners can be grouped into regulatory types using cluster analysis (Section 3.3).

Table 1. SRL behaviour metrics

Metric	Mean	Min	Max	SD
Sequence length	10.45	6.00	16.00	3.53
Behavioural diversity	6.73	4.00	10.00	1.79
Entropy	2.47	1.79	3.08	0.44

3.2. SRL behaviour transitions

To examine how learners navigated between SRL strategies, we computed a transition matrix capturing the frequency of behaviour-to-behaviour shifts across all participants (Figure 3).

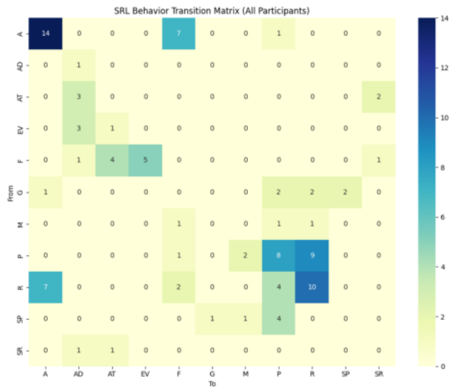


Figure 3. SRL behaviour transition matrix

The most dominant transitions occurred within the Performance phase, with highly frequent patterns such as repeated access to support (A → A, n = 14), repeated help-seeking (R → R, n = 10), and the sequence of problem presentation to help-seeking (P → R, n = 8). These indicate that many learners cycled between requesting support and repeatedly re-engaging with the same action rather than progressing toward resolution. Transitions from Performance to Reflection were observed but less frequent—such as F → EV (n = 4) and F → AT (n = 5)—suggesting that while some learners engaged in evaluative thinking post-task, this was not consistent across participants. Moreover, a few learners progressed across all SRL phases (e.g., G → SP → P, or EV → AT → SR), but such full-phase transitions were rare. Overall, the data suggest that many learners relied on reactive, cyclic behaviours centred on help-seeking and support access. This highlights the need to scaffold not only strategy use, but also phase progression and regulatory closure.

3.3. Learner profiles based on SRL behaviour

To identify distinct patterns in students' self-regulated learning (SRL), we performed a K-means cluster analysis (k = 3) using three behaviour sequence features: sequence length, behavioural diversity, and entropy. These variables reflect how long learners engaged in the task, how many types of SRL strategies they used, and how evenly their actions were distributed across strategy categories. The resulting clusters—visualised in Figure 4—capture meaningful differences in regulatory complexity and engagement styles.

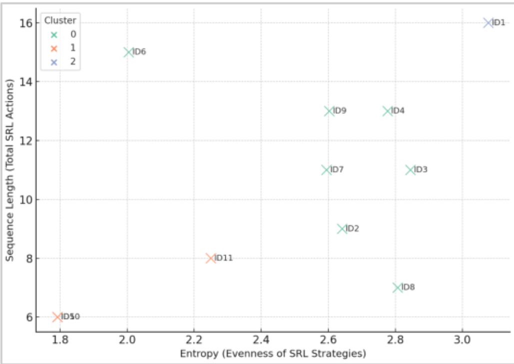


Figure 4. Learner clusters based on SRL behaviour metrics

Cluster 2 learners exhibited the most complex SRL profiles. They produced the longest (M = 16.0) and most evenly distributed sequences (entropy = 3.08), with high-frequency use of help-seeking (R = 4.00) and resource access (A = 3.00). They also engaged in forethought strategies like goal-setting (G = 1.00), strategic planning (SP = 1.00), and post-task reflection behaviors including evaluation (EV = 1.00) and attribution (AT = 1.00). This group represents strategic and highly engaged regulators, who distributed their behaviours across all SRL phases. Cluster 0 included learners with moderately long sequences (average length ≈ 11–13) and a more reactive behavioural profile. While they also relied on performance-phase strategies such as repeated help-seeking (R = 2.29) and support access (A = 2.71), they showed limited but present forethought (e.g., G = 0.86, SP = 0.43) and reflection (e.g., AT = 0.71, EV = 0.57). Their behaviour patterns indicate a performance-centred approach with occasional regulation across phases—more reactive than proactive. Cluster 1 learners demonstrated the shortest and least diverse sequences. With minimal use of planning (G = 0.00), support (A = 0.00), or metacognitive strategies, their behaviours concentrated on task presentation (P = 1.67), finalization (F = 1.00), and attribution (AD = 1.00). This group reflects a low-engagement, task-oriented style, suggesting minimal SRL involvement beyond basic task completion.

Figure 5 provides a radar visualization of SRL strategy usage across the three clusters. Cluster 2 displays the widest coverage, with strong presence in both forethought and reflection phases. Cluster 0 is centred on performance, with noticeable but lower engagement in the other phases. In contrast, Cluster 1 is tightly compressed, with very limited strategy diversity. These results illustrate that learners differ not only in how much they regulate, but also in which strategies they prioritize and how completely they enact the SRL cycle.

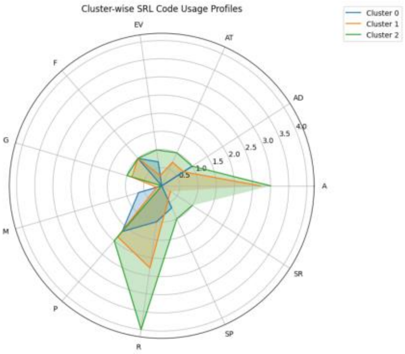


Figure 5. Cluster-wise SRL code usage profiles

4. Discussion and conclusion

This study investigated how learners regulate their behaviours during an immersive English-speaking task using LearningverseVR. The behavioural sequence data revealed notable variation in strategy use and ordering. While SRL behaviours from all three phases—forethought, performance, and reflection—were present across the dataset, their distribution and sequencing differed significantly between learners. Some sequences followed a structured pattern from planning to task execution and post-task reflection, while others remained narrowly focused on a subset of performance-phase strategies.

These findings contribute to ongoing discussions in SRL and immersive learning research. While previous studies have emphasized learners' affective responses and perceived engagement in iVR (Gruber & Kaplan-Rakowski, 2020; Kaplan-Rakowski & Wojdynski, 2018), our data-driven behavioural analysis extends this line of work by uncovering actual regulatory patterns. Consistent with concerns about increased cognitive demands in immersive environments (Wang et al., 2022), we found that many learners either skipped planning entirely or relied heavily on repeated help-seeking strategies without progression. However, as suggested by Sobocinski et al. (2024), behavioural sequences can offer a more granular understanding of SRL enactment in situ—a view our clustering results support.

Cluster analysis further clarified these differences. Cluster 2 learners engaged in a broad range of SRL strategies, including planning (G, SP), resource use (A), and reflection (EV, AT), suggesting a complete SRL cycle. Cluster 0 learners used similar strategy types but with lower frequency and less phase coverage, often looping within performance-phase behaviours before shifting to reflection. Cluster 1 learners demonstrated minimal SRL behaviours, with short sequences focused on problem presentation (P), task completion (F), and attribution (AD). These profiles illustrate that learner differences are not only in what strategies are used, but also in how those strategies are sequenced and sustained. These insights inform concrete design directions. For instance, learners with short, performance-focused sequences could benefit from system-triggered forethought prompts before task onset, especially when no planning behaviours are detected. For learners who show looping within R → A → R patterns, adaptive scaffolds that suggest closure (e.g., "Would you like to conclude your request?") may support regulatory progression. Post-task reflection can also be triggered only when F or M appears, rather than using static prompts. These timing-aware scaffolds respond to actual SRL behaviour patterns, as opposed to generic support.

Methodologically, the study contributes to SRL research by demonstrating how behavioural coding and sequence-based clustering can reveal regulatory profiles beyond what self-report captures. Still, the study is limited by its small, homogenous sample (N = 11), which limits generalizability. Additionally, the coding relied on observable, logged behaviours and could not capture internal planning or reflection that was not verbalized. Future work should integrate multimodal data (e.g., gaze, biosignals) to better detect implicit regulation.

Acknowledgements

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

This research was supported by the UGC's Fund for Innovative Technology-in-Education (FITE) (Project Codes: 05404 and 05408). We would like to thank the funding support, and all team members, including those contributing to the development of LearningverseVR.

References

Chen, C.-H., Hung, H.-T., & Yeh, H.-C. (2021). Virtual reality in problem-based learning contexts: Effects on the problem-solving performance, vocabulary acquisition and motivation of English language learners. *Journal of Computer Assisted Learning*, *37*(3), 851-860. https://doi.org/https://doi.org/10.1111/jcal.12528

Gruber, A., & Kaplan-Rakowski, R. (2020). User experience of public speaking practice in virtual reality. In *Cognitive and affective perspectives on immersive technology in education* (pp. 235-249). IGI Global.

Kaplan-Rakowski, R., & Gruber, A. (2021). One-On-One Foreign Language Speaking Practice in High-Immersion Virtual Reality. In Y.-J. Lan & S. Grant (Eds.), *Contextual Language Learning: Real Language Learning on the Continuum from Virtuality to Reality* (pp. 187-202). Springer Singapore.

Kaplan-Rakowski, R., & Wojdynski, T. (2018). Students' attitudes toward high-immersion. *Future-proof CALL: language learning as exploration and encounters–short papers from EUROCALL 2018*, 124.

Lee, S.-M., Yang, Z., & Wu, J. G. (2024). Live, play, and learn: Language learner engagement in the immersive VR environment. *Education and Information Technologies*, *29*(9), 10529-10550.

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, *66*(5), 1141-1164. https://doi.org/10.1007/s11423-018-9581-2

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education*, *70*, 29-40. https://doi.org/https://doi.org/10.1016/j.compedu.2013.07.033

Shen, B., Fanrong, W., Yi-Chao, J. M., Di, Z., & and Jong, M. S.-Y. (2025). Harnessing spherical video-based virtual reality to enhance EFL learners' writing performance and self-regulated learning strategy use. *Computer Assisted Language Learning*, 1-46. https://doi.org/10.1080/09588221.2025.2482148

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing Our Lives with Immersive Virtual Reality [Opinion]. *Frontiers in Robotics and AI*, *Volume 3 - 2016*. https://doi.org/10.3389/frobt.2016.00074

Sobocinski, M., Dever, D., Wiedbusch, M., Mubarak, F., Azevedo, R., & Järvelä, S. (2024). Capturing self-regulated learning processes in virtual reality: Causal sequencing of multimodal data. *British Journal of Educational Technology*, *55*(4), 1486-1506. https://doi.org/https://doi.org/10.1111/bjet.13393

Stemler, S. E. (2015). Content analysis. *Emerging trends in the social and behavioral sciences: An Interdisciplinary, Searchable, and Linkable Resource*, *1*, 1-14.

Wang, C.-H., Salisbury-Glennon, J. D., Dai, Y., Lee, S., & Dong, J. (2022). Empowering College Students to Decrease Digital Distraction through the Use of Self-Regulated Learning Strategies. *Contemporary Educational Technology*, *14*(4).

Yudintseva, A. (2024). An exploration of low- and high-immersive virtual reality modalities for willingness to communicate in English as a second language. *Computers & Education: X Reality*, *5*, 100076. https://doi.org/https://doi.org/10.1016/j.cexr.2024.100076

Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice*, *41*(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Exploring the Therapeutic Leisure Applications of Virtual Reality for Long-Term Hospitalized Patients with Chronic Mental Illness

Li-Yu CHEN[1], Chia-hui HUNG[2*]
[1,2]Chung Shan Medical University, Taiwan
chenliyu0983505705@gmail.com, chhung@csmu.edu.tw

Abstract: *This study investigates the effectiveness of Virtual Reality (VR) in enhancing leisure participation motivation and improving the physical and mental health of long-term hospitalized patients with chronic mental illness. A total of 25 patients were recruited and randomly assigned to either the experimental group (VR-enhanced leisure activities) or the control group (standard leisure activities) for an eight-week intervention. Assessment indicators included leisure motivation (Interest Scale, Volition Scale), cognitive and behavioral performance (Canadian Occupational Performance Measure, COPM), and emotional and problem-solving strategies (Positive and Negative Syndrome Scale, PANSS; Health Toolbox). Results showed significant improvements in the experimental group, with increased leisure motivation ($p < .000$) and COPM satisfaction ($p < .000$), enhanced cognitive and behavioral abilities ($p = 0.004$, Cohen's $d = 1.11$), reduced negative symptoms ($p = 0.002$), and strengthened stress coping strategies. In conclusion, VR-based leisure activities effectively enhance motivation for participation, promote physical and mental well-being, and support recovery.*

Keywords: Virtual Reality, Leisure, Chronic Mental Illness, Long-Term Psychological Care, Occupational Therapy

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

探究虛擬實境於長期住院慢精神疾病患者之治療性休閒活動應用成效

陳俐瑜 [1]，洪佳慧 [2*]
[1,2] 中山醫學大學職能治療學系，台灣
chenliyu0983505705@gmail.com，chhung@csmu.edu.tw

【摘要】本研究探討虛擬實境於長期住院慢性精神疾病患者休閒參與動機、生理與心理健康提升成效。招募 25 名患者，隨機分為實驗組 (VR 增強休閒活動) 與對照組 (標準休閒活動)，進行八週介入。評估指標含休閒動機 (興趣量表、意志量表)、認知與行為表現 (加拿大職能表現測量 COPM)、情緒與問題解決策略 (正性與負性症狀量表 PANSS、健康工具箱)。結果顯示，實驗組休閒動機 ($p < .000$)、COPM 滿意度 ($p < .000$) 均顯著提升，認知與行為能力改善 ($p = 0.004$, Cohen's $d=1.11$)，負性症狀減少 ($p = 0.002$)，壓力應對策略增強。綜上，VR 休閒活動有效提升休閒參與動機，促進生理、心理健康，幫助復原。

【關鍵字】 虛擬實境、休閒、慢性化精神疾病患者、長期心理照顧、職能治療

1. 前言

精神病人的照顧需求一直是備受社會關注之議題，臺灣慢性精神疾病患者人數逐年上升，尤其 50 歲以上患者佔總患者人口數 59.3% (臺灣衛生福利部統計處，2020)，且患者普遍住院時間過長 (王玲玲、林秋芬，2013)，反映臺灣精神疾病患者「高齡化與超長期住院」問題日益嚴重。住院精神病人之休閒活動受限於實體醫院環境，有諸多限制，尤其在疫情後，許多醫院實施分倉分流措施，加劇病人休閒需求未被滿足的情況，此現象不僅構成職能不平衡 (Occupational imbalance)，甚至已達到職能剝奪 (Occupational deprivation) 的程度，導致病人休閒參與動機日益削弱，最終對其生理與心理健康造成不容忽視的累積性損害。

然而，休閒對個體生理、心理健康至關重要，有助於疾病預防與控制，且休閒活動於精神疾病治療上具有重要價值，患者參與之休閒活動多樣性越高，其心理健康水平越佳 (Dupuis & Smale, 1995)。因此，如何提升長期住院病人之休閒參與動機，滿足其休閒需求，促進其身心健康，已成為當前長期住院精神病人照護中亟待解決的重要議題。

近年來，虛擬實境 (Virtual Reality) 已成為醫學領域的重要臨床工具，被廣泛應用於焦慮症、特定恐懼症、創傷後壓力症候群，以及腦損傷、中風訓練、幻肢疼痛等心理、生理領域。然而，VR 鮮少應用於住院精神病人的休閒治療活動，因此本研究旨在運用 VR 突破實體醫療場域限制，以虛實境增設計增強休閒治療活動，提供住院病人豐富且多樣的休閒體驗，盼能藉此滿足患者需求，提升其休閒參與動機，促進生理、心理健康，並蒐集虛擬實境治療活動執行過程中職能治療臨床治療師及患者之質性回饋，以提供臨床應用與未來相關研究之建議。

2. 文獻回顧

休閒為個體重要職能，從事休閒活動能紓解壓力、調適情緒，達到自我實現 (顧兆台、曾于誌，2012)，長期參與休閒有助於提升身心健康，反之則可能導致身心健康問題 (Tinsley, 1986) 。此外，休閒亦在疾病治療中扮演重要角色，能延緩疾病惡化、增強認知能力及自我價值感。然而，休閒的本質不在於活動，而在於其對於個人而言的主觀意義及感受，「休閒感受」由個體從事休閒活動時感受到的自由程度決定 (Seppo Roy, 2023)，形成「休閒態度」並影響「休閒參與動機」。休閒態度涵蓋認知、行為、情感三面向，是決定個體是否參與休閒活動最主要的因素之一 (Ragheb & Beard, 1982)。休閒參與動機由內在動機驅動，影響個體休閒參與程度 (Neulinger, 1974) 。綜上所述，個體從事休閒活動過程中的休閒感受會影響其休閒參與動機與休閒態度。因此，本研究以個體參與休閒過程中的「情緒感受」、「參與動機」、「認知行為表現」作為評估休閒成效之指標。

虛擬實境 (Virtual Reality, VR) 透過電腦模擬真實情境，製造仿真世界，以想像性 (Imagination)、互動性 (Interaction) 及沉浸性 (Immersion)為核心特性 (胡明強等，2021)，讓使用者透過視覺、聽覺、觸覺等感官獲得身歷其境的體驗。其視覺體驗由頭戴式顯示器隔絕外界，聽覺則以耳機模擬虛擬空間聲音 (李潤容，2017)，具備客製化、穩定一致、可重複及無錯學習等優勢 (Kim, 2005)。近年來，虛擬實境被應用於娛樂、教育與醫療領域，尤其在精神醫學方面，研究證明透過虛擬實境提供感官刺激，進行認知、社會技能等訓練，有效提升患者認知能力、改善情緒問題，幫助康復並提高生活品質，未來更可能成為治療特定精神症狀的創新臨床工具 (Park et al., 2019)。然而，虛擬實境亦存在限制，如傳輸線侷限行動、螢幕與人眼適應衝突引發眼睛疲勞、頭痛 (Kramida, 2015)，以及動暈症與後遺症等副作用 (Kennedy et al., 2003)。在臨床應用 VR 技術時，須仔細篩選參與者、規劃安全介入方案、選擇適當場域並密切觀察不適反應，以確保使用者健康與治療成效。因此，本研究欲於執行研究過程中嘗試調整虛擬實境介入策略，盼能提供未來臨床應用與相關研究之建議。

綜上所述，本研究以虛擬實境作為研究工具，設計「虛擬實境增強休閒治療活動」，介入長期住院精神病人之休閒領域，並以休閒動機、休閒態度 (個體參與休閒活動中認知與行為表現)、休閒感受 (個體參與休閒活動中情緒與問題解決策略) 為成效評估指標，探究虛擬實境對長期住院慢性化精神疾病患者之治療性休閒活動介入成效，以及對患者休閒參與動機、生理功能、心理健康之影響，期待能為未來精神醫療休閒治療模式之創新發展與臨床應用提供實證參考。

3. 研究方法

3.1 研究設計、對象與環境：

本研究採用「隨機實驗對照組準實驗設計」並以成效型研究方法輔助，實驗組參與虛擬實境增強休閒治療活動，對照組則參與標準休閒活動。「虛擬實境增強休閒治療活動」設計為透過虛擬實境的核心要素：想像性 (Imagination)、互動性 (Interaction) 及沉浸性 (Immersion) 影響參與者休閒感受，促進其休閒參與動機，改變認知行為，增加情緒問題解決策略，最終形成積極的休閒態度。「標準休閒活動」為患者於醫院內日常參與之傳統休閒活動，如：繪畫、讀報、韻律等。量性成效評估為比較實驗組及對照組前測、後測各項量表評估結果，質性成效評估為收集受試者及臨床職能治療師之半結構試訪談質性回饋。

本研究受試者納入標準為：住院達 6 個月以上，精神症狀穩定，能夠口語表達、自行移動，無疾病干擾之長期住院慢性化精神疾病患者。排除標準為：有視覺、聽覺障礙及其他生理疾病，有癲癇、易暈眩病史者，無法辨別現實與虛擬者。共納入 25 名受試者，分實驗組與對照組 (實驗組 13 名，平均 54 歲；對照組 12 名，平均 58 歲)，並以 G*power 統計分析法估算 「所需有效樣本數」 ，結果顯示，當 effect size=0.7，alpha 小於.05，則受試者總數 24 人即可。本研究通過臺灣國家科學技術委員會及中山醫學大學附設醫院之倫理審查委員會審查合格 (案號：CS1-24114)。

3.2. 評估工具：

(1) 職能治療臨床興趣量表：用以評估受試者休閒經驗、休閒動機，了解其喜好之休閒活動類型及數量。(2) 意志量表 (Volitional Questionnaire)：由職能治療師觀察、評估受試者從事休閒治療活動過程中內在動機之變化，及外在環境如何影響其意志狀況。(3) 加拿大職能表現測量 (Canadian Occupational Performance Measure)：評量個案在自我照顧、生產力及休閒三方面的表現職能勝任能力與滿意度之問題與變化情形。(4) 正性與負性症狀量表 (Positive and Negative Syndrome Scale)：評估受試者精神疾病症狀之嚴重程度，分為活性症狀、負性症狀、一般精神病生理症狀。(5) 健康工具箱 (Wellness Toolbox)：藉由半結構式訪談詢問受試者在

介入前、後壓力情緒因應策略，比較策略數量改變及分析其策略性質。深入了解虛擬實境介入對參與者情緒調節與壓力因應能力的影響，評估其促進心理健康的實際效益。

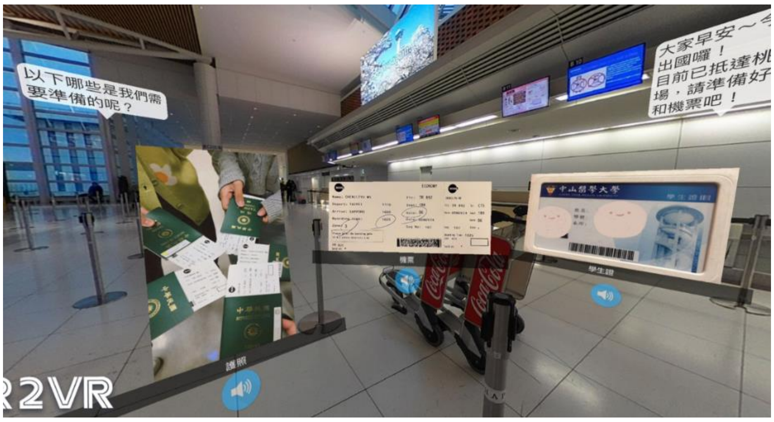
3.3. VR 休閒活動介入：

本研究以虛擬實境增強休閒治療活動介入，為期六週，每週一次，每次 40 分鐘。治療活動流程為暖身 (10 分鐘)，VR 休閒探索活動 (20 分鐘)，分享回饋 (10 分鐘)。每周選定一到二類休閒活動進行探索，活動前段依據每週主題設計不同的暖身活動，如：跳操 (舞蹈類)、揮棒練習賽 (球類) 、吟詩歌唱 (民俗社交活動類) 等。活動中段為 VR 休閒探索活動體驗，參與者配戴 VR 眼鏡進入擬真休閒情境，模擬在醫院外從事此休閒活動的情境，同時訓練其休閒參與技巧，如：空間知覺、認知、表達、問題解決、心理情緒調適技巧等。

本研究選用由阿特發互動科技有限公司所開發的 AR2VR 應用程式，作為建構 VR 休閒活動情境之平台。以 360 度全景相機拍攝現實世界休閒場域影像，將影像匯入 AR2VR 軟體中進行後製。於影片中鑲嵌多種互動元素，包括情境式問題、虛擬圖像、2D 輔助影片及背景音樂等，藉此提升虛擬情境之沉浸感與互動性 (圖一)。最後將完成的 VR 影片匯入雲端，傳送至簡易型 VR 紙盒眼鏡，供受試者配戴體驗 (圖二)。此外，VR 休閒情境依據興趣量表界定之八大休閒種類 (戲劇舞蹈美術類、運動類、視聽娛樂類、民俗社交活動類、球類活動、嗜好類、戶外活動類、益智遊戲類) 加以發展，依據每周欲訓練之休閒參與技巧，編排設計全景畫面，於情境中加入即時問答內容及具體任務，藉由視覺、聽覺、觸覺刺激，促使受試者主動觀察、回應情境問題，並激發自發性挑戰與完成任務之動機。透過上述方式，不僅使受試者能夠實際體驗多樣化休閒活動，亦可同時達到提升休閒參與技能與相關功能訓練之雙重目的。

然而，鑒於虛擬實境技術在實際應用上仍存在一定的使用限制，特別考量本次參與研究之受試者皆為慢性化精神疾病患者，其生理功能相較一般人較為脆弱，對外在感官刺激的耐受度亦較低。為避免使用傳統電子 VR 頭戴式裝置時，因螢幕高亮度、聲音強度過高或沉浸式體驗過於刺激，導致患者出現暈眩、噁心，甚至誘發癲癇等不良副作用，本研究特別選用簡易型 VR 紙盒眼鏡作為介入工具。具體操作方式為：於智慧型手機下載 VR 影片，透過 AR2VR 軟體開啟影片進行體驗，並將手機置入 VR 紙盒眼鏡專用之溝槽中。透過紙盒眼鏡內部設計的特殊鏡片，能夠讓受試者感受到立體 3D 畫面，達到基本的虛擬實境體驗效果，同時大幅降低過度感官刺激之風險，嘗試兼顧體驗品質，同時確保受試者在參與過程中的安全性，從而提高介入研究的可行性與實施效果。

▼圖一、VR 休閒情境畫面



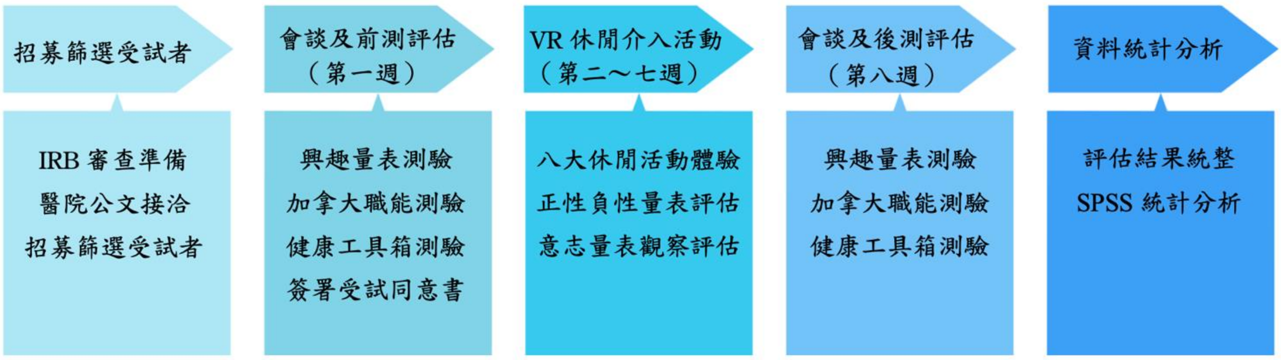
▼圖二、使用簡易 VR 眼鏡

3.4. 研究流程：

本研究經倫理審查委員會審查通過後，依照受試者納入/排除標準篩選出 25 位受試者，進行 8 週實驗。第一週進行各項評估工具前測及受試者訪談。第二至第七週介入期，實驗組及對照組分別參與 VR 增強休閒治療活動及標準休閒活動，由治療師評估活動過程中受試者之參與狀況及精神症狀改變程度。第八週進行後測：各項評估工具前測後測，並以訪談收集

受試者參與 VR 活動質性回饋。

▼圖一、研究流程圖



3.5. 資料分析：

(1) **描述統計 (Descriptive Statistics) 分析**將受試者基本資料以平均數與標準差或百分比進行分析。(2) **質性資料分析**為將意志量表治療師觀察評測結果及健康工具箱受試者自述情緒壓力因應策略使用紮根理論，進行開放編碼、組軸編碼、選擇性編碼，最後形成主題。(3) **推論統計 (Inferential Statistics)** 以**獨立樣本 T 檢定 (t-test, independent samples)** 比較實驗組及控制組之興趣量表、加拿大職能表現測驗 (COPM)、正性與負性症狀量表(PANSS) 之前測、後測結果，可知其休閒動機、休閒滿意度及精神病生理症狀嚴重程度之差異。以**重複量數 T 檢定 (t-test, repeated)** 比較每周治療師評測意志量表評測項目之變動差異。

4. 研究結果與討論

獨立樣本 t 檢定結果顯示，實驗組與對照組在各項評估工具前測分數皆無顯著差異，證實兩組休閒動機、認知與行為能力、休閒職能表現問題及休閒滿意度、精神症狀基線相當。無母數分析結果顯示，實驗組休閒動機於虛擬實境增強休閒治療介入後顯著提升（$F(5, 13) = 16.237, p < .000$），認知與行為能力有所改善（$p = 0.004$, Cohen's d=1.11）、負性精神症狀顯著減少（$t(13) = 3.88, p = 0.002$, Cohen's d = 0.53），健康工具箱評估結果顯示其情緒壓力應對策略增加。加拿大職能表現測驗 (COPM) 結果顯示實驗組休閒滿意度得分顯著提高（$t(13) = 5.182, p < .000$），且在介入後表現出休閒職能行為出現顯著變化（$t(13) = 4.335, p = 0.001$）。結果顯示，虛擬實境增強休閒治療活動能有效提升長期住院慢性精神病患之休閒參與動機、認知行為表現與情緒因應策略。實驗組於介入後對休閒需求更具覺察，並主動表達未來對休閒探索之期待，顯示此介入能促進休閒參與及培養積極休閒態度。

然而，本研究對象為長期住院之慢性精神病患者，因生理及心理功能較為脆弱，部分參與者於虛擬實境介入過程中出現步態不穩、疲乏及暈眩等現象。為確保安全，自第二週起改以「有椅背之滾輪椅」輔助，讓參與者以坐姿體驗 VR，提升穩定性。多數參與者對此調整表示舒適且易於旋轉，少數則偏好站姿之自由度。另有少數參與者於介入後出現頭暈、噁心等 VR 常見副作用，故本研究調整螢幕亮度、改以間歇休息模式進行體驗，並嚴格管理生理功能較弱個案之體驗時間。建議未來針對脆弱族群，除使用簡易紙盒 VR 眼鏡以降低過度沉浸造成不適反應以外仍須嚴格調控聲光刺激強度。此外，本研究亦觀察到年長參與者初次使用 VR 時因不熟悉操作而表現畏縮、緊張，甚至抗拒，但隨著操作熟練，部分個案逐漸展現成就感，並表達希望未來能有更多虛擬實境休閒活動。綜合以上，VR 影片設計應考量參與者年齡、心理調適狀況及挫折耐受度，並於情境問題設計上採分級制，且須謹慎安排影片中物件的位置與距離，避免誤觸或因操作角度困難而影響體驗品質。

▼

表 1 實驗組興趣量表「未來想嘗試」項目前測及後測 T 檢定分析

未來想嘗試	平均值 (Mean)	標準差 (SD)	P-value
前測	52.62	27.21	0.15
後測	72.23	32.54	

▼

表 2 實驗組加拿大職能表現量表「職能滿意度」項目前測及後測 T 檢定分析

職能滿意度	F 值	t 值	P-value
前測	5.668	1.124	.902
後測	1.043	5.182	.000

▼**表 3** 實驗組加拿大職能表現量表「操作能力表現」項目前測及後測成對 T 檢定分析

職能滿意度	自由度	t 值	P-value
前測-後測	12	4.335	.001

▼**表 4** 正性與負性症狀量表前測及後測成對 T 檢定分析

精神症狀	平均值 (Mean)	標準差 (SD)	P-value
正性症狀-前測	17.92	5.001	< .001
正性症狀-後測	15.69	4.11	
負性症狀-前測	20.92	4.82	.0002
負性症狀-後測	18.31	4.07	
一般症狀-前測	40.85	5.58	< .001
一般症狀-後測	37.31	4.91	
整體症狀-前測	79.69	10.14	< .001
整體症狀-後測	71.31	9.17	

5. 結論

本研究結果顯示，虛擬實境 (VR) 治療性休閒活動能有效提升長期住院慢性化精神疾病患者之休閒參與動機，並促進其生理功能、心理健康，幫助復原。實驗組對自身休閒感受度有更深刻的覺察，休閒動機顯著提升，且整體滿意度亦有所提高。此外，認知行為能力改善、負性症狀嚴重程度降低、壓力應對策略增加，證明以 VR 三大特性—想像性、互動性及沉浸性設計之休閒情境可有效促進認知，增進情緒與問題解決策略。儘管部分參與者在使用 VR 後有輕微不適狀況，未來需更進一步調適解決以減少不適。但總體而言，虛擬實境結合休閒與醫學兩領域的應用無疑是創新且有顯著成效的介入方式，盼再經過更多探討證實後可為長期住院患者慢性化精神疾病患者之生活品質及福祉帶來更深遠的改善與突破。

參考文獻

王玲玲、林秋芬（2013）。住院病人對病人權利的重視與獲得。**北市醫學雜誌，10** (3)，244–253。

李潤容（2017）。虛擬實境的沈浸感。**電腦與通訊，170**，1-5。

胡明強、黃鑲、李聖年、何穎柔（2021）。**沉浸式虛擬實境偽旅遊體驗對休閒之影響初探**。載

於國立臺北護理健康大學休閒產業與健康促進系（主編），休閒產業與健康促進學術研討會（頁 81-87）。國立臺北護理健康大學休閒產業與健康促進系，台北。

顧兆台、曾于誌（2012）。苗栗縣高中職學生休閒參與現況與休閒阻礙因素之研究。**育達科大學報，**(30)，75-102。

賴德仁、張家銘 (2020)。**精神病人社區照顧需求探討及評估**。衛生福利部國家衛生研究院。

Dupuis, S. L., & Smale, B. J. A. (1995). An Examination of Relationship Between Psychological Well-Being and Depression and Leisure Activity Participation Among Older Adults. *Society and Leisure, 18,* 67–92.

Kennedy, R. S., Drexler, J. M., Compton, D. E., Stanney, K. M., Lanham, D. S., & Harm, D. L. (2003). Configural scoring of simulator sickness, cybersickness, and space adaptation syndrome: Similarities and differences. In L. J. Hettinger & M. W. Haas (Eds.), *Virtual and adaptive environments: Applications, implications, and human performance issues* (pp. 247-278). Lawrence Erlbaum Associates Publishers. https://doi.org/10.1201/9781410608888.ch12

Kim HK. (2018). Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied ergonomics, 69,* 66-73.

Kramida G. (2015). Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE transactions on visualization and computer graphics, 22* (19), 12-31.

Neulinger, J. (1974). *The Psychology of Leisure.* Charles C. Thomas Publisher.

Park, M. J., Kim, D. J., Lee, U., Na, E. J., & Jeon, H. J. (2019). A Literature Overview of Virtual Reality (VR) in Treatment of Psychiatric Disorders: Recent Advances and Limitations. *Frontiers in Psychiatry, 10*, 1-9. https://doi.org/10.3389/fpsyt.2019.00505

Ragheb, M. G., & Beard, J. G. (1982). Measuring Leisure Attitude. *Journal of Leisure Research, 14,* 155-167.

Seppo Roy. (2023). Leisure and meaning in life. *Frontiers in Psychology, 14*, 1-11. https://doi.org/10.3389/fpsyg.2023.1074649

Tinsley, H. E., & Tinsley, D. J. (1986). A theory of the attributes, benefits, and causes of leisure experience. *Leisure Sciences, 8* (1), 1–45. https://doi.org/10.1080/01490408609513056

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Investigating the Potential of 6DoF AR Glasses in Educational Metaverse Applications

Yuehan ZHAI[1], Lijie ZHANG[2], Shuang JI[2*]
[1]Hangzhou Entel Foreign Language School, Hangzhou, China
[2]Zhejiang University, Hangzhou, China
Yuehanzhai@hotmail.com, lijiezhang1006@zju.edu.cn, shuangji@zju.edu.cn

Abstract: *The concept of an educational metaverse offers a future of immersive learning, yet current platforms often lack the spatial interaction needed for effective embodied experiences. While Augmented Reality (AR) bridges digital and physical, prevalent 3 Degrees of Freedom (3DoF) AR limits movement. Six Degrees of Freedom (6DoF) AR glasses track both position and orientation, offering enhanced immersion crucial for interactive learning. This paper investigates the potential of 6DoF AR glasses for the educational metaverse, focusing on game-based learning. We review the limitations of desk-based metaverse and 3DoF AR in education, explore 6DoF AR trends including gesture interaction, and outline a methodology for developing a proof-of-concept game-based learning scenario using XREAL Ultra glasses. This scenario demonstrates how overlaying digital content onto the real world can create complex, otherwise unattainable, learning contexts. We discuss the implications for immersive education and suggest future research directions.*

Keywords: Augmented Reality, 6DoF, Educational Metaverse, Game-Based Learning, Immersive Learning

1. Introduction

The field of education is increasingly exploring immersive technologies like Augmented Reality (AR) and Virtual Reality (VR) to enhance learning. This aligns with the vision of an "educational metaverse"—persistent, interconnected digital spaces for learning and collaboration (Zhai et al., 2023; Chu et al., 2025). While the educational metaverse promises embodied and interactive experiences, current "desk metaverse" platforms accessed via standard computers often lack the necessary spatial immersion and natural interaction crucial for many learning types, particularly those requiring physical movement and hands-on practice (Song et al., 2023).

AR offers a bridge between the digital and physical, allowing digital content to overlay the real world. This is promising for game-based learning, which benefits from interactive engagement (Jantanukul, 2024; Sakr & Abdullah, 2024). However, common AR applications utilize 3 Degrees of Freedom (3DoF) tracking, allowing users to look around but not move physically in space relative to virtual content (Zhang et al., 2024). This limitation hinders embodied interaction and dynamic scenarios needed for effective spatially-aware learning.

Six Degrees of Freedom (6DoF) AR glasses represent a significant advancement, tracking both position and orientation. This enables users to move freely within a physical space, with virtual content anchored realistically. This capability provides a much stronger sense of presence and allows for natural, embodied interactions, particularly beneficial for game-based learning requiring physical exploration and manipulation of virtual elements integrated into the physical environment. Advanced features like gesture identification further enhance intuitive interaction. Despite these advantages, the specific applications and practical realization of 6DoF AR glasses, especially leveraging natural interaction methods for game-based learning within the educational metaverse, remain less investigated.

This paper aims to explore this potential. We will review the limitations of current educational metaverse and 3DoF AR, examine 6DoF AR trends, and detail a methodology for a proof-of-concept game-based learning scenario using 6DoF

AR glasses to create learning contexts difficult in reality. Finally, we discuss implications and future research directions for utilizing 6DoF AR in the educational metaverse.

2. Literature Review

2.1. The Educational Metaverse: Aspirations and Current Limitations

The metaverse is moving from sci-fi to real tech, impacting fields like education. The educational metaverse aims for a persistent, shared, 3D digital space for immersive learning, interaction, and collaboration. This includes virtual lectures, group projects, exploring historical sites, and simulations. Benefits include better accessibility, global teamwork, increased engagement, and enabling learning too difficult/dangerous/expensive in reality (Chamola et al., 2025).

However, current educational metaverse examples are often "desk metaverse" (Nguyen et al., 2025). These use 2D or basic 3D via standard computers, with avatars navigated by keyboard/mouse. While offering social presence and shared space, they lack deep spatial immersion and natural interaction needed for much learning (Shadiev et al., 2024).

Desk metaverse limits are significant for learning needing hands-on or physical connection. Skills involving object manipulation, spatial understanding, or physical coordination are poorly supported by abstract inputs. Simulating experiments, practicing technical skills, or exploring environments requiring physical interaction and spatial awareness are difficult. This disconnect limits effective game-based learning scenarios primarily to those without embodied interaction or physical exploration (Chu et al., 2025).

2.2. Augmented Reality in Education: The Transition from 3DoF to 6DoF

Augmented Reality offers an immersive education alternative by overlaying digital content onto the real world (Tang et al., 2022). This keeps learners in familiar surroundings while adding relevant digital info (enhanced books, instructions, hunts).Much accessible AR uses 3 Degrees of Freedom (3DoF) tracking (Guo et al., 2022), tracking only orientation (pitch, yaw, roll). This lets users look around virtual content or view 360 media (Shadiev et al., 2024), useful when stationary. However, 3DoF lacks positional tracking (Microsoft, 2025) – it doesn't track physical movement. This limits dynamic learning needing physical exploration or manipulation. Users can't walk around virtual objects or have content respond to their location (Dutta, 2024). It's insufficient for AR metaverse blending and limits embodied, spatially aware learning needing physical navigation or location-tied tasks.

6DoF AR adds positional tracking via SLAM (Servières et al., 2021), tracking both orientation and position. This stably anchors virtual content to the physical world, and perspective updates with movement, providing strong presence and natural interaction (Yoon et al., 2025). 6DoF unlocks possibilities: learners can physically walk around virtual models, inspect from any angle, and interact with content anchored to real-world points. This enables scenarios like exploring virtual buildings in a room, conducting augmented experiments, or playing games requiring physical navigation/interaction. This freedom and accurate response to physical position are crucial for creating engaging, effective, embodied learning blending digital and physical realities.

2.3. Technological Trends in 6DoF AR Glasses: Towards Natural Interaction

Beyond 6DoF, modern AR glasses add features enhancing educational potential. Advanced Environmental Understanding includes plane detection for anchoring content on surfaces and depth meshing for realistic occlusion (Rohil & Prakash, 2023). Natural User Interaction (NUI) moves to controller-free methods like gesture/hand tracking (Luong et al., 2023). Cameras track hands and interpret gestures (pinch, swipe, point) as inputs. This provides intuitive, embodied interaction for manipulating virtual objects or performing tasks, aiding fine motor skills and technical learning by linking physical movement to digital outcomes.

Spatial Anchors allow virtual content to stay fixed at physical locations across different sessions or users (Zhang & Li, 2024). This is vital for persistent shared educational metaverse environments and resources in physical spaces.

Improvements in Optics, Form Factor, and Processing mean better display, more comfortable/lighter design, and increased onboard power for complex AR. These advancements in 6DoF AR glasses enable interactive, spatially aware, and naturally controlled educational experiences. Combining precise tracking, environmental understanding, and intuitive interaction addresses desk-based/3DoF limits, opening possibilities for embodied learning, especially in game-based paradigms.

3. Methodology

This research employs a design-based methodology focused on the development and description of a proof-of-concept implementation. The primary objective is to explore and demonstrate the practical application and potential pedagogical value of 6DoF AR glasses by designing and creating a tangible educational scenario that leverages their unique capabilities. This approach is suitable for investigating how the specific features of the technology can be utilized to support learning objectives and create novel educational experiences. This section outlines the hardware, software, and the process followed for developing the proof-of-concept game-based learning scenario.

3.1. Hardware and Software Environment

Hardware: The core hardware component utilized for this proof-of-concept is the XREAL Ultra AR glasses. The XREAL Ultra was selected because it represents a readily available, current generation of 6DoF AR glasses designed for both consumers and developers. It offers key technical specifications essential for this research, including integrated cameras for real-time environmental tracking and visual-inertial odometry, IMU sensors for robust motion tracking, and the capability to provide accurate 6DoF spatial tracking data.

Software: The Unity Engine serves as the primary development platform. Unity is a versatile and widely used game engine that provides a robust framework for creating interactive 3D and AR/VR content. Its extensive feature set, including physics simulation, animation systems, and support for various platforms, makes it an appropriate choice for developing complex interactive scenarios. Unity's ecosystem also includes tools and assets that can accelerate development.

SDK: The XREAL SDK for Unity is the crucial link that enables the Unity application to interface directly with the XREAL Ultra hardware and access its advanced functionalities. This SDK provides developers with APIs to utilize core features such as:Spatial Recognition (6DoF): Accessing the real-time position and rotation data of the glasses within the physical environment.Hand Tracking and Gesture Recognition: Receiving data about the position and articulation of the user's hands and interpreting specific hand poses or movements as recognized gestures.

3.2. Development of the 6Dof Learning Scenario

The central component of this methodology involves the detailed design and implementation of a specific game-based learning scenario engineered to leverage the full capabilities of 6DoF AR glasses, particularly spatial tracking and gesture interaction (Figure 1). The core logic enables 6DoF exploration: the game tracks the user's position and orientation, allowing natural movement around and through the virtual model to examine structures from any angle. Hand tracking and gesture recognition via the SDK enable intuitive interaction; specific hand gestures are mapped to actions like grabbing, separating layers, or using virtual tools for diagnosis. Game mechanics and learning flow are built around these interactions, guiding users through exploration and challenges with integrated AR UI for information display. This approach leverages 6DoF to create a hands-on, spatially aware learning experience difficult to achieve traditionally. The implementation results of this scene are shown in Figure 2.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

```
void UpdateTarget(XRGrabInteractable grabInteractable, ref Pose targetPose, ref Vector3 localScale)
{
    ComputeAdjustedInteractorPose(grabInteractable, out Vector3 newHandleBar, out Vector3 adjustedInteractorPosition, out  Quaternion
adjustedInteractorRotation);
    localScale = ComputeNewScale(grabInteractable, m_ScaleAtGrabStart, localScale, m_StartHandleBar, newHandleBar,
grabInteractable.trackScale);
    targetPose.rotation = ComputeNewObjectRotation(adjustedInteractorRotation, grabInteractable.trackRotation);
    ComputeNewObjectPosition(adjustedInteractorPosition,  adjustedInteractorRotation,
        targetPose.rotation, localScale, grabInteractable.trackRotation,
        m_OffsetPose.position, m_ObjectLocalGrabPoint, m_InteractorLocalGrabPoint,
        out Vector3 targetObjectPosition);
    targetPose.position = AdjustPositionForPermittedAxes(targetObjectPosition, m_OriginalObjectPose, m_PermittedDisplacementAxesOnGrab,
m_ConstrainedAxisDisplacementModeOnGrab);
}
```

- **EXPLAINATION**:
    1. calculate the target rotation and position based on the current Interactor's pose and manipulation mode.
    2. adjust the target position based on axis constraints.
    3. adjust the target position based on axis constraints.
    - This method synthesizes all interaction parameters and states to output the final object target pose and scale.

Figure1. Key Code - 6DOF object control through single-hand and multi-hand interaction

Figure 2. 6DOF spatial tracking for medical models and gesture interaction

3.3. Potential Evaluation Approach (Conceptual)

While this methodology focuses on the development of the proof-of-concept, a subsequent research phase would involve rigorous evaluation to assess its effectiveness. A mixed-methods approach could be employed, including: (1) Usability testing to gather qualitative feedback on the comfort of the glasses, the intuitiveness of the gesture interactions, and the overall user experience; (2) Observational studies to analyze how learners physically interact with the augmented environment and virtual content; (3) Pre- and post-assessments or in-game performance metrics to measure learning gains related to anatomical knowledge and diagnostic skills; and (4) Questionnaires to gauge student engagement, motivation, and their perception of 6DoF AR as a learning tool compared to traditional methods.

4. Discussion and Future Study

4.1. Potential of 6DoF AR Glasses in Education

The development of the forensics scenario illustrates the significant potential of 6DoF AR glasses for the educational metaverse and game-based learning. By providing full spatial tracking and natural gesture interaction, 6DoF AR overcomes the limitations of less immersive technologies. It enables embodied learning where physical movement and interaction are integral, fostering deeper understanding. Features like hand tracking make interactions intuitive, increasing engagement. The ability to overlay complex, interactive simulations onto real environments allows for safe, accessible practice of skills and concepts that are difficult or impossible to teach traditionally, thereby enriching the educational metaverse.

4.2. Challenges and Considerations

Despite the potential, challenges remain. High hardware costs limit accessibility. Developing 6DoF AR content requires specialized skills and tools. Technical issues like tracking robustness and performance optimization need continued improvement. Pedagogical frameworks specific to 6DoF AR are needed, and educators require training. Ensuring equitable access and addressing privacy concerns are also crucial for widespread adoption.

4.3. Future Study

Future research should empirically evaluate the impact of 6DoF AR learning scenarios on student outcomes and engagement. Investigating optimal pedagogical designs and interaction methods for 6DoF AR is needed. Research into scalable deployment models and addressing technical limitations will be important. Exploring collaborative 6DoF AR experiences in the educational metaverse is another promising area.

4.4. Conclusion

6DoF AR glasses offer a powerful new capability for the educational metaverse, particularly for creating immersive and interactive game-based learning experiences. By enabling full spatial tracking and natural gesture interaction, they overcome key limitations of existing technologies, allowing for the creation of rich, embodied learning scenarios integrated with the real world. While challenges exist, the potential for 6DoF AR to transform education by making complex concepts tangible and practice accessible is significant. Continued research and development are vital to fully harness this technology for the future of learning.

Acknowledgments

This study was funded by National Science and Technology Major Project (Grant No. 2022ZD0115904), 2021 National Natural Science Foundation of China (Grant No. 62177042), 2024 Zhejiang Provincial Natural Science Foundation of China (Grant No. Y24F020039), 2024 Graduate Student Science Popularization Competence Enhancement Program (Grant No. KXYJS2024008), 2024 Zhejiang Educational Science Planning Project (Grant No. 2024SCG247).

References

Chamola, V., Peelam, M. S., Mittal, U., Hassija, V., Singh, A., Pareek, R., Mangal, P., Sangwan, D., de Albuquerque, V. H. C., Mahmud, M., & Brown, D. J. (2025). Metaverse for Education: Developments, Challenges, and Future Direction. *Computer Applications in Engineering Education*. https://doi.org/10.1002/cae.70018

Chu, X., Wang, M., Spector, J. M., Chen, N.-S., Chai, C. S., Hwang, G.-J., & Zhai, X. (2025). Enhancing the flipped classroom model with generative AI and Metaverse technologies: insights from lag sequential and epistemic network analysis. *Educational Technology Research and Development*. https://doi.org/10.1007/s11423-025-10457-2

Dutta, R. (2024). Virtual Reality: An Overview of its Features and What the Future Holds. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, *12*(4), 1282. https://www.ijraset.com

Guo, H. J., Bakdash, J. Z., Marusich, L. R., & Prabhakaran, B. (2022). Augmented Reality and Mixed Reality Measurement Under Different Environments: A Survey on Head-Mounted Devices. *IEEE Transactions on Instrumentation and Measurement*, *71*, 5025115. https://doi.org/10.1109/tim.2022.3218303

Jantanukul, W. (2024). Immersive Reality in Education: Transforming Teaching and Learning through AR, VR, and Mixed Reality Technologies. *Journal of Education and Learning Reviews*, *1*(2), 51-62. https://doi.org/10.60027/jelr.2024.750

Luong, T., Cheng, Y. F., Möbus, M., Fender, A., & Holz, C. (2023). Controllers or Bare Hands? A Controlle d Evaluation of Input Techniques on Interaction Performance and Exertion in Virtual Reality. *IEEE Trans actions on Visualization and Computer Graphics*, *29*(11), 4633-4643. https://doi.org/10.1109/TVCG.2023.332 0211

Microsoft. (2025). Windows Mixed Reality The Working Principle of Inside-Out Tracking System. Microsoft L earn. https://learn.microsoft.com/zh-cn/windows/mixed-reality/enthusiast-guide/tracking-system

Nguyen, A., Huynh, L., Dang, B., et al. (2025). Conceptualizing and enhancing metaverse literacy for educatio n. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13486-9

Sakr, A., & Abdullah, T. (2024). Virtual, augmented reality and learning analytics impact on learners, and edu cators: A systematic review. *Education and Information Technologies*, *15*, 19913-19962. https://doi.org/10.1 007/s10639-024-12602-5

Servières, M., Renaudin, V., Dupuis, A., & Antigny, N. (2021). Visual and Visual-Inertial SLAM: State of th e Art, Classification, and Experimental Benchmarking. *Journal of Sensors*, 2021, 1-26. https://doi.org/10.11 55/2021/2054828

Shadiev, R., Chen, X., Reynolds, B. L., Song, Y., & Altinay, F. (2024). Facilitating cognitive development an d addressing stereotypes with a cross-cultural learning activity supported by interactive 360-degree video te chnology. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13461

Song, Y., Cao, J., Wu, K., Yu, P. L. H., & Lee, J. C. K. (2023). Developing "Learningverse"—A 3-D Metav erse Platform to Support Teaching, Social, and Cognitive Presences. *IEEE Transactions on Learning Techn ologies*, *16*(6), 1165-1178. https://doi.org/10.1109/TLT.2023.3276574

Tang, Y. M., Chau, K. Y., Kwok, A. P. K., Zhu, T., & Ma, X. (2022). A systematic review of immersive te chnology applications for medical practice and education-Trends, application areas, recipients, teaching cont ents, evaluation methods, and performance. *Educational Research Review*, *35*, 100429. https://doi.org/10.101 6/j.edurev.2021.100429

Yoon, S. H., Bianchi, A., Seifi, H., Kim, J. R., Vatavu, R.-D., Lee, J., & Lee, G. (2025). Beyond Glasses: F uture Directions for XR Interactions within the Physical World. In CHI EA '25: *Proceedings of the Exten ded Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1-6). Article 770. htt ps://doi.org/10.1145/3706599.3706715

Zhai, X., Chu, X., Chen, M., Shen, J., & Lou, F. (2023). Can Edu-Metaverse Reshape Virtual Teaching Com munity (VTC) to Promote Educational Equity? An Exploratory Study. *IEEE Transactions on Learning Tec hnologies*, *16*(6), 1130-1140. https://doi.org/10.1109/TLT.2023.3276876

Zhang, R. L., Yu, S. yu, GAO, R. feng, Wang, L. fei, & He, X. chun. (2025). How does technology-based e mbodied learning affect learning effectiveness? – Based on a systematic literature review and meta-analytic approach. *Interactive Learning Environments*, 1–24. https://doi.org/10.1080/10494820.2025.2479176

Zhang, X., & Li, S. (2024). A real-time, robust and versatile visual-SLAM framework based on deep learning networks. *arXiv preprint arXiv:2405.03413*. https://doi.org/10.48550/arXiv.2405.03413

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

The Art of Unlearning: An Exploration of the Affective Privacy Unlearning Model in Online Education

Yaowen KUANG [1], Junjie Gavin WU [2], Yiyu WU [1], Tao WANG [1*]
[1] Hubei Key Laboratory of Digital Education, Faculty of
Artificial Intelligence in Education, Central China Normal University, Wuhan
[2] Faculty of Applied Sciences, Macao Polytechnic University, Macao SAR
kuangyaowen@mails.ccnu.edu.cn, gavinjunjiewu@gmail.com, wuyy@mails.ccnu.edu.cn, tmac@ccnu.edu.cn,

Abstract: *With the rapid advancement of online education, the phenomenon of emotional deficiency has become increasingly pronounced. In recent years, there has been a growing emphasis within academia on researching the application of emotional interaction and affective computing technologies in online learning environments, with the objective of effectively mitigating the shortcomings in emotional communication prevalent in this domain. Affective computing aims to explore learners' emotional states by utilizing technological means to collect various state data during the learning process, thereby identifying and analyzing the emotional characteristics of learners. However, the collection and use of learner data raise a series of ethical and legal issues, including privacy and data security concerns. These privacy-related information and characteristic patterns are interrelated and highly sensitive. The leakage could pose numerous security risks, thereby restricting the full application and realization of the value of learning data in educational transformation. With the growing maturity of machine unlearning technology, we integrate it into the affective computing model within online learning environments. This study innovatively proposes the concept and technical framework of the Affective Unlearning Model (AUM), and demonstrates the implementation of the AUM within a facial affective computing model in online learning environments as an example. The research results indicate that this affective unlearning model can effectively achieve the unlearning of private data without occupying significant computational resources, thus safeguarding the students' right to be forgotten. Additionally, it ensures the subsequent usability of the model.*

Keywords: affective computing, machine unlearning, online learning, privacy preservation

1. Introduction

In recent years, AI has deeply integrated into education, especially with the rise of online learning platforms accelerated by the COVID-19 pandemic. While online classrooms play a vital role in teaching, the challenge of "affective-cognitive dissociation" can lead to learner isolation and fatigue, hindering online education's progress. Affective computing is a broad technology on education that utilizes artificial intelligence to learn and perceive student emotions., facilitating the delivery of personalized learning services. Personalized services require extensive collection of learners' data, such as facial expression. However, collecting and processing personal data poses significant ethical and legal risks, especially concerning privacy and data security. It is worth noting that even when private data have been removed from the database, the machine learning model can still retain and "remember" much of the underlying information(Nguyen et al., 2022). The "right to be forgotten" has been codified in privacy laws like the European Union's General Data Protection Regulation (GDPR)(2016) and China's Personal Information Protection Law (2021). In online education, this right ensures students' private information—including raw data in databases and traces in models like affective computing systems—is erased upon graduation or account termination. By minimizing unnecessary data retention and dissemination, such measures mitigate privacy breaches and safeguard user security.

To fulfill this requirement for forgetting, the simplest way is to retrain a new model using the retained data. However, this approach is often impractical because it consumes huge computing resources. Therefore, to effectively remove the specified data and its influence on the affective computing model, we propose an Affective Unlearning Model (AUM) from a technical perspective, utilizing the technical, "machine unlearning" (Aman et al., 2021). With our proposed the AUM, we can remove the specific student's privacy data from the model without the need for retraining while maintaining the usability. This approach not only safeguards intellectual property and privacy, but also ensure that the model retains its original predictive power even after the specified privacy information is removed.

2. Affective Unlearning Model

2.1. Definition of the Affective Unlearning Model

The AUM refers to the process by which a trained Affective Computing Model (ACM) intentionally forgets specific privacy data points from its training dataset, effectively rendering the model as though it had never encountered the forgotten data. Consider a cluster of data that we want to remove the training dataset from the trained ACM, denoted as x^*. An unlearning process U is defined as a trained Affective Computing Model (M_{ACM}), a training dataset D, and an remaining dataset D/x^* to a unlearned Affective Unlearning Model M'_{AUM}, which ensures that M'_{AUM} performs as though it had never seen the unlearning dataset x^*. In this manner, the unlearning process is defined as:

$$M'_{AUM} = U\left(M_{ACM}, D, D/x^*\right)$$

Based on the above definition of the Affective Unlearning Model, the core objectives of the affective unlearning model include the following three main aspects:

Data deletion and privacy protection. To ensure students' exercise of the right to be forgotten, institutions must implement data deletion protocols to eliminate unnecessary personal information, preventing misuse. This requires balanced mechanisms that maintain model accuracy and functional integrity while removing sensitive data, achieving optimal equilibrium between privacy protection and system performance.

Data poisoning response and harmful data deletion. When the ACM is subject to data poisoning attacks, it may misinterpret students' emotional states, resulting in incorrect classification results. We can maintain the security and integrity of the data by using unlearning to delete the harmful data. This prevents it from spreading further or causing additional damage. The model can more accurately assess students' emotional states, allowing it to offer resources and recommendations that better meet their needs and enhance the overall user experience.

Remove biases to ensure fairness. In the context of affective computing for online learning, models can be adversely affected by inherent biases or errors present in historical training data, resulting in biased predictions and unfair decisions. To address these potential sources of bias, we implement a forgetting mechanism. This approach not only significantly enhances the predictive accuracy of the models but also serves as a powerful catalyst for advancing educational decision-making toward a more just and equitable framework.

2.2. The technical framework of the Affective Unlearning Model

Based on the aforementioned definition of the AUM, this study constructed its technical framework, as illustrated in Figure 1. Specifically, the AUM comprises three main components: model training, model unlearning and model evaluation.

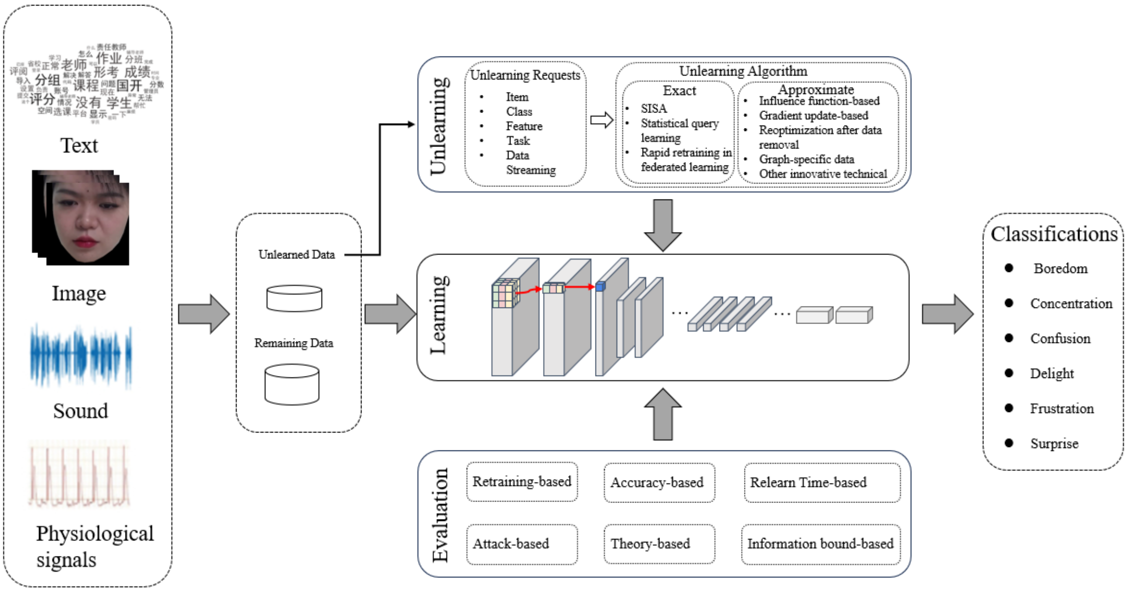


Figure 1. The framework of affective unlearning model.

①Model training: The meticulous extraction of facial expression features is the construction of an online affective computing model. First, capture facial images through the device. Next, it is essential to carefully extract key and relevant feature information from facial images. Good features are a vital part of an emotion recognition. Deep learning has proved to be a very effective feature extraction method. Last training the affective computing model (ACM) utilizing the extracted features. The trained ACM can be deployed in online learning scenarios to support personalized learning, facilitate educational assessment, and provide feedback. For example, Mehta et al.(2022) introduced a three-dimensional DenseNet self-attention neural network (3D DenseAttNet) for the automatic detection of students' engagement in E-learning platforms.

②Model unlearning: These unlearning algorithm include: item-level forgetting(Bourtoule et al., 2021); feature-level forgetting(Warnecke et al., 2021); class-level forgetting(Tanha et al., 2020); task-level forgetting(Liu et al.,2022); and data streaming forgetting(Nguyen et al., 2017). Within the framework of AUM in online learning, tailored forgetting algorithms can be effectively chosen to address the various types of forgetting requests from individual students. Specifically, in cases where a student requests the deletion of all their facial data for privacy protection, this data can be categorized as a class, allowing class-level forgetting algorithms to be used to completely erase all facial data associated with that student. Furthermore, when managing textual datasets that contain sensitive information, such as personal names and genders, these sensitive terms can be classified as features forgetting. By employing feature-level machine forgetting algorithms, these confidential details can be effectively removed from the textual data. Additionally, in light of the potential for racial bias in facial recognition software(Rhue, 2018), where the interpretation of emotions may vary according to a person's race, feature-level forgetting algorithms can be utilized to mitigate the influence of racial bias within such software systems.

③Model evaluation: A thorough evaluation of the AUM is crucial. Accuracy: as the primary measure of model performance, accuracy is crucial in this context. The AUM should consistently demonstrate efficient affective computing to ensure that data unlearning is accomplished without compromising predictive accuracy. Verifiability: to ensure effective protection of private data, a series of attack tests, including backdoor attacks, membership inference attacks, and model inversion attacks, are conducted. These tests aim to verify whether the AUM successfully defends against potential data leakage risks, thus demonstrating its effectiveness in safeguarding private information.

3. Application of the Affective Unlearning Model

This study explores facial expression-based affective computing models, which rely on extensive student facial image datasets. These images contain highly sensitive personally identifiable information which is not only unique and irreversible but also includes additional privacy details like the individual's age and gender, posing severe privacy risks if leaked. To mitigate such risks, we focus on implementing effective unlearning mechanisms to ensure thorough deletion and protection of specific facial image data from the model .

Figure 2. Example images that are sampled from our dataset.

In this study, we employed a self-constructed educational dataset. This dataset was meticulously developed from actual online learning environments. This dataset includes 11,442 samples, as shown in the Figure 2. We use the UNSIR algorithms (Choi et al., 2023) for unlearning. The method works as follows:

Error-Maximizing Noise: Introduce an error-maximizing noise. During the error maximization process, we maintain the weights of the pre-trained model in a fixed state. Given a noise matrix randomly initialized using a normal distribution N (0, 1), we aim to determine the noise that maximizes the model's classification loss function on unlearning data. We prepare for the subsequent erasure of information by determining the noise that is the inverse of the unlearning data.

Impair and Repair: For impair, we train the model on a small subset of data from the original distribution which also contains generated noise. This step effectively destroys the weight parameters of the initial model to remove the memory of the unlearning data from the initial trained model. For Repair, since the impair step may compromise the identification of the remaining data, we implemented a repair step. We retrain the impaired model on the remaining data for a single epoch to repair the model weights, ensuring that the model retains high performance.

We adopted top-2 accuracy as the primary metric. Additionally, we employed Membership Inference Attack (MIA)(Shokri et al. 2017) and established the forgetting score, where a lower score indicates better forgetting performance. Furthermore, We introduced the Normalized Machine Unlearning Score (NoMUS). This score effectively quantifies the overall performance of affective unlearning models by balancing the model's utility with the forgetting score.

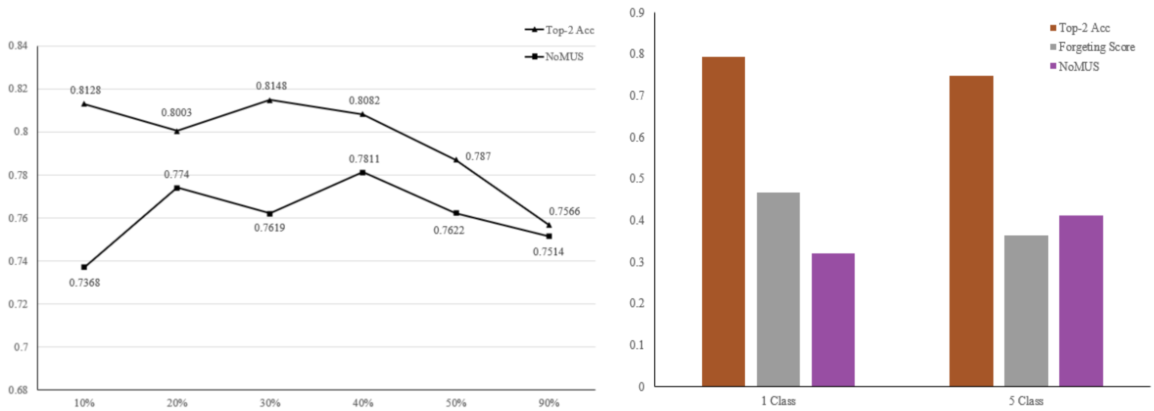


Figure 3. The performance of the AUM on forget class of data in the scenario of randomly forgetting face data and forgetting face data with specific identity.

(a) randomly forgetting face data (b) forgetting face data with specific identity.

We show the performance of our proposed method in the scenario of randomly forgetting face data and forgetting face data with specific identity. Results demonstrate the wide applicability of our method. Figure 3(a) depicts the performance of the AUM in the context of its forgetting mechanism following the random dropout of data at varying proportions. The observed results indicate a declining trend in accuracy as the proportion of forgotten data escalates. Nevertheless, amidst the pursuit of a balance between data privacy preservation and data availability, even post the elimination of 90% of data, the model's accuracy persists within acceptable bounds. Further scrutiny into the performance of NoMUS reveals that irrespective of whether 10% or 90% of data is randomly discarded, favorable forgetting performance are obtained, strongly suggesting the model's robustness concerning variations in data forgetting rates.

Within the framework of the AUM, we define the complete set of an individual's facial images as a class. Figure 3(b) meticulously illustrates the performance of the AUM after the forgetting of facial image data of one student (i.e., single class, 1 class) and five students (i.e., multiple classes, 5 classes). The method proposed in this study demonstrates outstanding unlearning outcomes in the task of forgetting data across multiple classes. This signifies the effective removal of relevant data from the model, rendering it unrecoverable through membership inference attacks, thereby robustly establishing the efficacy and security of the AUM model in data forgetting.

Table 1. Overall performance of various machine unlearning algorithms on our dataset with $\lambda = 1/2$. We emphasize the best score using bold and the second best score using italics. In the forgetting score, the lower is better.

Method	Test Acc	Forgetting score	NoMUS
Original Model	82.40%	0.2666	0.7916
Retrain Model	72.94%	0.0213	0.5302
FineTuningModel	75.04%	0.1321	0.6751
AUM	75.57%	0.1267	0.6654

Our results are compared with three baseline unlearning methods in Table 1. The original model exhibits the highest top-2 accuracy of 0.8240. However, the AUM model exhibits favorable performance in terms of top-2 Acc compared to other forgetting methods. For forgetting scores, the AUM model's efficacy in data privacy protection is approaching that of the retrain model, demonstrating its ability to maintain high model performance while safeguarding data privacy. For the NoMUS, the result indicates that the AUM model is capable of effectively implementing unlearning strategies without consuming large amounts of computational resources comparable to those of the retrained model, while maintaining a high level of recognition accuracy, thus achieving a good balance between performance and efficiency.

Thus, the AUM model can efficiently achieve the forgetting of students' private data, thereby effectively safeguarding the students' right to be forgotten and genuinely ensures the security of educational data.

4. Conclusion

The AUM framework integrates machine unlearning to enhance trust and security in affective computing, while promoting student engagement in digital learning. It optimizes data efficiency by reducing redundancy and streamlining management/analysis processes, thereby improving learning data quality for informed instructional decisions and personalized education. Through algorithmic fairness enhancements, the framework mitigates model bias and improves emotional recognition accuracy, eliminating demographic barriers (e.g., race/economic status) to educational resources while reducing systemic discrimination risks. This innovation fosters inclusive learning ecosystems that nurture socio-emotional skills and empathy through unbiased educational interactions.

AUM plays a pivotal role in online learning, which not only effectively protects students' privacy, but also optimizes the performance and reliability of the system, laying the foundation for a harmonious and efficient online learning environment. This study theoretically defines the core concepts of the AUM, constructs the framework, and explores specific practical application scenarios, thereby laying a solid foundation for further exploration of the affective unlearning. In the feature, we will continue to be dedicated to promoting the widespread application of this model in the field of education, accelerating the process of educational intelligence.

Acknowledgements

This study was funded by National Natural Science Foundation of China under Grants 62477017, 62372333, 62277021 and 62277029, and Fundamental Research Funds for the Central Universities, Central China Normal University, under Grants CCNUai017, CCNUai018, CCNU22JC011 and CCNU22QN013.

References

Choi, D., & Na, D. (2023). Towards machine unlearning benchmarks: Forgetting the personal identities in facial recognition systems. *arXiv preprint* arXiv:2311.02240.

Deng, H., Yang, Z., Hao, T., Li, Q., & Liu, W. (2022). Multimodal affective computing with dense fusion transformer for inter-and intra-modality interactions. *IEEE Transactions on Multimedia*, 25, 6575-6587.

Elatlassi, R. (2018). Modeling student engagement in online learning environments using real-time biometric measures: electroencephalography (EEG) and eye-tracking.

Gu, Y., Yang, K., Fu, S., Chen, S., Li, X., & Marsic, I. (2018, July). Multimodal affective analysis using hierarchical attention strategy with word-level alignment. *In Proceedings of the conference. Association for Computational Linguistics. Meeting* (Vol. 2018, p. 2225).

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Liu, B., Liu, Q., & Stone, P. (2022, November). Continual learning and private unlearning. In *Conference on Lifelong Learning Agents* (pp. 243-254). PMLR.

Mehta, N. K., Prasad, S. S., Saurav, S., Saini, R., & Singh, S. (2022). Three-dimensional DenseNet self-attention neural network for automatic detection of student's engagement. *Applied Intelligence, 52(12)*, 13803-13823.

Nguyen, T. T., Duong, C. T., Weidlich, M., Yin, H., & Nguyen, Q. V. H. (2017). Retaining data from streams of social platforms with minimal regret. In *Twenty-sixth International Joint Conference on Artificial Intelligence*.

Regulation, P. (2018). General data protection regulation. *Intouch*, 25, 1-5.

Rhue, L. (2018). Racial influence on automated perceptions of emotions. Available at SSRN 3281765.

Savchenko, A. V., Savchenko, L. V., & Makarov, I. (2022). Classifying emotions and engagement in online learning based on a single facial expression recognition neural network. *IEEE Transactions on Affective Computing, 13(4)*, 2132-2143.

Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017, May). Membership inference attacks against machine learning models. In *2017 IEEE symposium on security and privacy (SP)* (pp. 3-18). IEEE.

Standing Committee of the National People's Congress. Personal Information Protection Law of the People's Republic of China. 2021.

Tanha, J., Abdi, Y., Samadi, N., Razzaghi, N., & Asadpour, M. (2020). Boosting methods for multi-class imbalanced data classification: an experimental review. *Journal of Big data, 7,* 1-47.

TS, A., & Guddeti, R. M. R. (2020). Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks. *Education and information technologies, 25(2),* 1387-1415.

Warnecke, A., Pirch, L., Wressnegger, C., & Rieck, K. (2021). Machine unlearning of features and labels. *arXiv preprint arXiv:2108.11577*.

Zhang, Y., Tao, X., Ai, H., Chen, T., & Gan, Y. (2024). Multimodal Emotion Recognition by Fusing Video Semantic in MOOC Learning Scenarios. *arXiv preprint arXiv:2404.07484*.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Under the Perspective of Human-AI Co-Education: Construction and Application of the "Tripartite Co-Creation" Classroom Teaching Model

Hui-Min ZHANG[1] , Wei LIANG[2] , Xiao YANG[3]
[1,2,3] Shenzhen Education Information Technology Center, China
yangxiao_ouyi@163.com

Abstract: *Under the background of the deep integration of artificial intelligence (AI) and education and teaching, the human-AI collaborative classroom exists such realistic dilemmas as shallow application of AI and lack of student subjectivity. Based on the concept of human-AI co-education and with the goal of improving students' autonomous learning capabilities, this study constructs a classroom teaching model of "Tripartite Co-Creation" under the perspective of human-AI co-education from the aspects of teaching links, teacher-student activities, teaching evaluation, and resource tools, providing a new path to break the classroom dilemma. Based on this, this study takes Shenzhen junior high school mathematics as an example to carry out a quasi-experiment in teaching, and comprehensively verifies the promotion effect of this teaching model on students' autonomous learning capabilities from the dimensions of learning process and learning outcome, so as to provide reference for the efficient development of "Tripartite Co-Creation" classroom teaching.*

Keywords: Human-AI co-education, Tripartite co-creation, Human-AI deep integration, Classroom teaching model, Autonomous learning capability

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

人机共育视域下“三元共创”课堂教学模式的构建与应用

张惠敏[1]，梁为[2]，阳小[3]
[1,2,3]深圳市教育信息技术中心，中国
yangxiao_ouyi@163.com

【摘要】在人工智能与教育教学深度融合的背景下，针对当前人机协同课堂存在的“技术浅层化应用”“学生主体性缺失”等现实困境，本研究基于人机共育理念，以提升学生自主学习能力为目标，从教学环节、师生活动、教学评价、资源工具等方面构建了“三元共创”课堂教学模式，为破解课堂现实困境提供新路径。基于此，本研究以深圳市初中数学为例开展教学准实验，分别从学习过程和学习结果两个维度综合验证该教学模式对学生自主学习能力的促进作用，为“三元共创”课堂教学的高效开展提供参考。

【关键词】人机共育；三元共创；人机深度融合；课堂教学模式；自主学习能力

1. 研究背景

人机协同课堂作为学生核心素养培养和课堂变革的的核心实践范式，正经历着从工具性应用到认知架构的深层次变革。2025 年 4 月，教育部等九部门发布《关于加快推进教育数字化的意见》，明确提出要推动人工智能与教育教学全要素全过程深度融合，探索人机协同教学新模式，提高教育教学效率和质量（教育部，2025）。通过文献综述和实地调研发现，虽然各地开展了人机协同课堂教学实践，但仍然存在一些问题：一方面，人工智能技术应用存在“新瓶装旧酒”的异化风险，部分课堂正从“人灌”滑向“机灌”的认知陷阱；另一方面，人机协同的认知边界尚未厘清，技术形式化削弱学生主体性等问题日益凸显。人机共育强调以学习者的智慧增长为核心，重视人和机器之间的博弈与平衡。以人机共育助力实现高质量课堂教学和促进学生自主学习能力发展，是未来教育教学创新发展的重点方向(王一岩，2024)。

当前，人工智能正在重构教学场域中的主体关系网络，形成“师—生—机”三元主体共创共生的教育新格局。深圳市作为教育部“基于教学改革、融合信息技术的新型教与学模式”实验区和“智慧教育示范区”创建项目名单（简称教育信息化“双区”），通过构建数据驱动、人机共育的“双区联动”机制，在 100 所教育信息化“双区”实验校探索人工智能技术支撑下的三元主体共创共生新型教与学模式，为课堂教学变革提供了先行范本。本研究依托人机共育理念，构建人机共育视域下的“三元共创”课堂教学模式，并以深圳市初中数学为例开展教学准实验，探索“三元共创”课堂教学模式对学生自主学习能力提升的促进作用。

2. 人机共育的内涵和课堂特征

2.1. 人机共育的内涵

人机共育作为基于“师—生—机”互动关系构建的新型教育形态（李袁爽，2023），呈现出三重核心内涵：

主体关系的重构性共生。人机共育创新性地构建“师—生—机”三元主体协同关系。机器为师生提供丰富的资源与精准的学情分析，与教师的专业指导、学生的主动学习探索相互交织，共同推动教学活动的开展，实现三元主体在教学过程中的深度融合与共创共生发展。

认知过程的互补性进化。人机共育通过“生物脑—数字脑”的深度认知耦合，达成认知能力的跃迁式发展。机器凭记忆强化、模式识别等外源性认知功能，拓展学生的认知边界。教师则聚焦高阶思维引导、情感联结等助力学生核心素养能力提升。人机优势互补的认知模

式，能够充分汲取人机双方的优势，实现认知、能力的优化与进化。

教学机制的智适应演进。人机共育构建起“干预—自主”的动态平衡教学机制，智能设备实时采集学生的学习数据，通过智能算法进行深入诊断，洞察学生的学习问题所在。基于诊断结果，生成个性化的教学策略，依据学生学习过程中的实时反馈，动态调整教学方案。

2.2. 人机共育理念下的课堂特征

人机共育理念下的课堂中，“师—生—机”三元主体以人机共创方式开展教学活动，呈现以下核心特征。人工智能技术承担重复性、规律性等机械性认知任务，助力学生投身创造性思维与高阶认知活动，并推送元认知提示，提升学生自主学习能力（武法提，2024）。教师基于学生学习行为数据，并依此进行深度诊断分析，制定教学策略，实现数据驱动的教学决策。学生能利用人工智能技术突破传统节奏限制开展个性化学习。推动课堂实现“五个转变”，从统一到个性、从预设到生成、从讲授到自主学、从课时设计到单元统整、从知识本位到能力本位。最终形成以智能增强、数据驱动、多元转变为特色的三元共创课堂特征。

3. 人机共育视域下“三元共创”课堂教学模式的构建

综上对人机共育内涵和课堂特征的描述，本研究尝试构建人机共育视域下“三元共创”课堂教学模式（以下简称“三元共创”课堂模式），如图 1 所示。首先，教师、学生、机器三元的角色在模式中相互连结，教师转变为知识发现的引导与启发者，学生转变为探索与创造者，人工智能技术会发现和发展学生的最近发展区并充当创意激活者；其次，基于人机共育理念的课堂教学要坚持“以学习者为中心”，在课堂教学设计时，考量哪些活动能够支持学生核心素养、自主学习能力的提升；最后，致力于培养学生对知识的二次创造创新，通过项目式、探究性学习等方式贯穿课堂始末，促使学生使用知识创造性地解决真实生活问题。

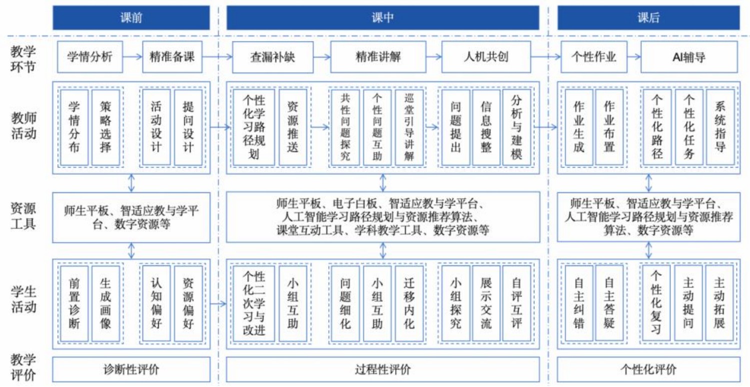


图 1 人机共育视域下“三元共创”课堂教学模式

课前阶段的教学环节包括学情分析与精准备课，智适应教与学平台基于学情大数据形成学生画像和班级学情分布，辅助教师评估并选择适合的教学策略，生成定制化的活动设计和课堂提问方案，以此促进学生的自主学习能力发展和减轻教师负担。课中阶段的教学环节包括查漏补缺、精准讲解与人机共创。首先进行查漏补缺，根据学生画像诊断结果为学生规划个性化学习路径，推送针对性学习资源，鼓励学生使用平板在智适应教与学平台上完成二次学习，并在小组内互助解决疑惑；随后在精准讲解环节则针对班级共性与个性问题，通过展示发现的问题引导学生在小组内探究解决方法；接着在人机共创环节进一步深化知识应用，教师提供背景材料后，学生利用生成式人工智能技术发掘单学科或跨学科问题，搜集整理信息，进行学科抽象建模及计算，并将结果应用于真实生活场景解释。课后阶段的教学环节包括个性作业与 AI 辅导，智适应教与学平台根据学生画像推送个性作业和学习资源，学生自主完成后，AI 辅导解答疑问，并基于兴趣推荐更多学习材料，持续支持学生的个性化学习与发展。整个过程以学生为中心，充分利用人工智能赋能实现个性化、精准化和高效化的教学。

4. 教学实践及效果检验

4.1. 研究对象

本研究选择深圳市 X 中学八年级 A 班和 B 班作为实验对象，以数学学科为例，开展一学期（2024 年 9 月—2025 年 1 月）的教学准实验。A 班、B 班的学生在总人数、性别占比、学习成绩等方面大致相当。A 班为实验班，B 班为对照班，两个班的学生人数均为 48 人。

4.2. 研究过程

为排除教师个体差异对实验结果的影响，本研究选取同一名数学教师对两个班级进行授课。对于实验班，采用本研究团队构建的人机共育视域下“三元共创”课堂模式开展教学；对于对照班，采用传统的教学模式开展教学（未采用师生平板、智适应教与学平台等工具，未进行基于人机共育理念下的教学设计等）。本研究聚焦学习过程和学习结果，以学生自主学习能力提升为目标，因此从过程维度和结果维度验证所构建的教学模式对学生自主学习能力提升的促进作用。过程评估考量学生在课堂上学习的自主性，具体表现为学生反馈、合作交流、自主探究、展示汇报四类课堂主动学习行为（李环，吴砥，朱莎，等.2023）。为此本研究将学生课堂主动学习时长占比作为过程维度指标。结果评估考量课堂上学生对知识的运用与创造，具体体现为学生在课堂上能够进行更高频次的深度学习。因此，本研究将学生在课堂上深度学习活动发生次数作为结果维度指标。研究者在准实验期内对实验组和对照组连续开展 20 节课的数据采集并利用 SPSSAU 网页数据科学算法平台系统进行独立样本 t 检验，比较两组在过程维度和结果维度的差异性。

4.3. 研究工具

4.3.1. AI 课堂智慧分析平台

关于学生课堂主动学习时长占比和深度学习活动发生次数的数据采集，本研究采用 AI 课堂智慧分析平台分析课堂录像视频。该平台通过智能技术对课堂录像文件进行分析，将视频切片、语音转录文本、音频片段、图像片段等多媒体素材标签化，将非结构化课堂行为转化为结构化数据指标并分析（刘梦君，2024）。输出教师单课报告、数据对比分析等形成循证教研的科学证据链。为排除环境和课时不同带来对数据的影响，本研究采集的实验组和对照组的每节课均为相对应的课时内容，并都在具备多机位采集音视频的录播教室中进行。

4.3.2. 知识深度模型

知识深度模型（Depth of Knowledge，DOK）由美国教育评价专家 Norman L.Webb 于 1997 年提出，是一种评估深度学习的有力测量工具（Herman, et al., 2013）。其理论框架融合了布卢姆目标分类学的认知分层逻辑，通过分层任务设计（DOK1—DOK4）推动学生从浅层记忆向高阶思维进阶。其中 DOK1 和 DOK2 又称为浅层学习，DOK3 和 DOK4 称为深度学习。每个层次反映了学生学习活动所需的知识深度。主要应用在学科教学、跨学科整合、技术融合教育等领域。AI 课堂智慧分析平台通过对课堂录像的视觉识别和语音转录，将非结构化文件转化为结构化文本、音频、视频片段等，基于模型关联不同层级，实现数据分析与统计。

表 1 知识深度模型框架

DOK 层级	认知水平	学习活动	学习层级
DOK1 回忆/重现	能回忆事实、信息或过程，处理低级别信息，需一步的思维活动	能列举现象或事实，并向大家展示或陈述	浅层学习
DOK2 技能/概念	能利用信息或概念，并能完成两步以上的任务	理解概念、规律，利用概念、规律解释简单现象，解决简单问题	
DOK3 策略性思维与推理	具备逻辑推理的思维能力，能制定复杂的计划，常常需要多步的思维过程	能从多角度思考问题，认识复杂的现象和过程，能综合性的解释问题，提出解决方案	深度学习

DOK4 拓展性思维	能通过调查、思考解决受多种条件影响的问题	能多源收集信息得出结论，能提出解决问题的改进或替代的方案

4.4. 研究结果

4.4.1. 人机共育过程维度的分析结果

实验组和对照组的课堂主动行为时长占比的 t 检验结果如下表所示。可以看到，组别样本在学生反馈、自主探究、展示汇报三方面具有显著性差异（$p<0.05$）。根据表 2 可以发现，实验组经过一学期的“三元共创”课堂教学实践后，学生会更加主动地在课堂上给予教师、同学反馈，更主动的进行自主探究，结合探究结果更积极进行展示和汇报。这说明“三元共创”课堂模式可以有效促进学生课堂主动学习行为的发生。

表 2 实验组和对照组过程维度指标数据独立样本 t 检验结果

分析项	组别	样本量	平均值	标准差	t	df	p
学生反馈	实验组	20	0.15	0.02	19.823	28.898	0.000
	对照组	20	0.08	0.01			
合作学习	实验组	20	0.10	0.01	-1.710	22.300	0.101
	对照组	20	0.11	0.00			
自主探究	实验组	20	0.19	0.01	3.187	28.181	0.004
	对照组	20	0.18	0.01			
展示汇报	实验组	20	0.07	0.01	12.425	23.999	0.000
	对照组	20	0.04	0.00			

4.4.2. 人机共育结果维度的分析结果

实验组和对照组的课堂深度学习行为发生次数的 t 检验结果如下表所示。可以看到，组别样本在 DOK1、DOK2、DOK3、DOK4 四个方面均具有显著性差异（$p<0.05$）。根据表 3 可以发现，实验组经过一学期的“三元共创”课堂教学实践后，学生在课堂的深度学习层次的学习行为相较对照组发生频次更多，浅层学习层次的学习行为相较对照组来说发生频次更少。这说明学生总体在课堂上更多由浅层学习转向深度学习，在“三元共创”课堂模式下能够促进学生更深度的学习和对知识的创造。

表 3 实验组和对照组结果维度指标独立样本 t 检验结果

分析项	组别	样本量	平均值	标准差	t	df	p
DOK1	实验组	20	2.50	0.51	-12.329	38.000	0.000
	对照组	20	4.50	0.51			
DOK2	实验组	20	6.50	4.73	-2.362	38.000	0.023
	对照组	20	10.50	5.92			
DOK3	实验组	20	19.50	5.92	2.673	38.000	0.011
	对照组	20	14.50	5.92			
DOK4	实验组	20	9.00	2.99	11.052	38.000	0.000
	对照组	20	1.50	0.51			

4.5. 研究结论

由上分析可知，“三元共创”课堂模式对学生自主学习能力提升有明显促进作用。具体来讲，实验组的学生在课堂上的主动学习行为发生时长更多，学生投入度、参与度、积极性更好；更深度的学习行为发生频次更高，学生逐步由浅层学习转向深度学习。可能的原因解释是：首先，此模式强调素养导向的个性化学习，在基于学生最近发展区为其规划学习任务，使得其学习难度适宜，激发学生学习积极性和自主性；其次，此模式相对传统教学模式发生了“五个转变”，更多地为学生提供了反馈、交流、展示的时间和空间，这能很好激发学生展示自我的动力；最后，教师在此模式下更多的变讲授为提问，在充分激发学生好奇心的同时，引导学生探究知识运用的本质，体会问题解决的成就感，这能极大地激发学生的热情，更积

极、主动地投入学习、探究中。因此，在基于人机共育的教学实践中，学生的热情、积极性、好奇心得到充分激发，主观能动性、自我效能感得到充分提升，其自主学习能力得到有效提升。综上，"三元共创"课堂模式对解决人机还未深度融合使得学生主体作用和自主性没能充分发挥的问题、促进学生自主学习能力和核心素养发展有一定参考价值。

5. 结语

本研究构建人机共育视域下"三元共创"课堂教学模式，为破解人工智能与课堂教学深度融合的问题提供创新路径。其次以深圳初中数学为例，开展教学准实验，验证了该教学模式可以有效促进学生自主学习能力提升。本研究结果虽然为人工智能时代的人机深度融合提供一定的课堂模式参考，但仍有不足和局限之处。如本研究教学准实验中，仅选取八年级学生作为研究对象，并以数学学科为例开展为期一学期的教学实验，存在教学周期较短、未采集定性数据、结论可推广性有待考证等问题。为此，在后续研究中，将通过学科拓展、实验周期拉长、学段和年级扩充、补充定性数据等进行更多、更广、更深入的教学实践，对该教学模式持续优化和迭代，以提升其适用性和可推广性。

基金项目

全国教育科学规划 2021 年度立项课题"基于智慧教育的高质量新型教与学体系构建与应用研究"（课题批准号：BCA210085）

参考文献

教育部等九部门（2025）。**关于加快推进教育数字化的意见**。

王一岩、朱陶、杨淑豪（2024）。人机协同教学：动因、本质与挑战。**电化教育研究，45**，51-57。

李袁爽、张子航、王运武（2023）。基于 ChatGPT 的人机共教、共育和共创模式。**中国医学教育技术，37**，383-389+396。

武法提、夏志文、高姝睿（2025）。以生成式人工智能重塑智慧学习环境：从要素改进到生态重构。**电化教育研究，46**，4-63。

李环、吴砥、朱莎、郭庆和罗志强（2023）。深度学习视域下智慧课堂教学模式的构建及应用研究。**现代教育技术，2**，61-70。

刘梦君、蔡雨菲、蒋新宇和杨兵（2024）。数字化转型下 AI 课堂观察系统使用对师生课堂行为的影响——基于 W 中学 412 节课例的循证研究。**现代教育技术，34**，64-73。

李璇律、田莉、张向众(2023)。学习者深度学习何以发生:从行为序列到行动进向。**远程教育杂志，41**(03)，84-93。

Herman, J., & Linn, R. (2013). On the road to assessing deeper learning: the status of smarter balanced and parcc assessment consortia. *National Center For Research on Evaluation Standards and Student Testing*.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

University Students' Perceptions of Language Learning in Immersive Learning Environments: An Artefact Analysis Perspective

Yin YANG*, Siu Cheung KONG
Department of Mathematics and Information Technology
The Education University of Hong Kong, Hong Kong SAR
yyin@eduhk.hk

Abstract: *This study examines postgraduate students' perceptions of language learning within immersive learning environments in the metaverse. With the metaverse increasingly integrated into higher education, it is essential to understand the specific features that enhance student motivation and engagement. The study involved 19 postgraduate students, with data collected through artefacts depicting their perceptions of learning in immersive environments and follow-up interviews. The findings contribute to the development of educational strategies for virtual learning environments, highlighting the critical importance of student-centered design and artefact creation. Furthermore, the study explores the broader implications of these findings, offering insights into how immersive technologies can be effectively utilised to enhance language education and improve the overall learning experience.*

Keywords: immersive learning environments, artefact analysis, learning perceptions, language learning

1. Introduction

In recent years, the integration of the metaverse into higher education has garnered significant interest from educators and researchers alike. The metaverse enables students to have immersive experiential learning opportunities and exposes them to diverse educational environments. Numerous studies have investigated various metaverse learning environments to enhance both teaching and learning (Beck et al., 2023; Wang et al., 2023). However, these studies often overlook the specific features and characteristics that students desire in the metaverse, which could potentially increase their motivation. Notably, there is a scarcity of research focusing on learner-created artefact creation and surveys to gather feedback within these virtual spaces, especially for language learning (Godwin-Jones, 2023).

The primary objective of this study is to examine postgraduate students' perceptions of language learning and teaching in the metaverse.

2. Literature Review

2.1. Metaverse Literacy

Metaverse literacy is an emerging concept critical to navigating and learning within immersive virtual environments. This literacy encompasses the skills and knowledge necessary to effectively interact, communicate, and learn in the metaverse. As the digital landscape evolves, metaverse literacy not only involves technical skills like navigating virtual spaces or understanding blockchain technology but also includes critical thinking and collaborative skills in digital settings (Kong & Kwok, 2023). Recent research by Latino et al. (2024) highlights the lack of metaverse-specific literacy programs in higher education, suggesting a gap between current educational curricula and the demands of immersive learning environments. Furthermore, the study by Nguyen et al. (2025) emphasizes the role of metaverse literacy in enhancing student engagement and motivation, indicating that students with higher levels of digital literacy tend to have more positive learning experiences in these environments.

2.2. Metaverse for Language Learning Purposes

The application of the metaverse in education has potential for pedagogical innovation and learning experiences. The immersive nature of the metaverse allows for hands-on, experiential learning that traditional classrooms cannot replicate. For instance, educational strategies incorporating the metaverse can cater to diverse learning needs, offering personalised learning pathways (Chen et al., 2023). However, challenges persist, such as the digital divide and the need for substantial infrastructural investments, which can hinder the widespread adoption of metaverse in education. Studies by Wang et al. (2023) have explored these challenges, suggesting that while the metaverse presents novel educational opportunities, it also requires careful consideration of accessibility and student readiness to ensure equitable educational benefits. The current literature has rarely explored learners' perceptions of language learning within immersive learning environments, leaving a critical gap in understanding how such settings influence language acquisition and engagement.

3. Research Design

An exploratory research design was adopted for this study. Students participated in a course based on a metaverse literacy framework (Kong et al., 2024), which included cognitive, metacognitive, affective, and social aspects. The study involved 19 master's degree students, and qualitative data were collected through student-created artefacts and focused group interviews.

(1) Students' created artefacts: After the 13 3-hour courses, students were invited to draw a digital / traditional poster showcasing their idea learning in the future metaverses. They described the pictures they draw with text.

(2) Focused group interviews: The interviews were conducted post-course to gather detailed feedback. Each interview was structured yet allow for explorative questions based on interviewee responses. The expected time was 25-30 minutes. Interviews were recorded (with consent) and transcribed for analysis.

To address the research aim — exploring postgraduates' perceptions of learning and teaching experiences within the metaverse — an artefact analysis approach was employed. A coding scheme was developed based on predefined categories relevant to metaverse learning experiences (Chang et al., 2020). The coding scheme for analysing metaverse interactions comprised four primary dimensions:

(1) Conceptual Understanding: This dimension focused on the technological aspects of the metaverse, including knowledge of VR/AR/Mixed Reality, metaverse technologies, video production, NFT/blockchains, digital twins, and AI applications.

(2) Engagement: Engagement referred to the nature of participants' experiences within the metaverse, such as interactions with teachers, AI tutors, peers, or individual activities. It also encompassed typical metaverse activities, such as teacher-led lectures, gaming, and general communication.

(3) Affection: Affection, or learners' behavioral attention, categorized participants as either developers or users and captured their emotional responses, ranging from positive to neutral.

(4) Social Aspects: This dimension explored participants' perspectives on the future of metaverse interactions, including their preferred settings (e.g., classrooms or self-defined spaces). It also emphasized the development of essential skills such as problem-solving, collaboration, and teamwork. This coding scheme enables identification and categorization of themes that emerge from the artefacts, providing insights into the students' perceptions and expectations of metaverse learning environments.

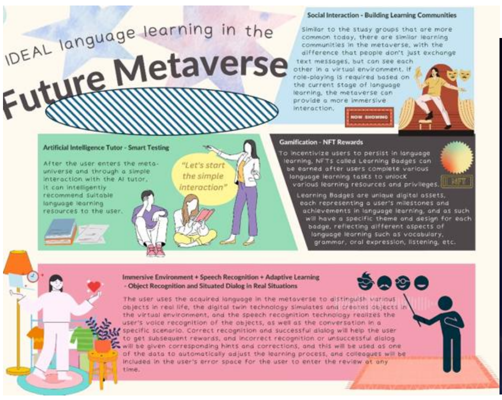
4. Results

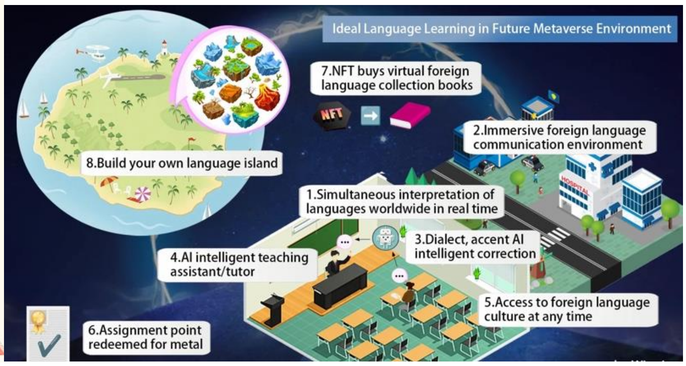
Based on the conceptual understanding reflected in the 19 artifacts analysed, 12 students (63.2%) highlighted the critical role of virtual reality (VR) in language learning. For instance, one student remarked in an interview, *"VR gives us the chance to experience real-life situations, like ordering food in a restaurant abroad, which makes learning more practical and engaging"* (Student 07). Another student added, *"It feels like you're actually in the culture, which helps you learn faster and remember better"* (Student 12). Eight students (42.1%) emphasised the importance of NFTs as a key technology for providing rewards and motivating learners to engage in language learning within the metaverse. As one student explained, *"Earning NFTs for completing tasks makes learning feel like achieving something tangible—it's like a game reward that keeps me coming back"* (Student 04). Another student mentioned that NFTs could provide personalized incentives, such as *"unlocking new challenges or tools that help me progress further in my language learning journey"* (Student 15). Additionally, 13 students (68.4%) acknowledged the significant role of AI-assisted non-player characters (NPCs) or agents in helping learners practice language skills. One student highlighted, *"Talking to AI characters feels less intimidating than speaking to real people, so I can practice without fear of making mistakes"* (Student 09). Another student shared, *"The AI agents can correct your grammar immediately and even suggest better phrases, which makes learning less stressful"* (Student 11). However, only two students (10.5%) identified the use of generated avatars as a means to depict their ideal language learning environments.

In terms of learner engagement, 15 students (78.9%) expressed that scenario-based learning was highly effective in motivating their interest in language learning. On the other hand, only six students (31.6%) preferred social learning in immersive environments. The majority of students expressed a preference for interacting with avatars or AI assistants, noting that *"learning can be like a game in immersive learning environments"* (Student 03, Student 06). The gamified approach was seen as a key driver of motivation and engagement.

Regarding learners' behavioural attention, most students visually represented their ideal learning environments in a way that demonstrated positive attitudes toward immersive language learning. These visualisations further supported their enthusiasm for such innovative methods.

Finally, concerning the social aspects and participants' visions for the future of immersive learning environments, five students (26.3%) depicted traditional classroom settings. However, the majority believed there would be no default settings for immersive learning environments, as learners could define their own environments based on various themes and preferences. One student commented, *"In the future, I think everyone will have their own personalized learning space—like a room designed specifically for their goals and interests"* (Student 10). Another student envisioned, *"a collaborative space where learners from different countries can meet virtually to practice languages together"* (Student 16). Table 1 shows four students' artifacts depicting their ideal language learning environments, illustrating diverse preferences such as gamified settings, culturally immersive scenarios, and AI-integrated tools.





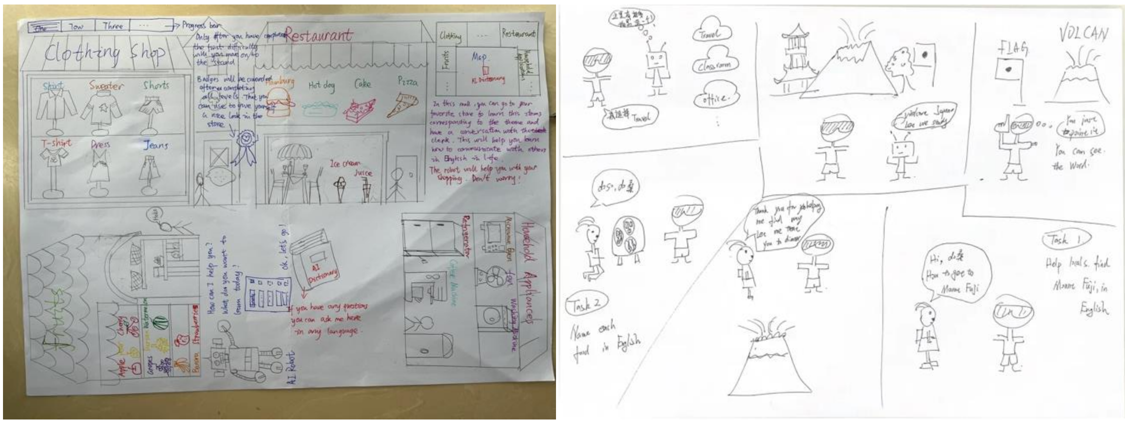


Figure 1. Examples of students' artefacts (student 02, 03, 08, and 10).

5. Conclusions and Implications

The findings of this study made contributions to the understanding of language learning in immersive learning environments within the metaverse, with a particular focus on learners' perceptions. By analysing artifacts created by postgraduate students and conducting focused group interviews, the study revealed key insights into how students perceive and engage with immersive technologies. VR was highlighted as a critical tool for contextual and experiential learning, allowing learners to practice language skills in real-life scenarios, such as interacting in a virtual restaurant or navigating a virtual airport. This aligns with previous research emphasising the role of VR in bridging the gap between theoretical learning and practical application (). Additionally, the use of NFTs as motivational rewards was seen as a novel way to sustain engagement, with students responding positively to gamified incentives that offer tangible, personalised achievements. AI-driven interactions with NPCs also emerged as a valuable feature, as they provided learners with a safe and adaptive environment for language practice, reducing anxiety and fostering confidence.

The findings have several implications for instructional design and metaverse-based educational strategies. For example, student-centered design is critical in developing virtual learning environments that align with learners' needs and preferences.

Despite of the implications, the study is not without limitations. First, the sample size of 19 postgraduate students is relatively small, which may limit the generalisability of the findings. Future research could involve larger and more diverse groups of participants. Second, the study relied on qualitative data from artifacts and interviews, which, while rich in detail, may be subject to interpretive biases. Triangulating these findings with quantitative data, such as pre- and post-assessment scores, could provide a more comprehensive understanding of the impact of immersive learning environments on language acquisition. Future research should continue to explore the evolving role of the metaverse in education, addressing its challenges and leveraging its opportunities to create inclusive and impactful learning experiences.

The study also revealed areas that require further exploration. For instance, while most students embraced VR, NFTs, and AI-assisted learning, only a small percentage found value in generated avatars or traditional classroom settings within the metaverse. This suggests that learners may prioritise functionality and interactivity over aesthetic or conventional elements.

Acknowledgements

Thanks for the funding support of MIT Departmental Research Grant 2024-25 (Project 0402I).

References

Beck, D., Morgado, L., & O'Shea, P. (2023). Educational practices and strategies with immersive learning environments: Mapping of reviews for using the metaverse. *IEEE Transactions on Learning Technologies*, *17*, 319-341.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Chen, X., Zou, D., Xie, H., & Wang, F. L. (2023). Metaverse in education: Contributors, cooperations, and research themes. *IEEE Transactions on Learning Technologies*, *16*(6), 1111-1129.

Godwin-Jones, R. (2023). Emerging spaces for language learning: AI bots, ambient intelligence, and the metaverse. *Language Learning & Technology*, *27*(2), 6-27.

Kong, S. C., & Kwok, W. Y. (2023). Metaverse literacy in primary to 12th-grade and higher education. In *The 1st International Conference on Metaverse and Artificial Companions in Education and Society*.

Kong, S. C., Yang, Y., & Zhang, H. (2024). A proposed framework for promoting metaverse literacy. In *The 2nd International Conference on Metaverse and AI Companions in Education and Society*.

Latino, M., Chiara De Lorenzi, M., Arcuti, M., & Corallo, A. (2024). Metaverse in education: What has been done so far? A systematic literature review to map benefits and limitations and to set future research and application directions. *CSEDU 2024*, *1*, 735-743.

Nguyen, A., Huynh, L., Dang, B., Pohjolainen, S., Mattila, J., Paajala, I. J., ... & Karppinen, P. (2025). Conceptualizing and enhancing metaverse literacy for education. *Education and Information Technologies*, 1-21.

Wang, Y., Su, Z., & Yan, M. (2023). Social metaverse: Challenges and solutions. *IEEE Internet of Things Magazine*, *6*(3), 144-150.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

AI-Driven Educator Persona Generation: An Intelligent Framework for Personalized Online Course Video Production

Jin-Huang QUE[1], Ya-Qi KONG[1], Xiao-Rong YU[1], Xu CHEN[1*]
[1]Central China Normal University, China
quejinhuang@mails.ccnu.edu.cn, 1277317986@qq.com, yuxr402@163.com, chenxu@ccnu.edu.cn

Abstract: *Artificial intelligence technologies powered by large language models are revolutionizing educational content creation. We propose a novel framework for personalized educational video generation that leverages AI-generated content (AIGC) to address diverse learning needs. Our three-module system integrates semantic-adaptive content transformation, identity-preserving vocal synthesis, and pedagogically-responsive visual synthesis to create customized virtual educator experiences. Experimental results demonstrate that contextually-aligned virtual teacher personas significantly enhance student engagement and learning outcomes while reducing knowledge gaps between learners with varied baseline abilities. This framework maintains low cognitive load requirements while improving knowledge retention and transfer, establishing a promising foundation for personalized education in the AI era*.

Keywords: personalized educational video generation, generative artificial intelligence, contextually-aligned virtual teacher

1. Introduction

The integration of artificial intelligence with educational technology has transformed knowledge dissemination, yet many digital platforms continue employing standardized approaches that fail to address diverse learning preferences (Cao et al., 2023). This one-size-fits-all model persists despite evidence that personalization significantly improves learning outcomes. OECD's PISA 2022 assessment reveals only 51% of students can effectively evaluate online educational resources (OECD, 2024), highlighting the urgent need for sophisticated AI-driven educational approaches.

Conventional adaptive learning systems operate within limited parameters that inadequately capture the complexity of human learning processes, resulting in suboptimal knowledge transfer (Kabudi et al., 2021). To address these limitations, we present an integrated AI-driven framework that generates personalized educational videos through virtual teacher personas.

Our approach combines semantic-adaptive content transformation using large language models, identity-preserving vocal synthesis, and pedagogically-responsive visual synthesis with appropriate nonverbal communication elements. This framework dynamically aligns virtual teacher characteristics with both subject matter and learner preferences to enhance engagement and knowledge retention.

This research demonstrates how advanced generative AI can create contextually appropriate learning experiences that enhance knowledge acquisition while reducing cognitive disparities among diverse learner populations—contributing to educational platforms that effectively address personalization challenges in increasingly digital learning environments.

2. Framework for Generating Personalized Educational Videos

Developing effective personalized educational videos presents three critical challenges: appropriate content generation, coherent narrative construction, and engaging teacher persona design. Our framework addresses these challenges through a novel integration of multimodal AI technologies across three interconnected modules, as illustrated in Figure 1.

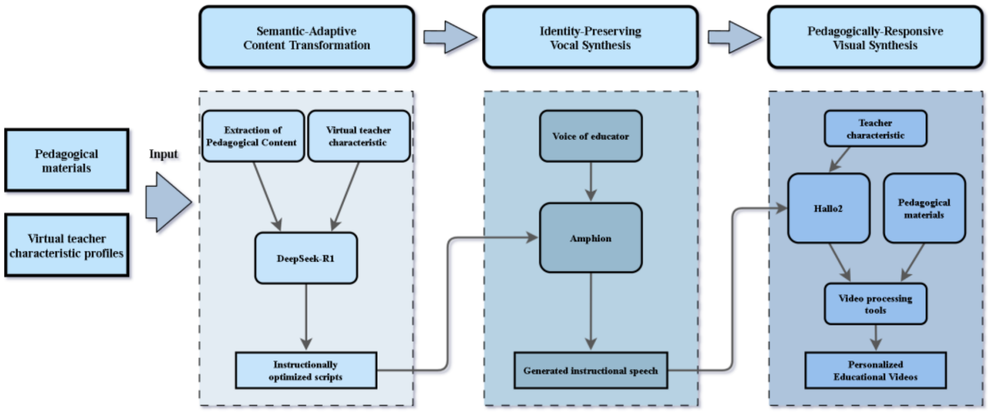


Figure 1. A personalized educational video generation framework consisting of three modules.

2.1 Semantic-Adaptive Content Transformation Module

This module transforms teaching materials to match each virtual teacher's preferred teaching methods through a structured process. The system begins by processing two inputs: original teaching content and the virtual teacher's profile listing preferred teaching approaches (problem-based learning, storytelling, etc.).

The content analysis breaks down materials into knowledge units, identifying concepts, supporting details, examples, and assessment questions. Natural language processing creates concept maps showing relationships between ideas. The module then matches teaching methods from the virtual teacher's profile with specific content restructuring patterns—transforming content into scenarios for problem-based learning, narrative structures for storytelling teachers, or debate points for discussion-focused instructors.

Using DeepSeek-R1, the transformation process preserves key information while adapting presentation style, sequence, and emphasis to match the virtual teacher's natural teaching approach. The module maintains appropriate difficulty levels and learning progression while producing teaching material that feels authentic to each virtual teacher's style.

2.2 Identity-Preserving Vocal Synthesis Module

Our speech generation component employs Amphion model (Zhang et al., 2024) to create authentic educator vocal identities with crucial paralinguistic features essential for effective knowledge transmission. This module's distinctive contribution comes from its handling of prosodic elements that convey pedagogical intent—emphasis patterns, rhythmic variations, and emotional modulations that signal conceptual importance.

By preserving these subtle communication aspects while enabling historical figures to deliver contemporary content in multiple languages, our system creates a cognitive bridge between learner expectations and educational material. This cross-modal consistency enhances content authenticity and cognitive processing throughout the learning experience.

2.3 Pedagogically-Responsive Visual Synthesis Module

The visual synthesis module uses Hallo2 (Cui et al., 2024) with Latent Diffusion Models (Rombach et al., 2022) to create visually coherent educational presentations. Operating in latent space rather than pixel space reduces computational complexity while maintaining high performance across critical metrics including Fréchet Inception Distance, Fréchet Video Distance, and synchronization consistency.

Hallo2 processes semantic text labels to dynamically adjust facial expressions and head movements corresponding to specific pedagogical moments. This allows virtual teachers to express enthusiasm when introducing key concepts, adopt contemplative demeanor when explaining complex ideas, or display encouraging expressions during challenging material. The system's multilingual support ensures broad applicability across diverse educational contexts.

The integration of these three specialized modules creates a comprehensive pipeline for producing educational videos that are both personalized to student preferences and pedagogically optimized—establishing a new paradigm for AI-enhanced educational content that addresses fundamental limitations of conventional online learning approaches.

3. Experiment Design and Results

3.1 Experiment Design

We conducted a controlled experiment to evaluate how different virtual teacher personas affect learning outcomes and cognitive load. Using a randomized design with three treatment groups, participants engaged with identical course content on Confucian thought presented by different virtual teachers: Confucius , Albert Einstein, and a generic 30-year-old Chinese male teacher, as shown in Figure 2.

Participants completed pre-test assessments to establish baseline knowledge, viewed their assigned virtual teacher lecture, and then completed post-tests measuring: learning retention (recall and comprehension of lecture content), learning transfer (application of concepts to novel scenarios) and cognitive load (mental effort expended during learning).

Figure 2. Virtual teacher videos utilized in the experiment.

3.2 Analysis of Experimental Results

Our study included 90 participants evenly distributed across three virtual teacher conditions. Pre-test results (maximum: 9) showed a mean score of 7.11, indicating moderate baseline knowledge of Confucian thought. Cognitive load measurements averaged 9.92 (out of 25), with only 10% of participants reporting high cognitive load, suggesting minimal mental burden across conditions.

Post-test assessments revealed strong learning outcomes: mean scores of 8.17 (out of 10) for learning retention and 3.61 (out of 7) for learning transfer. Analysis by knowledge level uncovered a notable convergence effect—participants with lower pre-test scores showed greater improvement, resulting in more homogeneous post-test performance: the low-knowledge group (0-6 points, n=29) achieved mean learning retention and transfer scores of 7.59 and 3.34; the medium-knowledge group (7-8 points, n=33) scored 8.52 and 3.60; and the high-knowledge group (9 points, n=28) scored 8.39 and 3.89.

Most significantly, despite variations in pre-test scores across groups, participants viewing the Confucius virtual teacher achieved the highest learning retention, as shown in Figure 3. This finding is particularly notable given that the Confucius group did not have the highest baseline knowledge. The contextual alignment between Confucius and Confucian philosophy appears to have created cognitive resonance that enhanced learning effectiveness, suggesting that strategic matching of virtual teacher characteristics to course content optimizes learning outcomes.

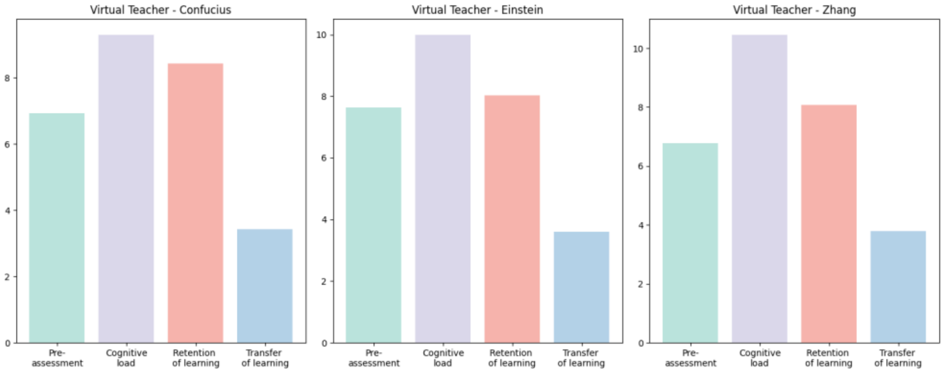


Figure 3. Average experimental scores of participants grouped by the virtual teacher's image in the video.

4. Conclusions and Future Perspectives

Our AI-driven framework advances personalized education by integrating large language models, speech synthesis, and video generation to create tailored virtual teaching experiences. By aligning virtual teacher characteristics with subject matter and learner preferences, we enhance engagement and learning outcomes.

Experimental results demonstrate that our approach reduces knowledge gaps between students with varying baseline abilities while maintaining low cognitive load. Notably, the superior performance observed with contextually-aligned virtual teachers (Confucius teaching Confucian philosophy) highlights the importance of strategic persona-content matching beyond mere visual customization.

However, the framework still has three primary limitations: lack of real-time responsiveness to student confusion; constrained domain-specific adaptability across different subject disciplines; and the need for robust ethical data governance to protect student information and ensure educational equity.

As AI capabilities continue to advance, thoughtfully integrating these technologies into educational contexts promises to transform teaching and learning in the digital age, creating more engaging, accessible, and effective educational experiences for diverse learners worldwide.

References

Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt. *arXiv preprint arXiv:2303.04226*.

Cui, J., Li, H., Yao, Y., Zhu, H., Shang, H., Cheng, K., Zhou, H., Zhu, S., & Wang, J. (2024). Hallo2: Long-duration and high-resolution audio-driven portrait image animation. *arXiv preprint arXiv:2410.07718*.

Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and education: Artificial intelligence*, 2, 100017.

OECD. (2024). *PISA 2022 results (Volume V): Learning strategies and attitudes for life*. OECD-Publishing.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10684-10695.

Zhang, X., Xue, L., Gu, Y., Wang, Y., Li, J., He, H., Wang, C., Liu, S., Chen, X., & Zhang, J. (2024). Amphion: an Open-Source Audio, Music, and Speech Generation Toolkit. *2024 IEEE Spoken Language Technology Workshop (SLT)*, 879-884.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Investigating the Application of Augmented Reality in Enhancing Memory Function in Patients with Schizophrenia

Siang-Ruei LI[1], Chia-Hui HUNG[1*]
[1]Chung Shan Medical University, Taiwan
larry133728@gmail.com, chhung@csmu.edu.tw

Abstract: *This study aims to investigate the effects of augmented reality (AR) activities, designed based on daily life context, on the memory of individuals with schizophrenia. A total of 25 participants were recruited and screened using the Mini-Mental State Examination (MMSE, >24 points) before being assigned to either the experimental or control group. The experimental group underwent six weeks of AR memory training based on daily life context, and the intervention outcomes were assessed using a Contextual Memory Test (CMT). The results indicated no significant difference in MMSE scores between the two groups. However, a non-parametric analysis of the pre- and post-test results of the CMT revealed a significant difference ($p = 0.006$), with a large effect size (Hedges' $g = 1.2$), suggesting a substantial improvement in memory. Additionally, participants showed improvements in work performance and attention. Nonetheless, AR interventions should consider potential risks of dizziness or seizures, which may be mitigated by reducing display brightness and minimizing rotational effects. In conclusion, this study supports the effectiveness of AR in enhancing memory in individuals with schizophrenia, contributing to their social adaptation.*

Keywords: augmented reality, schizophrenia, memory, daily life context, occupational therapy

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

應用擴增實境於促進思覺失調患者記憶力之成效探討

李祥瑞 [1]，洪佳慧 [1*]
[1] 中山醫學大學職能治療學系，台灣
larry133728@gmail.com，chhung@csmu.edu.tw

【摘要】本研究目的為以日常生活情境設計的擴增實境(Augmented Reality, AR)活動，探究對思覺失調患者記憶力的影響。參與者以 MMSE(>24 分)篩選後 25 位，分實驗與對照組。實驗組參與 6 週的日常生活情境之 AR 記憶力活動，並以情境記憶力測驗評估成效。結果顯示，兩組於 MMSE 分數無顯著差異，但在情境記憶力測驗前、後測於無母數分析(p=0.006)具顯著差異，且效果量(Hedges' g=1.2)，顯示在記憶力提升有大效果，此外，參與者在工作表現與注意力亦有提升。然，AR 介入應注意參與者暈眩或癲癇，可以降低顯示器亮度和減少旋轉以避免。綜上所述，AR 可有效提升思覺失調症患者的記憶力，有助於其社會適應。

【關鍵字】擴增實境、思覺失調、記憶力、日常生活情境、職能治療

1.前言

擴增實境(AR)為元宇宙重要技術，透過在真實影像中加入虛擬物體，讓使用者在真實環境與虛擬物體互動，不同於虛擬實境的全虛擬環境(Mystakidis, 2022)。思覺失調是一種慢性精神疾病，特徵包含認知和知覺損傷。認知障礙是影響患者生活品質的核心問題，其中，記憶力是理解、學習和推理的基礎，因此記憶力缺失會影響患者融入社會的能力(Lepage et al., 2014)，因此處理記憶力缺失對職能治療至關重要。

研究指出，患者的情境記憶力缺失，造成結合情境線索以形成連貫記憶方面有困難，導致記憶碎片化(Waters et al., 2004)。然而目前仍缺乏介入情境記憶力的方法。因此，本研究目的為運用 AR 的沉浸性和情境互動等特點，訓練患者運用情境線索的能力，並探究其成效。

2.文獻探討

擴增實境(AR)技術因其即時互動性與沉浸式體驗，近年來逐漸被導入醫療與復健領域。尤其在職能治療中，AR 不僅能提升治療的趣味性與參與度，更提供多感官的學習機會，成為輔助治療的新興工具，例如 De Cecco 等人(2023) 運用共享 AR 技術，讓治療師與個案身處相同的虛擬與真實環境，藉此觀察與評估個案執行擺設餐桌的過程。結果顯示，共享 AR 可以即時回饋個案的表現，進而提升個案的參與動機與治療師的同理心。

此外，Wang 等人(2025)設計了 AR 任務提示系統，協助認知障礙者完成速食餐點製備。該系統以 AR 圖像提示餐點製備步驟，聲音提示正確與否，結果顯示參與者的工作能力顯著提升，並減少體力與技巧需求，增強自我控制感與生活品質。

除了功能性訓練以外，陳威全等人(2023)則利用 AR 提升自閉症患者的情緒理解力，該研究的第一部分透過 AR 顯示 8 種社會情緒的 3D 臉部表情，藉此讓參與者逐一辨認各表情所代表的情緒。第二部分則運用社會情境腳本訓練，讓參與者學習在互動過程中選擇適當的表情。結果顯示，所有參與者的情緒推論能力顯著改善，證實 AR 在教導情緒辨識的有效性。

綜觀上述研究可見，AR 技術具備即時互動、多感官輸入與沉浸式體驗等特性，有助於提升學習動機、技能表現與社交理解。這些研究多聚焦於工作調整與情緒辨識，但針對思覺失調症患者記憶力訓練的應用仍屬少見，顯示此領域尚有研究空間。有鑑於此，本研究擬以 AR 為介入工具，探討其對思覺失調症患者記憶力提升的影響，期望為職能治療的精神領域提供

創新且實用的介入策略。

3.研究方法

3.1 研究設計、對象與環境：本研究採「非隨機實驗控制組前後測設計」，比較受試者在參與六周 AR 記憶力訓練前後的表現。研究於某精神專科醫院招募思覺失調患者，納入標準包含症狀穩定且 MMSE24 分以上。排除標準包括認知障礙、癲癇及暈眩病史。共納入 25 名參與者，分實驗組與對照組。(實驗組 13 名，平均 54 歲；對照組 12 名，平均 58 歲)。本研究經過中山醫學大學附設醫院第一人體研究倫理審查委員會審查合格後才進行收案，計畫編號為 CS1-24082，許可有效期間為西元 2024 年 9 月 11 日至西元 2027 年 9 月 10 日。

表一、人口學基本資料

	實驗組(n=13)	對照組(n=12)	
	n (%)	n (%)	*p* 值
年齡 [a]	54.17	58.08	0.200
性別 [b]			0.312
男性	8 (62%)	4 (33%)	
女性	5 (38%)	8 (67%)	
教育程度 [b]			0.753
碩士	0 (0%)	1 (8%)	
大學/大專	1 (8%)	2 (17%)	
高中	8 (62%)	5 (41%)	
國中	3 (22%)	2 (17%)	
國小	1 (8%)	2 (17%)	
住院年數 [a]	5.19	6.99	0.609
診斷 [b]			0.534
F200	6 (46%)	7 (58%)	
F201	0 (0%)	1 (8%)	
F208	2 (17%)	2 (17%)	
F250	1 (7%)	1 (8%)	
F258	1 (7%)	0 (0%)	
F259	1 (7%)	0 (0%)	
F312	2 (17%)	0 (0%)	
F333	0 (0%)	1 (8%)	

[a] 獨立樣本　t 檢定，數值為平均值

[b] 卡方檢定，數值為數量，括號內為百分比

3.2. 測量工具：簡易心智量表(MMSE)：評估受試者認知功能，共 30 分，高分表示認知功能較佳。本研究用以作為納入標準篩選，排除認知障礙者。情境記憶測驗(CMT)：採用劉倩秀等人孫淨如(2018)中文版，CMT 根據記憶模式理論(The Stage Model of Memory)及資訊處理理論(Information Processing Framework of Memory)發展，用於評估受測者與日常生活情境有關的記憶功能及使用情境策略進行記憶的能力。CMT 包含早晨版與餐廳版圖片卡(各 20 項物品)，施測時給予 90 秒記憶，並以回憶出的物品數量作為評分指標。CMT 已針對 50 歲以上沒有明顯記憶生能的成人建立標準化常模，且已完成輕度認知障礙個案之先導試驗，信效度方面，本研究所採用的立即回想部分，針對健康族群和輕度認知障礙者所建立的組內相關係數 ICC 分別為 0.79 和 0.91，具備良好之信度，此外，有多篇研究文獻驗證其效度，顯示 CMT 具備

良好的跨文化效度、建構效度和區辨效度，因此選用 CMT 作為本研究之評估工具。

3.3. AR 記憶力活動設計：AR 活動使用 Sketchfab 取得 3D 素材，以 MAKAR 軟體建立場景與互動模式，最後使用平板輸出操作，活動結合視覺與聽覺刺激，提供沉浸式體驗。本研究 AR 包含六大情境(準備行李、旅行、節慶文化、購物、準備餐點和盥洗)，及四類記憶力訓練(短文記憶、空間關係記憶、圖像記憶、聲音辨識)，以情境線索與物品的連結提升記憶力。第一週主題為「準備行李」，參與者與虛擬人物互動後，系統會顯示如「襪子和衣服在衣櫃裡」等提示，參與者需記住物品與其位置，並在空間中尋找。此活動訓練參與者物件與空間的聯結記憶，並協助熟悉 AR 互動模式。第二週以「國外旅行」為主題，場景設於機場，著重聽覺記憶訓練。參與者需點擊虛擬物體(如飛機、免稅店等)記住對應聲音，之後聆聽一段聲音序列並回答相關順序問題。活動中提供字首記憶法輔助記憶。第三週延續旅行主題，場景包括花店、餐廳與神社，著重空間與短文記憶。參與者需記住場景中物品的位置與排列，並回答如「在拉麵左方的物品是什麼？」等問題，訓練情境記憶與三維空間關係理解。第四週場景為動物園，包含企鵝館、獅子館與禮品店三關。企鵝館訓練圖像記憶與數量辨識；獅子館訓練是否重複出現的記憶判斷與順序回憶；禮品店則強化物品與位置的對應記憶。第五週主題為「節慶文化」與「購物」，設有市集、室內與廣場三場景。市集中記憶購物清單並選出正確項目；室內場景中記住聖誕樹裝飾樣式與物品位置；廣場中記憶三首聖誕歌曲旋律並辨識對應曲目。第六週則聚焦日常活動「盥洗與準備餐點」。浴室場景中記住 8 樣盥洗用品並回憶；廚房場景需根據短文描述依序完成早餐製作步驟；客廳場景則要求記憶餐桌上物品的空間配置並回答相對位置。

圖一、第五周活動畫面

圖二、研究流程

3.4. 研究流程：本研究流程共分歷時 9 週，分為三個階段如圖二所示。

3.5. 資料分析：本研究使用 IBM SPSS 進行數據分析。以獨立樣本 t 檢定比較兩組前測差異，CMT 的前後測數據經 Shapiro-Wilk 檢定後，顯示不具常態分佈，且樣本數較少，故使用無母數檢定分析前後測變化。若 $p < 0.05$ 表示實驗組情境記憶力有顯著提升。

4.研究結果與討論

表二、各評估工具的數據和統計結果

	實驗組 (n=13)		對照組 (n=12)		
	Pre-test	Post-test	Pre-test	Post-test	
評估工具	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	*p* 值
CMT	7.00(2.858)	10.15(3.579)	8.00(3.490)	6.17(3.070)	0.006
MMSE	28.54 (1.450)	-	27.67(1.969)	-	

獨立樣本 t 檢定顯示，實驗組與對照組在 MMSE(p=0.112)與 CMT 前測分數(p=0.444)上無

顯著差異，確認基線認知功能相當。無母數分析結果顯示，實驗組 CMT 分數顯著提升(p=0.006，Hedges' g = 1.2)，平均分數從 7.00 提升至 8.00，顯示 AR 記憶訓練有效增進情境記憶。質性分析發現，多數參與者認為記憶力提升，且記憶策略使用增加，反映其社會適應能力增強。然而，部分參與者出現暈眩，建議調整顯示器亮度與減少旋轉以降低不適風險。

5.結論

本研究結果顯示，運用 AR 進行記憶力訓練能有效提升思覺失調患者的情境記憶能力，並促進記憶策略的運用。實驗組在 CMT 測驗中顯著進步，且多數參與者自覺記憶力與專注力有所提升，顯示 AR 的沉浸性與互動性有助於改善認知功能。然而，部分參與者出現暈眩反應，未來應優化 AR 設計以提升安全性與使用舒適度。整體而言，AR 記憶力訓練為思覺失調患者認知復健提供了一項具潛力的介入方式。

參考文獻

陳威全、程于芳、王欣宜（2023）。運用擴增實境技術增強自閉症類群障礙症孩童理解情緒辨識與情緒推論能力。**國立臺灣科技大學人文社會學報**，**19**（2），159-186。

De Cecco, M., Luchetti, A., Butaslac, I., Pilla, F., Guandalini, G. M. A., Bonavita, J., Mazzucato, M., & Hirokazu, K. (2023). Sharing Augmented Reality between a Patient and a Clinician for Assessment and Rehabilitation in Daily Living Activities. *Information*, *14*(4), 204.

Lepage, M., Bodnar, M., & Bowie, C. R. (2014). Neurocognition: clinical and functional outcomes in schizophrenia. *Can J Psychiatry*, *59*(1), 5-12.

Mystakidis, S. (2022). Metaverse. *Encyclopedia*, *2*(1), 486-497.

Wang, X., Lu, H. Y., Lau, C. Y. C., & Tong, R. K. (2025). Augmented reality rehabilitation for stroke. *Digital Technology in Public Health and Rehabilitation Care*, 219-237.

Waters, F. A. V., Maybery, M. T., Badcock, J. C., & Michie, P. T. (2004). Context memory and binding in schizophrenia. *Schizophrenia Research*, *68*(2), 119-125.

Sun, J. R., Lioa, H. M., Chang, L. W., Chen, K. H., & Liu, C. H. (2018). Developing the Pictograms of the Contextual Memory Test for Cultural Suitability in Taiwan. 職能治療學會雜誌, 36(2), 203-227.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

The Effect of AI-Assisted Technology on Students' Mandarin Speaking Proficiency

Feng-Huang LIN[1], Yu-Ju LAN[1*], Chung-Hao HUANG[2]

[1]Department of Teaching Chinese as a Second Language, National Taiwan Normal University, Taiwan

[2]Department of Electrical Engineering, Chung Yuan Christian University, Taiwan

yujulan@gmail.com

Abstract: *This study aims to explore the effects of AI-assisted technology on Mandarin-speaking proficiency and speaking anxiety in language learning. The study adopted a case study approach, involving five Mandarin learners. During the experiment, five participants attended in an eight-session speaking course over two months. The first four weeks followed a traditional teaching method, and the later four weeks we integrated AI-assisted speaking training into the curriculum. The AI system incorporated course-related exercises and provided real-time feedback to the learners. Evaluations were conducted before and after each phase to define the effectiveness of AI-assisted teaching on speaking performance and anxiety. The findings indicate that incorporating AI technology into speaking courses has practical implications. The results of this study can be a reference for future applications of AI technology in conjunction with traditional Mandarin-speaking instruction.*

Keywords: Mandarin teaching, speaking courses, AI-assisted instruction, Mandarin speaking skills, speaking anxiety

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

AI 科技技術輔助對學生華語口說能力之成效影響

林鳳凰[1]，籃玉如[1*]，黃崇豪[2]
[1]國立臺灣師範大學 華語文教學系，臺灣
[2]中原大學 電機工程學系，臺灣
yujulan@gmail.com

【摘要】 本研究旨在探究 AI 科技技術融入華語口說教學之口說能力與口說焦慮的影響。研究採用個案研究法，以五名華語學習者為對象，進行兩個月共八節課之華語口說教學課程，前四週採用傳統教學法，後四週加入 AI 教學互動系統以輔助教學，系統整合了課程內容相關之題目，並能提供學生及時反饋。在每階段開始前後皆進行評量，以探討 AI 技術輔助教學之成效與口說焦慮之影響。結果顯示，使用 AI 技術輔助口說課程具有實踐意義，研究結果可作為未來運用 AI 科技技術搭配傳統口說教學之參考。

【關鍵字】 華語教學；口說課程；AI 科技輔助教學；華語口說能力；華語口說焦慮

1. 研究動機與目的

隨著全球化發展，跨國企業要求員工有良好的溝通能力，許多駐台外商企業不僅提供外籍員工額外的華語課程補助，甚至要求員工通過華語程度考試，而在語言掌握的聽說讀寫四項技能中，企業普遍重視口說技能，且期待員工能展現即時性的語言學習成效。然而，在教學現場，老師也多能注意到學習者在口說方面容易產生焦慮，進而影響學習動機與成效（林佳臻，2017）。

2022 年 11 月 ChatGPT 問世，帶動生成式 AI 快速發展，引起大眾關注並被廣泛運用在多元領域，推出僅兩個月，2023 年 1 月時 ChatGPT 擁有超過 1 億用戶，已成為增長最快的消費者軟體應用程序之一。市面上相繼出現許多 AI 學習功能型軟體，尤其在語言學習方面，生成式 AI 學習工具強調即時回饋與互動式口說訓練，提高了學習者的參與度和積極性，能獲得更好的學習效果（張靜淑，2023； Chen, 2024），其中，即時的回饋亦符合 Schmidt (2012) 所提出的注意力假說(noticing hypothesis)，學習者必須注意到，習得才會發生。過往研究也指出，當學習者注意到糾錯回饋的當下，即使沒有立即更正，學習者卻能在後續的話輪中，產出接近或完全正確的目標與形式（潘冠霖，2018）。顯示即時回饋在於引導學習者注意語言形式，促進語言學習的重要性。

而當學生感受到 AI 學習工具的實用性後，更會提升學生持續使用的意願（Zou et al., 2023）。也有研究指出，AI 口說工具能協助學習者逐步克服口說焦慮，增強自信心與學習動力（劉慧雯、蔡佩珊，2023）。且 AI 學習工具的應用與發展不限於商業領域，政府部門也積極倡導，教育部國教署 113 年度推動「AI 輔助自主學習」計畫，將 AI 融入「自學、共學、互學、導學」的教學策略中。結合上述背景與資源，本研究透過將 AI 技術融入教學，設計 AI 口說學習程式並提供即時回饋，提升學習者在口說課程的成效與動機。同時也將透過問卷，了解 AI 技術輔助口說學習是否能達到減少學習者的口說焦慮，強化學習者信心。

2. 研究方法與系統設計

本研究採個案研究法，探討五位學習者，包含三名英語為母語者、一名西班牙母語者與一名印尼為母語的學習者在傳統教學融入 AI 科技前後的口說能力成效差異。五位研究對象之華語文能力介於初學到Ａ2 級別間，具基本的華語口說能力，可以用基本語法和簡單的生詞

與他人就有限範圍的日常話題進行交談。然而，語音語調較不準確，詞彙有限，表達時經常犯錯且說話速度有點慢。本研究為期兩個月，分為兩階段進行八次教學，第一至四週採用傳統式教學，每次一小時，第五到八週，引入 AI 科技，透過 AI 口說教學互動系統提供即時回饋，以觀察其對學習者口說能力的影響。並於每階段實驗前後進行華語口說程度測試，口說測試考題與評分方式皆模擬 TOCFL 華語文能力測驗口說測驗準備級之題型、範圍與評分指標，並由三位華語老師進行審閱評分，以了解融入 AI 輔助教學後，學習者的口說進步狀況。此外，在 AI 輔助教學前後，以問卷調查方式探討研究對象的口說焦慮變化，進一步了解 AI 技術在語言學習中的應用成效。

本研究開發一款華語 AI 教學互動系統，如圖二所示。本系統是以 Gradio 框架開發前端 Web 介面，系統中包含輸入介面與歷史紀錄等頁面。當受測者進入輸入介面時，可以在此頁面下，觀看授課教師所給的課堂教材(如圖片、影片等)，接著受測者輸入個人基本資料(如班級、姓名等)，以便進行記錄。當個人資料輸入完成後，學生可以根據 AI 虛擬教師的提問，利用華語進行口說回答，當回答完畢後，本系統會將學生所錄製的聲音傳送到 ChatGPT，透過 OpenAI 公司的 Whisper 模型，將使用者錄音轉換為文字後進行分析。分析完畢後，AI 虛擬教師將會依照實體授課教師所給的評分標準，進行判斷給分，評分過程中，AI 虛擬教師會將分析後的結果，以語音及文字的方式，告訴受測者他在該題的測試結果、改善建議與評估分數，這可讓受測者提升自主學習華語的意願。當受測者進入歷史紀錄頁面時，如圖三所示。受測者便可以知道此次練習的所有題目的測試結果與改善建議，幫助學生提升自身的華語能力。

圖一、AI 教學互動系統之輸入介面示意圖

圖二、AI 教學互動系統之歷史紀錄示意圖

3. 結論與未來研究

全球化發展，語言能力的重要性不減反增。隨著生成式 AI 技術的成熟，各式語言學習 AI 工具面世，並透過大量廣告強調 AI 對語言學習的高效性，然而，尚未嘗試的學習者往往無法確定廣告之真實性；或是當學習者在進行口說學習，並面對能力成效壓力時，亦無法適時運用符合所需的 AI 工具來輔助學習，達到期望的成效。本研究從實作精神出發，以兩階段之研究方式，實證 AI 口說教學互動系統輔助學習的成果，在口說成效方面，研究預期搭配 AI 口說教學互動系統提供即時性的回饋，有助於學生調整偏誤，提升句子準確率，強化其口說能力，具有實踐意義。此外，AI 輔助學習能減緩學習者口說焦慮方面，與先前的相關研究結果相符，能有效改善學習者的自信與流暢度，對學習動機與成效達到正向影響。而 AI 所提供的個人化回饋與適應學習機制，預期能進一步提升學習者長期學習成效，使學習者更能主動投入語言習得過程。透過本研究，可更全面理解 AI 工具輔助口說課程的學習成效影響，並作為未來 AI 融入語言學習應用的參考依據。

參考文獻

林佳臻（2017）。**探討口語練習回饋形式與先備知識於數位遊戲式英語學習環境對學習成效與口語焦慮之影響**。〔碩士論文，國立中央大學〕。https://hdl.handle.net/11296/h98t7h

張靜淑（2023）。人工智慧在國中小教學現場之運用。**台灣教育**，744，73-80。https://www.airitilibrary.com/Article/Detail?DocID=18166482-N202312190007-00008

潘冠霖（2018）。**華語線上互動中糾錯回饋與二語習得研究**。〔碩士論文，國立臺灣師範大學〕。https://hdl.handle.net/11296/sunt29

劉慧雯和蔡佩珊（2024）。以 AI 提升大學生英語口說表現。**臺灣教育評論月刊**，13(11)，68-74。https://www.airitilibrary.com/Article/Detail?DocID=P20130114001-N202411050015-00011

Chen, M.-R. A. (2024). Metacognitive mastery: Transformative learning in EFL through a generative AI chatbot fueled by metalinguistic guidance. *Educational Technology & Society, 27*(3), 407–427. https://www.jstor.org/stable/48787038

Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. *Perspectives on individual characteristics and foreign language education*, *6*(27), 27-49. https://www.degruyter.com/document/doi/10.1515/9781614510932/pdf?licenseType=restricted#page=37

Zou, B., Lyu, Q., Han, Y., Li, Z., & Zhang, W. (2023). Exploring students' acceptance of an artificial intelligence speech evaluation program for EFL speaking practice: an application of the Integrated Model of Technology Acceptance. *Computer Assisted Language Learning*, 1–26. https://doi.org/10.1080/09588221.2023.2278608

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

The Impact of Voice-Based Embodied English Chatbot on Learners' English Learning Motivation

Ying-Qi SHIH[1], Hui-Chun HUNG[2*], Shu-Fen HUANG[3]
[1,2]Graduate Institute of Network Learning Technology, National Central University, Taiwan
[3] National Central University, Taiwan
a038038456123@gmail.com, hch@cl.ncu.edu.tw, jessiehuang@gmail.com

Abstract: *This study explores artificial intelligence in language learning, focusing on the impact of embodied voice agents on English learning motivation. An interactive system with real-time chat functions was developed to support English practice. Participants were students from a university in northern Taiwan, completed pre- and post-tests of the English Learning Motivation (ELM) Scale. The experimental group, using an embodied chatbot, and the control group, using a non-embodied chatbot. The experimental group advanced in attitude toward learning English, desire to learn English and overall ELM, while the control group has no significant across all dimensions. Between-group differences were found only in desire to learn English, with no significant differences in other dimensions. These findings suggest embodied chatbots enhance motivation, particularly in attitude and desire, though the unique effect of embodiment requires further exploration.*

Keywords: English Learning Motivation, Embodied Chatbots, Generative Artificial Intelligence, Language Learning

1. Introduction

As the learning environment for English as a second language is gradually improved, many students still lack confidence and have low participation in real oral communication environments, indicating that there are still challenges in improving their willingness to express themselves orally and their confidence in communication (Su, 2021).

According to the research results of Mayer et al. (2003), the humanized characteristics of embodied teaching agents (such as natural voice and affinity) can effectively reduce cognitive load, allowing learners to devote more cognitive resources to in-depth processing of information, thereby improving learning outcomes.

However, research on the embodiment of such applications within English language teaching contexts remains relatively limited. Therefore, this study aims to develop a voice-based embodied English learning system and investigate: Does this system significantly impact university students' English learning motivation? It also seeks to provide new perspectives and approaches for English language teaching, fostering students' confidence and expressive abilities in spoken communication, while offering empirical evidence for future instructional design, ultimately enhancing students' overall proficiency in English.

2. Literature Review

2.1. Embodied Chatbots and Generative Artificial Intelligence (GenAI)

Embodied chatbots, incorporating humanized features like voice and visual cues, enhance learners' social presence, supporting interactive learning in language acquisition (Mennecke et al., 2010). Huang et al. (2022) highlighted chatbots' technological affordances (timeliness, personalization) and pedagogical uses (interlocutors, simulations) of chatbots, noting their ability to simulate natural dialogues and improve speaking proficiency through interactive practice. However, technological limitations and novelty effects may reduce long-term effectiveness. De la Vall and Araya (2023) emphasized AI language tools' personalized learning and cultural exposure benefits, yet highlighted their lack of embodied interaction, limiting social engagement in EFL settings. Pan et al. (2024) introduced ELLMA-T, an LLM-based embodied agent in social VR, enabling contextualized role-plays and dynamic feedback to enhance EFL speaking skills.

2.2. English Learning Motivation

Motivation is crucial for English as a Foreign Language (EFL) learning, shaping persistence and outcomes (Gardner, 1985). Gardner (1985) Attitude/Motivation Test Battery (AMTB) defines motivation via attitude toward learning English (ALE), motivational intensity (MI), and desire to learn English (DLE). Tai and Chen (2024) found that GenAI chatbots like CoolE Bot enhance engagement and sustained dialogue, boosting learners' motivation to interact. Zhang and Wang (2023) showed that motivation predicts academic performance, with low agreeableness amplifying this effect. Ebadi and Amini (2024) reported that embodied chatbots' social presence fosters enthusiasm and engagement in EFL learning.

3. Method

3.1. System Usage Scenarios

The voice-based embodied English chatbot system developed by the concept of instant messaging in chatrooms, allowing students to create different chatrooms to practice conversations on different topics. During practice, users can input their responses through either voice recording or text typing. Additionally, they can replay the chatbot's replies multiple times to improve their pronunciation. Upon completing a conversation topic, users can review past chat records to enhance their English communication skills.

By leveraging predefined rules configured within the OpenAI Assistants API, the chatbots can adopt different roles based on the learning objectives, thereby enhancing students' communication skills and interactive learning experience.

3.2. Participants and Experimental Design

This study adopted a quasi-experimental design, involving 66 students from two university English classes in northern Taiwan. One class (n=35) served as the control group (CG), and the other (n=31) as the experimental group (EG). All participants provided informed consent, with English proficiency levels ranging from A2 to B1 (CEFR). The EG interacted with an embodied voice chatbot, featuring facial expressions, real-time audio responses, and AI-driven replayable feedback. The CG interacted with a non-embodied chatbot, where the left side of the chat information panel did not feature an embodied display.

Preparations included teacher interviews and design of conversational themes and prompts. Both groups practiced dialogues in class and selected topics for extra practice outside class, as shown in *Figure 1*.

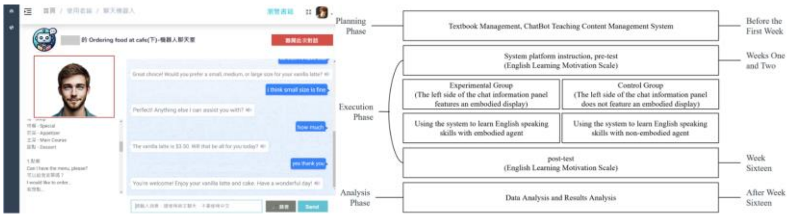


Figure 1. Embodied Chatbot Interface and Classroom Activity Design

3.3. Research Measurement

To assess students' English learning motivation, this study used an adapted version of Gardner (1985) AMTB, focusing on three dimensions: Attitude Toward Learning English (ALE), Motivational Intensity (MI), and Desire to Learn English (DLE). The scale was simplified to suit CEFR A2-B1 participants and time constraints, reducing ALE items from

10 to 4, MI from 6 to 5, and DLE from 6 to 5. A five-point Likert scale (1 = strongly disagree, 5 = strongly agree) was employed. The adapted scale assessed motivation differences between the experimental and control groups.

4. Results

For English learning motivation, EG showed significant improvements in ALE, DLE, and overall when compared to pre-test results, with *p*-values of .007, .020 and .012(p=.007<.01, p=.020, .012<.05), respectively, as shown in *Table 1*. The CG, on the other hand, showed no significant across all dimensions, as shown in *Table 2*. Subsequently, a Mann-Whitney U test was conducted to compare the EG and the CG, revealing significant superiority of the EG over the CG in DLE, with *p*-value of .031 (p=.031<.05), but no significant differences in other dimensions, as shown in *Table 3*.

Table 1. Wilcoxon signed-rank result of ELM in EG

Dimensions	Test	*N*	*M*	*SD*	*Z*	*p*
ALE	Pretest	31	15.61	3.333	2.708**	.007
	Posttest		17.58	2.363		
MI	Pretest	31	18.58	3.492	1.717	.086
	Posttest		20.13	3.233		
DLE	Pretest	31	19.26	3.794	2.324*	.020
	Posttest		21.23	3.471		
Overall	Pretest	31	53.45	10.188	2.522*	.012
	Posttest		58.94	8.406		

**p<.05, **p<.01, ***p<.001*

Table 2. Wilcoxon signed-rank result of ELM in CG

Dimensions	Test	*N*	*M*	*SD*	*Z*	*p*
ALE	Pretest	35	15.97	2.728	1.202	.229
	Posttest		18.51	3.745		
MI	Pretest	35	19.60	3.415	1.338	.181
	Posttest		54.09	8.942		
DLE	Pretest	35	16.57	2.062	.300	.764
	Posttest		19.31	3.169		
Overall	Pretest	35	19.34	2.869	1.185	.236
	Posttest		55.23	7.096		

Table 3. Results of the Mann-Whitney of the ELM

Dimensions	Groups	*N*	*Mean Rank*	*Sum of Ranks*	*Mann-Whitney U*	*Z*	*p*
ALE	EG	31	17.58	2.363	405.000	1.797	.072
	CG	35	18.51	3.745			
MI	EG	31	20.13	3.233	488.000	.706	.480
	CG	35	54.09	8.942			
DLE	EG	31	21.23	3.471	377.000	2.153*	.031
	CG	35	19.31	3.169			
Overall	EG	31	58.94	8.406	425.000	1.515	.130
	CG	35	55.23	7.096			

$*p<.05$, $**p<.01$, $***p<.001$

5. Conclusion

This study investigates the impact of embodied chatbots on English learning motivation compared to non-embodied chatbots, addressing a gap in comparative research. The embodied chatbot significantly enhanced students' desire to learn English and attitude toward learning English, demonstrating the potential of embodied technology in language education. By simulating natural interactions, embodied agents create engaging learning experiences that foster motivation. The control group, using a non-embodied chatbot, showed no significant improvement across all dimensions, highlighting the embodied design's advantage in motivation. There were between-group differences only in DLE, which suggests the unique benefits of embodiment in this dimension require more research, possibly due to the fact that the non-embodied chatbot is not as effective.

The absence of improvement in MI in EG, likely due to complex dialogues for beginner learners, subjective survey items reflecting cultural modesty, or limited study duration, represents a limitation. Future research should simplify interactions and adopt objective effort-based questions to strengthen MI and clarify embodied technology's benefits.

Acknowledgements

This research was supported by the National Science and Technology Council under the research project grant (Project No. NSTC113-2628-H-008-001-MY3 and NSTC113-2423-H-008-005).

References

De la Vall, R. R. F., & Araya, F. G. (2023). Exploring the benefits and challenges of AI-language learning tools. *International Journal of Social Sciences and Humanities Invention*, *10*(01), 7569-7576.

Ebadi, S., & Amini, A. (2024). Examining the roles of social presence and human-likeness on Iranian EFL learners' motivation using artificial intelligence technology: A case of CSIEC chatbot. *Interactive Learning Environments*, *32*(2), 655-673. https://doi.org/10.1080/10494820.2022.2096638

Gardner, R. C. (1985). Social psychology and second language learning: The role of attitudes and motivation. *Edward Arnold*, *205*.

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, *38*(1), 237-257.

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of educational Psychology*, *95*(2), 419. https://doi.org/10.1037/0022-0663.95.2.419

Mennecke, B. E., Triplett, J. L., Hassall, L. M., & Conde, Z. J. (2010). Embodied social presence theory. *2010 43rd Hawaii international conference on system sciences*, 1-10. https://doi.org/10.1109/HICSS.2010.179

Pan, M., Kitson, A., Wan, H., & Prpa, M. (2024). ELLMA-T: an Embodied LLM-agent for Supporting English Language Learning in Social VR. *arXiv preprint arXiv:2410.02406*. https://doi.org/10.48550/arXiv.2410.02406

Su, Y.-C. (2021). College students' oral communication strategy use, self-perceived English proficiency and confidence, and communication anxiety in Taiwan's EFL learning. *Educational Studies*, *57*(6), 650-669.

Tai, T.-Y., & Chen, H. H.-J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers & Education*, *220*, 105112.

Zhang, Y., & Wang, H. (2023). Effect of English learning motivation on academic performance among English majors in China: The moderating role of certain personality traits. *Psychology Research and Behavior Management*, 2187-2199. https://doi.org/10.2147/PRBM.S407486

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

AI or Not AI, That Is a Question

Meimei WANG[1], Chunlin LEI[2*], Wen SHEN[3]
[1,2,3]Shanghai University of International Business and Economics, Shanghai
leichunlin@suibe.edu.cn

Abstract: *This study was conducted in an English writing class at a university in the eastern part of China. Through questionnaires and interviews, it explored learners' usage of AI tools and their attitudes towards such usage in foreign language writing. The study found that students generally hold a positive attitude towards using AI to assist with foreign language writing, yet a small number of students still chose not to use AI in their writing. After conducting interviews with them, this study categorized main reasons for students' non-usage of AI, including concerns that relying on AI may waste the opportunity to practice and hone writing abilities, concerns that AI may cause rigid written patterns, distrust in the writing quality provided by AI, and considerations about integrity risks.*

Key Words: artificial intelligence; students' attitudes towards the use of AI; foreign language writing

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

人工智能：使用还是回避？

汪美美[1]，雷春林[2*]，沈文[3]
[1,2,3]上海对外经贸大学，上海
leichunlin@suibe.edu.cn

【摘要】本研究通过问卷和访谈，调查了我国东部某高校商务英语写作课的学习者对于生成式人工智能（AI）的使用情况及其使用态度。研究发现，学生普遍对 AI 技术辅助外语写作持积极态度，但有少数学生在写作过程中选择不使用 AI 工具，其原因包括担心浪费锻炼写作能力的机会、担心固化写作思维、不信任 AI 写作质量以及对诚信风险的考虑等。

【关键词】人工智能；使用态度；外语写作

1. 研究背景

近年来，人工智能语言处理模型迅速兴起，人机交互模型的涌现为教育行业带来了诸多机遇。例如，这些技术不仅能为外语学习者提供语言智能测评（郑春萍等，2024）、还能为使用者提供即时高效的反馈（魏爽、李璐瑶，2023）以及个性化的学习方案（郑春萍等，2024）。然而，AI 工具的普及也带来了诸多挑战，例如学术诚信问题、数据安全、学习依赖性等（马武林，2023）。面对这些潜在的风险，学界也在不断探索 AI 赋能外语教学的新路径。然而现有研究多聚焦于从教师角度出发的理论层面探讨(任伟等，2024；陈茉、吕明臣，2024)，对学生在外语学习中使用 AI 的认知关注稍显不足（徐林林等，2023）。事实上，语言学习者对于 AI 工具的使用态度在很大程度上影响着其在外语学习场景中的使用行为（王树胜、王俊菊，2024）。而当前有关研究更多地聚焦于探讨和分析使用 AI 工具的态度和策略，忽略了少部分不使用 AI 工具的群体。在此背景下，本研究从语言学习者的视角出发，具体探讨学生对于不选择使用 AI 工具的原因，希望为未来的 AI 课堂教学建设提供一定的参考。为此，本研究具体提出了以下两个问题：（1）在面对写作任务时，学生倾向于采用何种 AI 使用策略？（2）如果学生不选择 AI 工具来辅助其完成其课程作业，其具体原因是什么？

2. 研究设计

本研究以中国东部某高校开设的商务英语写作课程为背景。参与研究的是三年级学习商务英语写作课程的 50 名本科生。本研究采用行动研究范式，首先开展了学生在写作任务时使用 AI 工具的现状调查。具体而言，教师要求学生填写关于在推荐信的写作过程中借助 AI 辅助写作的情况说明表，该表设置了四个选项供学生选择，分别是：（1）我自己独立完成了本次写作任务；（2）我没有全盘抄袭人工智能提供版本；（3）我自己完成的部分占到了 70%；（4）我大幅使用人工智能进行写作和润色（比例高于 50%）。通过这几个选项，来了解学生在写作过程中，倾向于采取何种策略来使用 AI 工具。随后经教师对上述 50 份写作任务的梳理和统计，发现有 8 位同学选择“独立完成本次写作任务”。在人工智能普遍盛行的当下，学生不选择使用人工智能工具的原因是什么，反而是一个非常值得探究的问题。因此，在征得这 8 位同学的同意后，由助教对这 8 位同学分别进行了 12-15 分钟的半结构化访谈。访谈主要围绕他们对使用 AI 工具的态度来进行。访谈记录经过转写后，采用定性编码分析，来回答本研究的第二个问题。

3. 研究结果

通过梳理 50 名学生提交的作业和 AI 使用情况说明表，发现在写作任务中没有明确说明“必须使用”还是“禁止使用”的情况下，绝大多数学生（84%）对 AI 呈积极的使用态度，并会选择使用 AI 来辅助完成课业。但本研究中，有 8 名（16%）同学在完成课业的过程中，完全没有使用 AI 工具，本研究通过访谈法对其中的原因进行了探究。访谈记录经过转写后，笔者对数据进行了由上至下的梳理，之后再由下至上进行分析，从中归纳提炼出四个主题，用以解释学生选择不使用 AI 工具的原因。具体描述如表 1。

表 1：学生不使用 AI 工具的原因

主题	代表性描述
害怕陷入思维定势	“如果自己写有错误被纠正出来能记得更深一点。但如果用了 AI...很难跳出这个框架了(S1)”；“因为我怕我用了 AI 我的思维会被那个 AI 给改变,就是我的思维会跟着 AI 走（S8）”
担心弱化写作能力	“在这样的一个课程上，我去大幅度地使用 AI 来帮助我完成这个作业的话，就有点浪费这个机会。哪怕说我写的没有 AI 润色过的文章那么完美。但我能在后续修改的过程中能够更好地锻炼我的写作能力吧（S7）”
不信任 AI 的写作质量	“这个 recommendation letter 要结合自身的特殊的情况来写。我觉得 AI 不太能够写好（S5）”；“AI 写出来的东西就比较千篇一律吧（S6）”；“AI 有时候可能会夸大甚至是扭曲（S2）”
担心触碰诚信底线	“我就感觉 AI 写的，它会有很多这种长难句。语言比较正式感觉不像是我这种能力能写出来的。我就不太好意思去用（S3）”

这一结果表明，学生对于使用 AI 工具的负面影响是有一定认知的，但毕竟智能时代已来，我们不能因噎废食，要加强引导，使 AI 能真正赋能学生外语写作能力的培养。

参考文献

陈茉, & 吕明臣. (2024). ChatGPT 环境下的大学英语写作教学. **当代外语研究**, (1), 161-168

马武林, & 徐梦云. (2023). 21 世纪以来我国智能语言教学研究述评. **外国语文**, **39**(6), 140-150.

任伟,刘远博, & 解月. (2024). 英语写作教学中 ChatGPT 与教师反馈的对比研究. **外语教学理论与实践**, (4), 30-38, 60.

魏爽, & 李璐遥. (2023). 人工智能辅助二语写作反馈研究——以 ChatGPT 为例. **中国外语**, **20**(3), 33-40. https://doi.org/10.13564/j.cnki.issn.1672-9382.2023.03.007

王树胜, & 王俊菊. (2024). 动机驱动的 AI 赋能外语学习接受度研究——基于技术接受模型. **外语电化教学**, (6), 23-29, 108.

徐林林, 胡杰辉 & 苏扬. (2024). 人工智能辅助学术英语写作的学习者认知及行为研究. **外语界**, (3), 51-58.

郑春萍,于淼 & 郭智妍.(2024). 人工智能在语言教学中的应用研究:回顾与展望. **外语教学**, **45**(01), 59-68.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982-1003.

Li, K. (2023). Determinants of College Students' Actual Use of AI-Based Systems: An Extension of the Technology Acceptance Model. *Sustainability*, *15*(6), 5221.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Artificial Intelligence Education of Hong Kong Secondary Schools – Cultivating Ethical Values

Chi Wai Jimmy LAM
Buddhist Yip Kei Nam Memorial College, Hong Kong SAR
lam@byknmc.edu.hk

Abstract: *In recent years, artificial intelligence (AI) has developed rapidly. As global society continues to explore its positioning and applications, Hong Kong's secondary education system must establish a dedicated learning and teaching framework to suit the needs of students: "Technical Knowledge – Ethical Consideration – Social Responsibility". Beginning with the technical knowledge, students should understand AI's fundamental principles and the relationship between big data, data training and AI operations. Students should recognize that AI-generated answers and decisions may be flawed. The framework then systematically expands to ethical perspectives, using practical tools and real-world cases to demonstrate the complexity of ethical judgement in human societies. Finally, we must broaden students' perspectives to examine AI's societal impact and applications. Starting with legal issues it generates, enabling them to critically analyze the social implications and the respective responsibilities of different stakeholders.*

Keywords: Artificial Intelligence, Ethical Consideration, Values, Secondary Education

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

香港中學的人工智能教育 - 培育倫理價值觀

林志煒
佛教葉紀南紀念中學，香港
lam@byknmc.edu.hk

【摘要】近年，人工智能的發展一日千里，全球社會仍在摸索人工智能的定位和應用，香港中學教育極需要建立有關的學習和教學框架 -「技術認知 – 倫理批判 – 社會責任」。 由技術認知開始，透過明白人工智能的背後原理以及大數據訓練與人工智能的運作關係，讓學生明白人工智能的答案或決定並非絕對可靠，然後推展至倫理價值觀的視野，透過工具和新聞事件，讓學生了解人類社會的倫理判斷的複雜性。最後我們需要給予學生更廣闊的視野，看見人工智能在社會中的應用範疇，由它所衍生的法律問題開始，明白有關技術對社會的影響和不同持份者的有關責任。

【關鍵字】 人工智能；倫理；價值觀；中學學習框架

1. 前言

在這個瞬息萬變的時代，今天的學生除了傳統的知識外，還需要學習更多為未來而準備的知識、技能和態度。

人工智能的逐漸普及由多年前開始，當時人工智能集中應用於高新科技以及一些工商業研究和產品，民眾慢慢地從使用不同類型的產品中認為人工智能能為生活帶來正面的改變。然而，生成性人工智能企業 OpenAI 於 2022 年 11 月 30 日發表了旗下 ChatGPT 產品，震驚了全世界，也讓整個人工智能應用更迅速地滲透到每家每戶，人們紛紛嘗試和使用有關的產品，以作為協助和製作文書類別的工具。時至今日，它更能應用於製作圖片、影片以及以對話形式作出生成的部署。

近年中學至大學對於人工智能教育往往過分集中於如何以編程方式使用人工智能以協助設計產品，以及認識人工智能背後的數學和科學模型理論和方法等技術層面的教育。但是，以整個人工智能的生態圈而言，尤其以今天的青年而言，他們更需要掌握有關的倫理部份，在面對未來人工智能廣泛應用的社會，他們需要明白如何將之與社會中不同的產品、文化、法律、知識產權等不同的範疇融合，才可以更進一步地以自身的技能發展或善用人工智能。

2. 人工智能走入課堂

循序漸進的教學框架設計尤其重要，教師需要逐步帶領進行反覆地提問，進行高階思維的建構。人工智能在學生的層面上而言，尤以生成式人工智能，是既神秘又新奇的事物。他們普遍會由媒體報導、網紅人物的分享以及自己的經驗作出分析和推斷。普遍而言，中學生會認為人工智能已具備自主思考的能力。因此，整個教育框架需要包含以下環節：

1. 理解人工智能基礎技術
    a. 透過工具 (例如: Teachable Machine) 讓學生明白訓練模型與人工智能的關係
    b. 明白大數據的意義 (海量、速度、多樣性和價值)以及與人工智能的關係
    c. 人工智能誤判的真實個案
    d. 還原基本步，Garbage in, garbage out 的基礎電腦運算知識
2. 理解人工智能關於倫理的安排

a. 沒有工具能清楚了解黑箱中的關連網絡
b. 以情境建構故事，讓學生明白人類的倫理價值觀的複雜性和不確定性
c. 透過工具 (例如:Moral Machine) 讓學生透過情景，了解不同人士對同一件事情的看法，再推展出人工智能或無法滿足所有人的需要
d. 演算法的不同會帶來不同程度的偏頗

3. 認知人工智能如何影響社會
   a. 以新聞報導帶出不同種類由人工智能帶出的倫理問題 (例如:深度偽冒的造假，不論是文字、圖片、聲音、影片等)
   b. 以情景帶出由人工智能定奪的結論未必能滿足人類對倫理、道德的看法
   c. 人工智能對私隱的保護
   d. 人工智能對全球化法規上的衝擊
   e. 探討將來人類對人工智能工具和發展等規範的需要和方法

3. 框架下的教育

由於學生的學習是循序漸進的，他們需要先明白人工智能的行為是源於大數據給予的資料，以先進的演算法將之分析再融合，最後以神經網絡 (黑箱) 定義結果。通過不同的教育安排，學生會慢慢掌握到人工智能是可以犯錯的，也可以因為一些原因，作出偏頗或不合適的結論，尤其當有巨大的錯誤資訊傳入時，「指鹿為馬」的情況是可以由人為方法產生。

因此，學生會傾向較易理解人工智能的推論能力源自數據、演算法、使用者的回應以及其他不同類型的因素所造成。每個人對倫理和價值觀等也有不同的看法和理解，親疏亦有別。故此，學生在過程中明白有關的複雜性。面對人工智能作為通用的工具，他們需要理解其局限性，以及有關倫理之意義和安排。此外，作為今天的年青人，他們很有可能會在未來的日子從事有關業務，我們必須先讓他們建立起一個意識，在未來他們設計或選擇使用人工智能時，可以更能理解箇中情況和問題，作出更全面的考量、分析和判斷。

對於認知人工智能對社會的影響，學生在上述的安排下明白人工智能和倫理間的關係和相依性，認知它在社會應用可造成的影響，將能更進一步地培養他們對未來人工智能發展的看法、知識和法規層面的認識，從而加強他們對人工智能的應用和產品的辨識力，從而根據人工智能的利與弊作出有效的分析，將其優點發揮和發展。

參考文獻

Gouseti, A., James, F., Fallin, L., & Burden, K. (2024). The ethics of using AI in K-12 education: a systematic literature review. *Technology, Pedagogy and Education*, *34*(2), 161–182.

Fokas, A. S. (2023) Can artificial intelligence reach human thought? *PNAS Nexus*, *2*(12), pgad409.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Developing a Family AI Book Talk Companion: Leveraging AI to Improve Parent-Child Reading Discussion Quality

Yi-Cheng TSAI[1*], Chang-Yen LIAO[2], Tak-Wai CHAN[3]
[1,2,3] Graduate Institute of Network Learning Technology, National Central University, Taiwan
charlietsai11313@gmail.com

Abstract: *This study examines the effects of an AI-driven family book talk system on improving the efficiency and quality of parent-child reading interactions. Previous research shows that joint reading has positive impacts on children's reading abilities and family relationships. However, traditional reading practices can be hampered by communication barriers and external distractions, which lower participation willingness. In response, this study proposes a Family AI Book Talk Companion System grounded in a structured dialogue process, integrating large language models (LLMs) and dialogue generation techniques to boost interaction quality and engagement. The findings reveal that the system effectively enhances the structural organization and quality of reading discussions, while also increasing participants' sense of involvement and willingness to read as a family. Future research directions include strengthening AI-based facilitation strategies and diversifying feedback mechanisms.*

Keywords: AI Book Talk Companion, Family Book Talk, Book Talk Chatbot, Family Reading

1. Introduction

Chambers (1985) proposed the concept of "Book Talk," emphasizing the importance of expressing ideas and engaging in discussion. However, such practices are not common in parent-child reading contexts. While joint reading can enhance children's language development, cognitive skills, and emotional bonds (Bus et al., 1995), modern families often face time constraints that limit the time available for reading. Additionally, a lack of effective guidance, such as insufficient questioning techniques or the absence of scaffolding strategies, further reduces the frequency and quality of reading sessions (Bingham, 2007; Hindman & Morrison, 2012). Additionally, communication barriers and external distractions further impede meaningful engagement. To address these challenges, this study introduces a Family AI Book Talk Companion System that leverages natural language processing to structure and guide discussions, improving interaction quality and deepening parental involvement in children's reading experiences.

2. System Design and Application

This AI-based system aims to improve the quality of parent-child reading interactions through a structured approach. The system architecture comprises three main components: the user interface, the AI dialogue engine, and the data processing and analysis module. 1) User Interface: A web-based application accessible via tablets and computers, designed to minimize disruptions during reading activities. 2) AI Dialogue Engine: Integrates GPT-4, semantic understanding, and Web Speech API to support speech recognition and natural conversation. This reduces the need for manual text entry, thereby lowering barriers to parent-child interactions. 3) Data Processing and Analysis Module: Gathers interaction data and performs semantic analysis to generate feedback and reports for further use.

The reading interaction process is divided into three stages: pre-reading preparation, during-reading interaction, and post-reading discussion. In the preparation phase, the system suggests themes and offers scenario-based prompts to spark interest. During reading, it employs guided, open-ended, and goal-oriented questioning strategies, paired with narrative and associative prompts, to encourage deeper conversations. In the post-reading phase, it facilitates reflection and parent-child communication, enriching the overall reading experience. Acting as an interactive coordinator, the AI system helps reduce communication barriers, providing structured dialogue guidance for more meaningful discussions. Moreover, it records and analyzes these interactions to produce a "Parent-Child Book Talk Report," highlighting key insights and future activity suggestions to help families continually refine their reading quality.

3. Preliminary Results

A preliminary exploration was conducted to optimize the design and functionalities of the proposed system. Ten pairs of participants—with no prior system exposure—were recruited to simulate parent-child interactions, ensuring unbiased feedback.

In a single-session evaluation, 92.9% of participants indicated a willingness to continue using the system, citing its effectiveness in reducing communication barriers and improving interaction flow. In terms of overall satisfaction, 64.3% were highly satisfied, 28.6% were satisfied, and 7.1% were neutral. Regarding enhanced reading engagement, 35.7% strongly agreed that the system added enjoyment to reading, 50% agreed, and 14.3% observed no significant change, suggesting potential areas for increasing entertainment value.

Most participants reported improved structure and organization in book discussions, thanks to features like book recommendations, guided questions, and detailed interaction logs. Furthermore, 50% strongly agreed and 42.9% agreed that the system improved comprehension of the reading material, demonstrating its potential to deepen understanding.

Nonetheless, participants recommended further personalization and entertainment features, including richer voice interactions, more expressive feedback, and diverse questioning styles (e.g., humorous or creative prompts) to elevate the fun factor. Suggestions also included refining the book recommendation module to better match individual reading levels and interests. Future work should therefore enhance AI-driven personalization and entertainment elements, expand sample sizes, and conduct long-term usability testing to validate sustained effectiveness and stability. Overall, results indicate that the system offers a promising approach to improving structure, interaction quality, and engagement in family reading settings, laying a foundation for broader applications and future development.

Acknowledgements

This study was funded by the National Science and Technology Council of Taiwan (113-2410-H-008 -020 -) and by Research Center for Science and Technology for Learning, National Central University, Taiwan.

References

Bingham, G. E. (2007). Exploring the role of maternal literacy beliefs in mother–child joint book reading and home literacy environment. *Early Education and Development*, *18*(1), 23–49. https://doi.org/10.1080/03004430.2013.788813

Bus, A. G., van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Joint Book Reading Makes for Success in Learning to Read: A Meta-Analysis on Intergenerational Transmission of Literacy. *Review of Educational Research*, *65*(1), 1-21. https://doi.org/10.3102/00346543065001001

Chambers, A. (1985). Booktalk: Occasional writing on literature and children. *Harpercollins Childrens Books*.

Hindman, A. H., & Morrison, F. J. (2012). Family involvement and educator outreach in Head Start: Nature, extent, and contributions to early literacy skills. *The Elementary School Journal*, *112*(3), 359–386. https://doi.org/10.1086/657651

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Fostering Computational Thinking and AI Literacy Through Cross-Disciplinary STEAM Education: A Case Study of the "AI Chameleon" Project for P.6 Students

Chi Yan WONG [1*], Wai Lung MUNG[2], Wai Lam CHU[3]
[1,2,3]Fung Kai No.1 Primary School, Hong Kong SAR
astomwong@gmail.com

Abstract: *This study examines the effects of a cross-disciplinary AI curriculum on computational thinking and artificial intelligence literacy development among 101 fifth-grade students in Hong Kong. The instructional design incorporated STEAM elements with formative assessment strategies, emphasizing understanding of AI technologies and responsible usage attitudes. Pre-test measurements indicated students demonstrated significant proficiency in stepwise problem-solving ($M=3.85$, $p<.01$) and identifying key programming elements ($M=3.79$, $p<.01$), while exhibiting comparative weaknesses in flowchart creation ($M=3.42$, $p<.05$) and problem decomposition conceptualization ($M=3.45$, $p<.05$). Post-implementation assessments revealed statistically significant improvements in problem decomposition (8.4%, $p<.01$), algorithm design (7.6%, $p<.01$), and systems thinking (6.9%, $p<.05$). Qualitative analysis showed 65% of participants could articulate specific limitations of AI recognition technologies, compared to 48% prior to instruction. These findings suggest cross-disciplinary AI education effectively enhances primary students' computational thinking abilities and AI literacy.*

Keywords: AI Literacy, Self-Directed Learning, Six-step STEM Pedagogy, Computational Thinking, STEAM Education

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

結合 AI 技術與跨學科 STEAM 課程培養學生運算思維能力與人工智慧素養：以五年級「人工智慧奇異變色龍」課題為例

黃智仁[1*]，蒙韋綸[2]，朱偉林[3]
[1,2,3]鳳溪第一小學，中國香港
astomwcy@fk1ps.edu.com

【摘要】本研究探討跨學科AI課程結合形成性評估對101位香港小五學生的影響。前測顯示學生在「解決程式問題步驟」(M=3.85, p<.01)和「發現問題關鍵」(M=3.79, p<.01)較強，「繪製解決流程」(M=3.42, p<.05)和「思考問題拆解」(M=3.45, p<.05)較弱。課程後，問題分解(8.4%, p<.01)、演算法設計(7.6%, p<.01)及系統性思考(6.9%, p<.05)能力顯著提升，能說明AI辨識技術限制的學生比例從48%增至65%。研究證實此教育方式有效促進小學生運算思維與AI素養發展。

【關鍵字】 人工智慧素養；自主學習；STEM 六步曲教學法；運算思維；STEAM 教育

1. 前言

在數位科技快速發展的時代，培養小學生的AI素養與運算思維已成教育重點。Black與Wiliam(2009)指出，形成性評估是促進學生自我調節學習的關鍵教學理念。Wing(2006)提出的運算思維被視為21世紀核心素養，Grover與Pea(2013)強調應將其融入K-12課程。本研究針對五年級學生設計「人工智能奇異變色龍」跨學科課程，探討如何通過STEAM教育與形成性評估策略培養學生的運算思維與AI素養。

2. 課程設計理念

本課程基於科技藝術融合、形成性評估與跨學科整合三大理念。通過「奇異變色龍」設計，將AI圖像辨識與生物學、美術、編程知識結合，創造有意義的學習情境。課程遵循Black與Wiliam(2009)的形成性評估框架，強調明確學習目標和成功標準，重視課堂對話、提問和即時回饋。這種「協助學習的評估」通過教師引導、同儕互評及自我反思，幫助學生改善作品、深化AI理解並發展自我調節能力，呼應Dewey(1938)的體驗式學習理論。

3. 課程目標與學習成果

課程目標分三個維度：知識（理解AI圖像辨識原理及限制）、技能（培養運算思維核心能力）和態度（發展對AI技術的批判思考及負責任使用意識）。預期學習成果是學生能設計製作創意「奇異變色龍」作品，並解釋其AI原理與功能設計邏輯。這些目標與Black與Wiliam(2009)的形成性評估理念一致，旨在促進學科概念深度理解及發展自我評估和調節能力。

4. 學與教策略

課程採用Black與Wiliam(2009)的形成性評估框架，包含五大策略：闡明分享學習目標與成功標準、設計有效課堂討論獲取學習證據、提供促進學習的回饋、激發學生成為彼此學習資源、促使學生成為學習主人(Wiliam & Thompson, 2007)。教學過程始於引導性提問與示範建立共識，再透過小組實作應用所學。各環節設置回饋點（教師觀察、同儕互評、自我反思），幫助學生優化作品。此設計符合「教師、

同儕和學習者」三方評估角色，促進學生從被動接受評估轉為主動參與學習調控。

5. 課程內容與實踐

研究對象是於香港就讀小學的101位五年級學生（男54女47，10-11歲）。學生用平板結合Makecode設計「奇異變色龍」，經歷探索、設計、開發及測試改良四階段。每階段配合形成性評估活動（討論、分享、互評、自評），符合Black與Wiliam「評估循環」，並採取「成功標準共建」策略建立明確評估指標。

6. 課程結果與分析

6.1. 運算思維能力的變化

本研究採用混合研究法，結合定量與定性數據分析。定量數據通過前後測問卷收集，問卷採用李克特五點量表(1=非常不同意，5=非常同意)，評估學生在運算思維和AI素養方面的自我認知變化。定性數據則通過學生作品分析、課堂觀察記錄及半結構化訪談獲取。

6.2. 課程學習成效

在課程學習成效方面，學生對「工程思維能力有所提升」(M=3.71, p<.01)的認同度最高，其次為「自主學習能力有所提升」(M=3.68, p<.01)。「守法意識有所提升」(M=3.62, p<.05)雖然相對較低，但整體而言仍呈現正面評價。

6.3. 教學反思與建議

研究顯示，結合形成性評估的跨學科AI教育能有效提升小學生運算思維與AI素養。建議教育工作者：

1. 設立明確學習標準：依Black與Wiliam(2009)「成功標準」概念促進自我監控。
2. 建構多元回饋系統：結合教師指導、同儕評價及自我反思，培養建設性回饋能力。
3. 給予充分修正時間：強調評估「用於改進」而非僅作結果判斷。
4. 關注不同起點學生需求：根據評估信息提供適切鷹架支持，實現「適應性教學」。

參考文獻

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31. https://doi.org/10.1007/s11092-008-9068-5

Dewey, J. (1938). *Experience and education*. Kappa Delta Pi.

Grover, S., & Pea, R. (2013). Computational thinking in K-12: A review of the state of the field. *Educational Researcher, 42*(1), 38-43. https://doi.org/10.3102/0013189X12463051

Kong, S. C. (2024). Cross-disciplinary STEM learning for Asian primary students: Design, practices, and outcomes. In W. W. M. So, Z. H. Wan, & T. Luo (Eds.), *Cross-disciplinary STEM learning for Asian primary students* (pp. 147-163). Routledge.

Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners* (2nd ed.). ASCD.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53-82). Erlbaum.

Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33-35. https://doi.org/10.1145/1118178.1118215

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Harnessing Retrieved Augmented Generation for Data-Driven Learning

Eric Yu Kin CHAN
Hong Kong Baptist University Affiliated School Wong Kam Fai Secondary and Primary School, Hong Kong SAR
eykchan@hkbuas.edu.hk

Abstract: *This paper explores the application of Retrieved Augmented Generation (RAG) in enhancing data-driven learning among Grade 8 students. By integrating advanced AI tools like GPT, the study aims to elevate students' cognitive abilities from lower-order thinking skills to higher-order thinking skills, as outlined in Bloom's Taxonomy. Through project-based learning, students utilize AI to address social issues, culminating in the development of chatbots that incorporate guided documents for more accurate responses. The findings indicate a positive shift in students' perception and effective use of AI in education, highlighting its potential to transform traditional learning methodologies.*

Keywords: Retrieved Augmented Generation, Data-Driven Learning, Bloom's Taxonomy, AI in Education

1. Introduction

Using pretrained Large Language Models (LLMs) offers several opportunities in education. They can assist students in gaining insights into their problems, thereby enhancing students' creativity during project-based learning.

During the process of basic prompting with LLMs, users receive simple feedback. According to Bloom's Taxonomy – Cognitive Domain, this ranges from lower-order thinking skills such as remembering, to higher-order thinking skills like creating. This suggests that students may bypass some stages necessary for developing higher-order thinking skills when they rely on results generated by AI. This hierarchical structure reflects the increasing complexity and depth of cognitive processes involved in learning. Lower-order thinking skills involve basic recall and comprehension, while higher-order thinking skills require more sophisticated cognitive abilities, including critical thinking, problem-solving, and the ability to synthesize and generate new ideas. (Anderson, 2001)

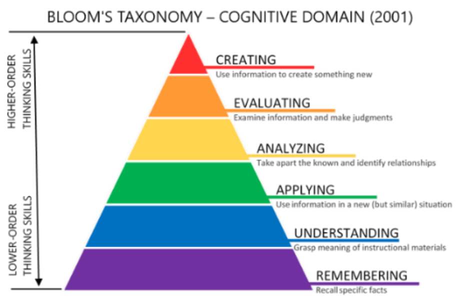


Figure 1. Bloom's Taxonomy – Cognitive Domain

This study aims to enhance student cognition level through Retrieved Augmented Generation. The lessons were designed to help students apply, analyze, and evaluate data to improve their thinking skills. Students engaged in project-based learning using AI for social good by creating chatbot.

2. Experience of using GPT in Grade 8 students

To understand the use of GPT among grade 8 students, a survey was conducted at the beginning of the school year. A total of 187 grade 8 students completed the survey. The results showed that 83% of students had used GPT before the start of the school year, and 66% of students used GPT for learning. However, some students provided feedback on why they do not use GPT for learning. They stated that using GPT feels like cheating and does not help with learning new things. It is hoped that, with the adoption of new teaching methods, students will recognize the potential benefits of AI in education and its positive impact on society.

3. From Prompt Engineering to Retrieved Augmented Generation

Grade 8 students may lack knowledge about effective practices for using GPT through well-constructed prompts. They might use simple ideas in their prompts. The lesson was dedicated to building experience in creating good prompts. Elements such as persona, objective, audience, context, and boundaries were introduced and practiced using GPT.

A project titled “AI for Social Good” has been assigned as a group task to be completed within this school year. Students are expected to identify a topic of personal relevance related to this issue. They will utilize GPT to explore potential solutions for the chosen topic. This approach aims to provide students with insights that may assist in developing viable solutions.

Subsequently, students have endeavored to address their identified issues by developing a chatbot. This chatbot does not rely solely on user prompts; instead, students are required to provide guided documents as references for GPT to generate responses. Figure 2 illustrates an online development tool designed to assist students in building a chatbot using Retrieval Augmented Generation. To enhance the chatbot's performance, students may add additional nodes to incorporate reference documents or include logical elements to refine the predicted outcomes.

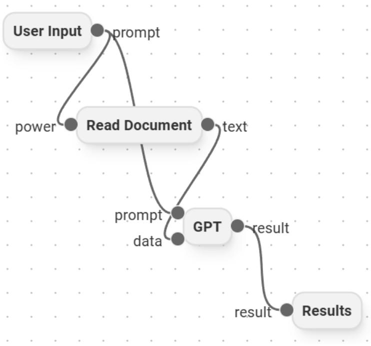


Figure 2. Basic Retrieved Augmented Generation Structure

4. Evaluating Students' Learning Outcomes

The topic of "AI for social good" is broad. Students created various chatbots on diverse topics, leading to multiple outcomes. They shared their ideas and solutions using sources they found, aided by RAG to present them in the chatbot. Students were inspired by each other's presentations. Project completeness is one of the assessment criteria. Teachers evaluated their learning process using students’ submitted reference resources, AI flow design, GEN AI adoption, chatbot design, and teamwork.

5. Conclusion

Through this learning process, students must apply information in new situations and analyze and evaluate the performance of the chatbot they have built. This approach assists students in progressing from lower-order thinking to higher-order thinking. Students achieved recognition for their participation in generating results by evaluating predicted outcomes through their contributions (their reference documents). This participation can enhance students' learning experiences by exposing them to new concepts and knowledge.

References

Anderson, L., & Krathwohl, D. A. (2001). *Taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Learning for Good: Exploring Co-Creative Interdisciplinary Learning in the Era of Generative AI

Ka Ki Spike HO[1*], Tse Shan Suzanne SUEN[2*]

[1,2] Buddhist Sum Heung Lam Memorial College, Hong Kong SAR

[3] Association of I.T. Leaders in Education (AiTLE), Hong Kong SAR

spikeho0083@gmail.com, kkho@bsc.edu.hk

Abstract: *With the rapid advancement of generative artificial intelligence (AI) technologies, our school has responded to the Hong Kong Education Bureau's initiative to promote STEAM and AI education by designing and implementing an innovative interdisciplinary program titled "Co-creative Learning." This program aims to enhance students' linguistic thinking and creativity. Beginning with basic prompt engineering, students progressively master image generation techniques using Stable Diffusion and extend their learning to creative activities that integrate Chinese culture and writing. The curriculum emphasizes the transformation from technological understanding to creative application, adopting inquiry-based and project-based learning approaches to foster students' problem-solving skills. In the featured activity "If I Were in an AI Parallel Universe," students combine AI-generated images with expressive writing to portray their emotions and imagined worlds through personal associations and creative thinking.*

The program also encourages students to reflect on the ethical implications of AI in cultural preservation and creative practices, fostering positive values and a responsible attitude toward emerging technologies. This prepares students for future learning and life in an AI-driven society.

Keywords: Generative AI; Prompt Engineering; STEAM Education; Interdisciplinary Learning; Stable Diffusion

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

探索生成式人工智能時代 如何驅動跨學習領域的同創共學

何嘉琪[1*]，孫芷珊[2*]
[1,2]佛教沈香林紀念中學，香港
[3]資訊科技教育領袖協會 (AiTLE) ，香港
spikeho0083@gmail.com，kkho@bsc.edu.hk，tssuen@bsc.edu.hk

【摘要】 隨著生成式人工智能（AI）技術迅速發展，本校設計並實施一個名為「同創共學」的校本課程，提升學生創意思維。由基礎的提示工程開始，逐步掌握圖像生成技術，並延伸至結合中華文化與寫作的創作活動。課程設計強調由「科技理解」到「創意實踐」的轉化，並以探究學習與項目式學習為主軸，培養學生的解難能力。在創作活動「假如我在 AI 的平行時空」中，學生生成圖像與寫作，透過個人創意描繪其情感與想像世界。課程亦讓學生反思 AI 在文化傳承、創作倫理中的角色，建立正向價值觀與負責任的科技態度，為未來 AI 社會的學習與生活打好基礎。

【關鍵字】 生成式 AI；提示工程；STEAM 教育；跨學科學習；Stable Diffusion

1. 階段一:「我是 AI 提示工程師」學習與 AI 溝通 提升語文思維

隨著科技迅速發展，生成式 AI 被廣泛應用。本課程透過教授與 AI 互動的技巧，裝備學生面對未來，並培養其辨識 AI 產出資訊真偽的能力，進而強調資訊的倫理運用與媒體素養的培養。課堂將引導學生學習撰寫準確提示詞，作為與 AI 溝通的基礎，以獲得更具針對性的回應，這也為第二階段的「AI 生成圖像課程」奠定核心能力。學生需掌握與 AI 互動的基本策略，重點在於撰寫有效提示（prompt），並透過追問與優化技巧，提升回應品質。同時，課程將引導學生認識 AI 的運作原理，並意識到運用生成式 AI 獲取知識時需具備查證意識。生成式 AI 回應的品質與學生語文能力密切相關，而提示詞的設計正是跨領域語文素養的實踐場域。

2. 階段二:「我是 AI 咒語繪畫師」學習 AI 生成圖像技能

當學生掌握了與 AI 生成文本的經驗後，第二階段學習生成式 AI 圖像創作與設計平台 Stable Diffusion(SD)，它是一個進階的開放源碼 AI 生成圖像平台，通過教授 AI 生成式圖像技術，培養他們的創新思維，為未來的職業生涯做好準備。

2.1. 理解 AI 生成圖像的特性

在課堂中，學生需輸入一則簡單指示「Hong Kong」並生成圖像，接著比較自身與同儕的作品，思考：「同一句提示詞是否會產出完全相同的圖像？」此活動旨在讓學生理解，為何多數生成圖像平台設計為同一提示可產出多個不同結果。學生觀察到，即使輸入相同語句，AI 每次產出圖像皆有所不同，這是生成式 AI 獨有的隨機性設計。因此，若僅依賴單一句提示，難以精準控制輸出結果。透過此活動，學生認知到學習進階生成技巧的重要性，能有效降低生成的不確定性，提升圖像輸出精準度。

2.2. 讓學生明白不是懂得英文提示便懂得 AI 生成圖像

僅懂提示詞不足以完全掌握圖像輸出；在教師引導下，學生學習如 CFG、Seed 等 SD 進階參數，透過探究方式逐步掌握生成式 AI 的關鍵控制技術。當學生能靈活運用這些技巧後，將進行主題式解難活動，結合中英文跨學科創作與寫作任務：「假如我在 AI 的平行時空」。

3. 階段三:「假如我在 AI 的平行時空」圖像生成配合中英文寫作進行創作

此階段是一項解難活動，運用逆向工程的學習策略。學生不再輸入提示詞生成圖像，而是由教師先展示數張現實場景照片，引導學生細緻觀察與想像，並從中選擇最具個人感受與聯想潛力的一張作為創作藍本。學生運用 AI 進行風格模仿與創意擴展，生成屬於自己心中「AI 平行時空」的圖像作品。此作品需呈現其內在想像與情感投射，並完成一篇短文，內容須結合實景描寫、聯想意象與感情色彩。整體作品將採用進展性評估，全面評量學生在中文與英文提示詞撰寫表現，以及科技層面中 AI 進階參數的應用能力。此活動融合觀察力、創造力、AI 應用技巧與語文表達能力，除提升學習動機外，亦拓展學生敘事深度與寫作內涵。

圖 1 學生創作出描繪理想社區作品，成功串聯中英文與科技科，展現跨學科整合成果

圖 2 其中以屯門社區及輕鐵為題的作品，更獲港鐵公司邀請，在港鐵博物館中展出，肯定學生的創意表現及學以致用的能力

教育的本質在於促進學生的整全發展，而不僅是追求學術成績。在科技迅速演進的時代，我們更需反思教育的真正使命：不只是傳授知識與技能，更是培養學生的共通能力與正向價值。因此，課程設計應著重實踐性與可轉化性，使學生能將所學應用於真實世界情境中，發展出終身受用的能力，這些核心素養將不會被科技所取代。創新教育的目標，是協助學生在瞬息萬變的世界中建立綜合能力、創新思維、正確價值觀與解難能力，進而為社會與國家未來的創新發展貢獻力量。"當你堅持投入鑽研一件事，它將逐漸成為你的藝術"

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Minecraft Metaverse: Transforming STEAM Education Through Interactive, Game-Based Learning for Enhancing Student Engagement and Skill Acquisition

Manpreet SINGH[1]
The Education University of Hong Kong, Hong Kong SAR
manpreet@eduhk.hk

Abstract: *This research examines the use of Minecraft Education Edition as an immersive Metaverse platform for creating student-centered classrooms where undergraduate students engage in game-based learning for STEAM Education. During a three-hour lesson, students navigate a virtual world, completing various tasks and quests through their digital avatars, which fosters an interactive and engaging learning environment. Instructor also participate as an avatar, enhancing collaboration and facilitating real-time interactions. To monitor student progress and provide timely interventions, students are required to capture screenshots of their completed activities and share them on a visual collaboration platform called Miro. This approach not only documents student achievements but also encourages accountability and reflection. Furthermore, student engagement is assessed by observing avatar movements within the virtual space, offering insights into participation levels. Findings indicate that Minecraft Education Edition effectively serves as a Metaverse by providing an immersive, interactive environment that promotes active learning and collaboration. This study highlights the potential of game-based learning in higher education, suggesting that such platforms can enhance student engagement, facilitate skill acquisition, and transform traditional teaching methods. The implications for future educational practices and the integration of immersive technologies in curriculum design are also discussed.*

Keywords: Metaverse, Minecraft Education Edition, Game-based Learning, STEAM Education

1. Introduction

The traditional teaching methods in STEAM Education (Science, Technology, Engineering, Arts, and Mathematics) often rely on passive learning approaches (Park and Kim, 2022), which may not effectively engage students and promote active learning. The integration of immersive technologies, such as virtual reality (VR) and augmented reality (AR), has the potential to create interactive and engaging learning environments that promote student-centered learning. Minecraft Education Edition, a specialized version of the game with additional features tailored for STEM subjects, has been shown to enhance learning, engagement, and collaboration among secondary students through project-based learning activities (Callaghan, 2016) is a popular game-based learning platform that has been used in the present work to implement STEAM education.

2. Methodology

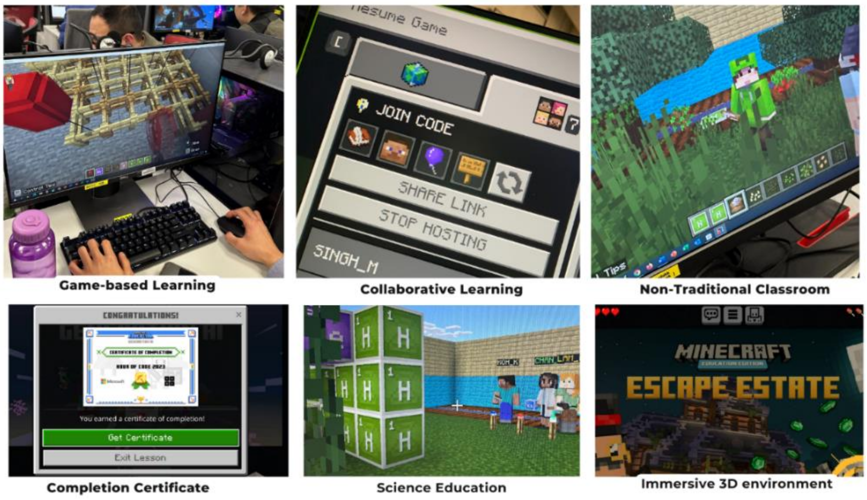
This study used a qualitative approach to examine the use of Minecraft Education Edition as a Metaverse platform for STEAM education. A three-hour lesson was designed for undergraduate students, where they navigated a virtual world, completing various tasks and quests through their digital avatars. The instructor also participated as an avatar, enhancing collaboration and facilitating real-time interactions. Students were required to capture screenshots of their completed activities and share them on a visual collaboration platform called Miro. This approach not only documented student achievements but also encouraged accountability and reflection.

3. Results

The findings of this study indicate that MEE effectively serves as a Metaverse by providing an immersive, interactive environment that promotes active learning and collaboration. Students demonstrated high levels of engagement and participation, with their avatar movements indicating a high level of activity within the virtual space. The use of MEE also facilitated skill acquisition, with students demonstrating a deeper understanding of complex concepts and principles.

4. Discussion

This study highlights the potential of game-based learning in higher education, suggesting that such platforms can enhance student engagement, facilitate skill acquisition, and transform traditional teaching methods. The findings of this study have implications for future educational practices and the integration of immersive technologies in curriculum design. Educators can use MEE as a Metaverse platform to create student-centered classrooms that promote active learning, collaboration, and engagement.



5. Conclusion

In conclusion, this study demonstrates the potential of Minecraft Education Edition as a Metaverse platform for STEAM education in agreement with Singh and Sun (2025). The findings of this study highlight the importance of game-based learning in higher education and suggest that such platforms can transform traditional teaching methods. Educators can use Minecraft Education Edition to create interactive and engaging learning environments that promote student-centered learning and facilitate skill acquisition. These findings also agree with Caldas et al., 2024, which show that this Minecraft Education Edition is able to promote positive academic engagement and achievement among students.

References

Callaghan, N. (2016). Investigating the role of Minecraft in educational learning environments. *Educational Media International*, *53*(4), 244–260.

Caldas, O. I., Mauledoux, M., Aviles, O. F., & Rodriguez-Guerrero, C. (2024). Breaking presence in immersive virtual reality toward behavioral and emotional engagement. *Computer Methods and Programs in Biomedicine, 248*, 108124.

Park, S., & Kim, S. (2022). Identifying world types to deliver gameful experiences for sustainable learning in the metaverse. *Sustainability*, *14*(3), 1361.

Singh, M., & Sun, D. (2025). Evaluating Minecraft as a game-based metaverse platform: Exploring gaming experience, social presence, and STEM outcomes. *Interactive Learning Environments*. Advance online publication.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Practical Experience of Artificial Intelligence + Digital Education in Integrating Learning and Teaching in Primary Schools

Ka Man WAT[1], Oi Yuen CHENG[2], Long Yuen YU[3*], Ching Long NGAN[4]
[1,2,3,4]King's College Old Boys' Association Primary School, Hong Kong SAR
yuly15@kcobaps1.edu.hk, watkm15@kcobaps1.edu.hk, chengoy@kcobaps1.edu.hk, ngancl16@kcobaps1.edu.hk

Abstract: *Since 2016, our school has been promoting the programmed curriculum of digital education and planning school-based Artificial Intelligence (AI) education in 2018 to cope with the needs of technology education and social development. This paper explores the application of AI+ in digital education with the goal of enhancing students' technological literacy, creativity, and responsible attitudes towards the use of technology. The integration of AI technology can significantly enhance learning effectiveness and motivation.*

Keywords: Artificial Intelligence, Digital Education, Learning and Teaching

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

人工智慧+數字教育在小學融入學與教的實踐經驗

屈嘉曼[1]，鄭愛遠[2]，余朗源[3*]，顏正朗[4]
[1,2,3,4]英皇書院同學會小學，中國香港
yuly15@kcobaps1.edu.hk, watkm15@kcobaps1.edu.hk, chengoy@kcobaps1.edu.hk, ngancl16@kcobaps1.edu.hk

【摘要】 自 2016 年起，本校開始推展數字教育的編程課程，並於 2018 年規劃校本人工智慧（AI）教育，以應對科技教育及社會發展的需求。本文探討 AI+在數字教育中的應用，目標在於提升學生的技術素養、創造力及負責任的科技使用態度。AI 技術的整合可顯著提升學習效果與學習動機。

【關鍵字】 人工智慧；數字教育；人工智能；學與教

1. 發展方向

本校 AI 教育發展可歸納為以下幾個方向：

學生能掌握編程和 AI 的運用，培養他們對技術的負責任態度與明辨性思維，使學生能應用、學用，甚至善用及靈活運用 AI。

- 編程與應用設計：學生運用簡單的編程語言編寫程式，設計解決實際問題的應用，並以製作模型或辨識系統進行創意發明和設計。
- AI 輔助教學：於學習、教學及評估三方面，提升學習動機及學習效能。
- AI 素養與道德教育，培養學生對科技的負責任態度與明辨性思維。

2. 教學實踐

2.1 學用人工智慧

在 AI 課程設計中，學生透過編寫模型或識別系統學習科學和數學，並利用 AI 編程及組裝工具解決實際問題。以下是幾個實例：

2.1.1 科學科 – 製作人工智能識別系統

利用 Machine Learning for Kids (文字辨識) (Scratch 3)製作「脊椎動物類別辨識系統」，透過辨識不同類別動物的特徵強化知識。本項目旨在利用機器學習技術，幫助學生了解脊椎動物的分類。系統將基於學生所上傳的圖像進行分析，並將圖像分類為動物五大類別。而學生應用 Machine Learning for Kids 的文字辨識，學習通過數據訓練模型進行動物分類，學生再利用 Scratch 3 編程語言來構建該系統。（圖 1）

圖 1 Machine Learning for Kids 範例

2.1.2 跨科專題研習

學校亦在跨科專題研習（圖 2）中引入人工智慧，藉「人口變化：長者護老大行動」課題，透過深入研究長者的生活需求，學生們設計和發明切合長者需求的實用物品。在這個過程中，

學生結合設計思維和運算思維，運用 AI 套件進行編程，應用文字、圖形和語音辨識技術來創造創新產品。他們還利用 AI 繪圖生成器，製作海報、簡報和多媒體材料。

圖2 長者護老大行動專題研習範例

2.2 應用人工智能

隨着人工智慧工具發展成熟，學校發展人工智慧+，即學生學習人工智慧外，更以 AI 技術深度與科目融合和創新應用，於學習、教學及評估三方面，提升學習動機及學習效能

2.2.1 對學生學習而言

o 個性化學習 -AI 可以分析學生的學習習慣和成績，並提供量身定制的學習計劃和資源，幫助學生以自己的節奏學習。

o 即時反饋 - AI 可以隨時回答學生的問題，提供即時反饋， 並協助解決學習中的困難。

o 語言學習 - AI 語言學習應用程序提供語音識別和即時糾正，幫助學生 提高語言能力。

2.2.2 對教師教學而言

o 自動化評分 - AI 可以自動評分作業和評估，節省教師的時間， 使他們能夠專注於分析學生在學習上的強弱項和學習難點，並提供適時的回饋。

o 學習分析 - 利用數據分析，AI 可以幫助教師了解學生的學習進度，識別問題區域，並根據數據調整教學內容及策略。

o 內容生成 - AI 可以創建教學內容，包括課程材料、 測驗和練習題，幫助教師減輕備課負擔。

2.2.3 對評估而言

o 促進學習的評估：有系統和便捷的模式收集學生學習成果的顯證，讓教師更容易分析學生在學習上的強弱項和學習難點，並提供適時的回饋

o 作為學習的評估：學生在學習過程中學習檢視自己或同儕的學習效能，反思和調整學習策略，以發展他們自主學習的能力

3. 結語

我們探討了如何將人工智慧與數字教育結合，以提升小學生的學習體驗。學校一直推展數字教育的編程課程及校本人工智慧（AI）教育，以應對科技教育及社會發展的需求。透過這些實踐經驗，我們期待能夠培養出對科技有負責任態度的學生，並促進他們的創造力和技術素養。未來，我們將繼續探索 AI 在教育中的潛力，為學生創造更豐富的學習環境。

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

附錄

製作人工智能識別系統教學範例

AI 語言學習應用程序範例

AI語言學習應用程序範例

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Promoting Technological Innovation Education in K-12: Cultivating Artificial Intelligence, Metaverse, and Innovative Thinking

Fu Yuen CHUNG[1*], Po Chi IU[2*]
[1,2]Concordia Lutheran School, Hong Kong SAR
pe13hk@gmail.com, pciu@cls.edu.hk

Abstract: *This paper explores how to integrate artificial intelligence and metaverse education at the K-12 education stage to cultivate students' innovative thinking. Based on Piaget's cognitive development theory and Kolb's experiential learning theory, combined with Papert's constructionism and design thinking, an interdisciplinary innovation curriculum framework is established. Through three main implementation strategies: "Artificial Intelligence Education," "Mixed Reality Learning," and "Computational Thinking Curriculum," along with the "Think-Creative-Day" outcome display and business-school collaboration model, students can apply technology to solve problems in real-life situations. This blended educational model not only cultivates students' abilities to face the future but also encourages them to actively participate in innovative practices, preparing a competitive new generation of talents for the digital age.*

Keywords: Artificial Intelligence Education; Metaverse Education; Design Thinking; Experiential Learning; Business-School Collaboration

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

在 K-12 推動科技創新教育：K-12 人工智能、元宇宙與創新思維培育

鍾富源 [1*]，姚普致 [2*]
[1,2] 路德會協同中學，香港
pe13hk@gmail.com, pciu@cls.edu.hk

【摘要】 本文探討 K-12 教育階段如何整合人工智能與元宇宙教育，培養學生的創新思維。以 Piaget 認知發展 理論和 Kolb 經驗學習理論為基礎，結合 Papert 構建主義與設計思維，建立跨學科創新課程框架。通過「人工 智能教育」、「虛實結合學習」和「計算思維課程」三大實施策略，配合「想·創·日」成果展示與商校合作模式，讓學生在真實情境中應用科技解決問題。這種融合教育模式不僅培養學生應對未來的能力，更讓他們主動參與創新實踐，為數位時代培育具競爭力的新一代人才。

【關鍵字】 人工智能教育 ；元宇宙教育 ；設計思維 ；經驗式學習 ；商校合作 5

1. 前言

在數位化迅速發展的當代，人工智能(AI)、元宇宙(Metaverse)與創新設計思維正重塑社會各個層面，包括教育生態系統。面對急速變化的社會環境，教育工作者必須重新思考如何裝備學生應對未來的挑戰。在香港教育局的課程指引「學會學習 2.0+」提出「開拓與創新精神」的重要性，強調學生需具備創新能力、主動精神及抗逆力，即使面對不確定性，仍能保持堅毅並把握機遇。本文探討在 K-12 教育階段如何透過跨學科課程與實踐式學習，整合人工智能 與元宇宙教育，培養學生的創新思維和二十一世紀所需技能。

2. 理論基礎與教育範式轉變

2.1. 認知發展與經驗學習

根據 Piaget 的認知發展理論，K-12 學生正處於具體運算到形式運算的關鍵過渡期。人工 智能教育能提供抽象思維的具體載體，而元宇宙則創造了沉浸式學習環境，使抽象概念具象 化。與此同時，柯爾布(Kolb, 1984)的經驗學習理論強調，有效學習應形成一個完整循環：具體經驗、反思性觀察、抽象概念化和主動實踐。這兩種理論在科技創新教育中的結合表現為 學生通過 AI 與元宇宙技術獲得具體經驗，引導反思科技應用過程中的問題與挑戰，連結設計 思維與 AI 概念進行抽象思考，最後在真實或虛擬環境中實踐解決方案。

2.2. 構建主義與設計思維的融合

Papert 的構建主義(Constructionism)強調學習者需通過創造有意義的作品來建構知識。設計思維(Razzouk & Shute, 2012)作為一種系統性解決問題的方法，與構建主義高度契合。 在科技創新教育中，這種融合表現為五步教學流程：同理心(empathize)、定義問題(define)、腦震盪(ideate)、原型設計(prototype)和測試(test)。當這一流程應用於 AI 與元宇宙教育時，能有 效培養學生的創造力、批判性思考與問題解決能力。

3. 人工智能與元宇宙融合教育具體實施

3.1. 初中全面推動「人工智能(Artificial intelligence)」教育

由 2019 年開始，學校已加入人工智能教育課程，在 2022 年更被香港中文大學獲入選為「中 大賽馬會「智」為未來計劃——領袖學校」十間之一。學校在初中及高中加入人工智能元素， 並使用設計思維及了解社區作切入點。其中一個課業是學生分組透過網上人工智能

使用 NLP 及 GAN 技術，輸入相關完整字句，產生藝術作品，讓師生及同儕互評。將來更會與藝術科進 行跨科合作。

3.2. 虛擬與現實結合的學習模式

元宇宙教育應重視虛實結合，避免與現實脫節。有效策略包括將現實世界問題帶入元宇宙， 讓學生在安全環境中尋求解決方案；鼓勵學生將元宇宙中習得的技能應用於現實情境；以及 將沉浸式體驗與實體製作相結合，在虛擬設計後進行實體原型製作。在我校的七十週年校慶 中，我們設計了「重建舊校舍」的學習單元讓學生創建原本校舍的虛擬導覽，「文物保育與 科技」活動則結合文化保育與科技應用，都是將現實問題帶入虛擬環境的範例。

3.3 計算思維課程的整合

計算思維作為 AI 與元宇宙教育的基礎，可採用四步曲教學流程。首先是示範與體驗階段，使用完成的 AI 應用或元宇宙場景；其次是分析與思考，理解底層原理與運作邏輯；然後是編程與創建，實際編寫程式或構建元宇宙環境；最後是反思與改進，評估作品限制並思考改進方向。

3.4. 創新的實踐與展示模式

本校舉辦成果展示的「想·創·日」：與社會參與為激勵學生並強化學習成效。「想·創·日」作 為一種創新展示模式，邀請教育工作者、科技企業專家和社區代表組成跨界評審團，學生向 真實觀眾展示 AI 與元宇宙作品，並根據評審意見現場調整方案。這種公開展示的機會讓學生 跳出舒適區，培養演講技巧和應變能力。實踐表明，許多平日表現被動的學生，在此類活動 中往往展現出驚人的積極性和創造力。

3.5. 創新的商校合作模式

為強化學習與實際應用的連結，我校引入了創新的商校合作模式。學生可為真實企業解決技 術或業務挑戰，與社會企業合作解決社區問題，甚至將優秀項目轉化為初創企業。過往合作 案例包括設計食品安全人工智能 App、開發老人智能輔助設備等。透過這些合作，學生不僅 應用所學知識，更體驗到為社會創造價值的成就感。

4. 結論與未來展望

K-12 階段推動人工智能、元宇宙與創新設計教育的融合，不僅是應對技術變革的必然選擇， 更是培養學生未來核心競爭力的關鍵途徑。通過構建系統性課程架構、發展教師專業能力、 解決實施挑戰，我們能夠為學生創造真正面向未來的學習體驗。未來發展方向應關注打造更 完整的 K-12 科技創新教育生態系統，開發 AI 與元宇宙技術的教育專用版本，建立國際標準 化的科技創新教育評估框架，以及加強全球教育工作者在這一領域的協作與經驗分享。在這 場教育變革中，教育工作者需秉持開放創新精神，同時堅守教育核心價值，確保技術服務於 育人目標，而非反之。唯有如此，我們才能培養出能夠在數位時代中不僅適應變化，更能主 導創新的新一代。

參考文獻

Razzouk, R., & Shute, V. (2012). What is design thinking and why is it important? *Review of Educational Research*, *82*(3), 330–348. https://doi.org/10.3102/0034654312457429

Kolb, D. A. (1984). *Experiential learning*. Prentice-Hall.

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

Using Metaverse Cospace to Design a Virtual Church Tour

Ka Yin CHAN

Sacred Heart Canossian School, Hong Kong SAR

c_chan@shcs.edu.hk

Abstract: *The widespread adoption and evolution of virtual reality (VR) technology has significantly expanded opportunities for the application of technology in education. Traditionally, virtual education courses have focused on "field trips," allowing students to experience virtual representations of locations that were previously only accessible through flat images. Teachers have used canned VR content to conduct virtual lessons for students. The author designed a course combining VR technology with programming skills, enabling students to understand VR technology and learn how to create virtual reality environments. This course encourages students to design and produce interactive processes themselves, transforming them into Makers—actively building, executing, and repeatedly refining their plans while learning, rather than passively experiencing someone else's creations (like viewing canned videos). By applying VR technology, the course broadens students' learning horizons. This particular project not only allows students to create virtual environments but also helps them understand Hong Kong's historic architectural landmarks—such as Catholic churches. Through a first-person perspective, students introduce the unique features of these churches. Each student adds their own interactive elements to the virtual environment, showcasing their individual creativity.*

Keywords: Virtual environment, programming, computational thinking, creativity, church

Kong, S. C., Lan, Y. J., Zhao, J. H., Song, Y. J., & Mitrovic, T. (Eds.). (2025). *Proceedings of the 3rd International Conference on Metaverse and Artificial Companions in Education and Society*. Hong Kong: The Asia-Pacific Society for Computers in Education.

透過Metaverse Cospace以虛擬環境創造聖堂導覽

陳嘉妍
嘉諾撒聖心學校，香港
c_chan@shcs.edu.hk

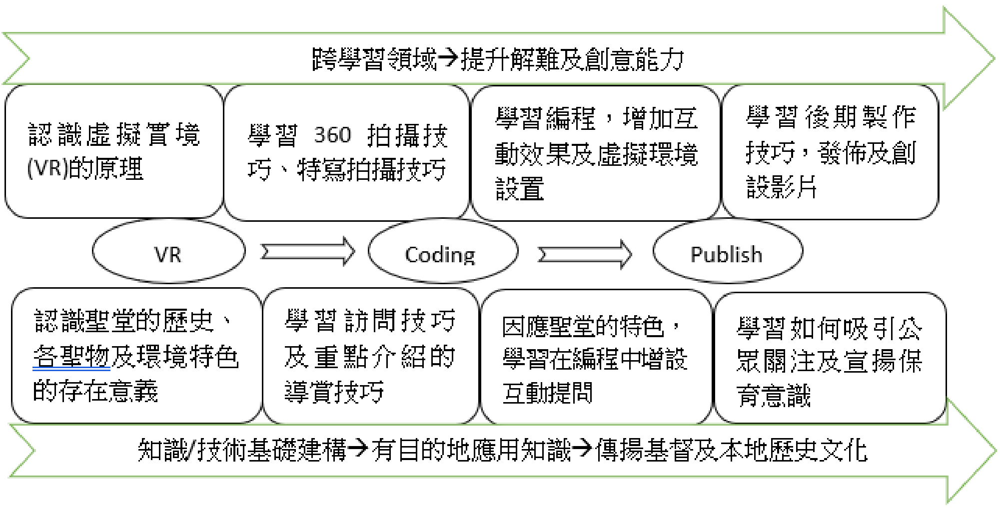
【摘要】虛擬實境技術的普及與衍生，讓教育應用科技的機會就更遼闊。以往虛擬教育課程內容集中在「戶外教學」（field trips）領域，透過虛擬實境技術，使原本只能透過平面圖像體會的景點更加逼真，教師應用現有罐頭式的VR素材來為學生進行虛擬教學。筆者設計了應用虛擬技術結合編程技術的課程，讓學生了解VR技術，並真切了解如何製作虛擬實境，讓自己規劃及製作互動過程，讓學生成為創客(Maker)主動建立、執行、反覆修正自己的計畫並從中學習，而非只被動閱讀別人的體驗(察看罐頭式的影片)，將應用虛擬實景的技術來拓寬學生的學習領域。是次計劃既能讓學生創造虛擬環境外，更進一步讓學生了解香港的歷史建築物-天主教聖堂，以第一身視角，向各人介紹聖堂特色。每個學生都在這個虛擬環境中加入獨有的互動，呈現各學生獨有的創造力。
【關鍵字】虛擬環境；編程；運算思維；創造力；聖堂

1.前言-計劃緣起

本校為一所天主教學校，秉持嘉諾撒會祖辦學重要理念-讓人愛慕及認識耶穌基督。宗教培育對學生的心靈發展尤為重要。聖堂是天主教教會的重要寶地，是天主與信友們的交流地，蘊藏著濃厚的宗教價值與信息。本校學生會在聖堂參與彌撒聖祭，即使是非教友，也會在天主教禮儀節令的四旬期到聖堂以拜苦路作祈禱。

早年由於疫情問題，未能進行大型的聚集活動。學生便未能親見聖堂內部，進行宗教活動。「嘉諾撒仁愛女修會聖堂」乃香港一級歷史保護文物，富有豐富的歷史價值，內裡的聖像、聖物及玻璃窗都富有特色及有深層意義。故此，筆者便設計「VR 導覽聖堂-培訓小小導賞員」的課程。透過重現虛擬實景，讓學生更深入地認識天主教聖堂的特色，並以創新科技的方式(VR+編程)，讓學生創設獨有的聖堂介紹。是次計劃既能訓練學生思考如何介紹聖堂，成為小小導賞員，亦能提升學生對基督徒價值觀的身份認同，亦能拓寬學生掌握科技的創意，讓學生們學會思考如何將科技應用於生活上以達致更美好的效果，保留香港歷史古蹟的文化。

1.1. 實施計劃



圖一： 整個活動流程概念圖

學生會先習得聖堂歷史、建築特色，再觀身參觀聖堂，利用360鏡頭及攝影機拍攝影像，製造虛擬的聖堂環境，其後與修女訪談再進一步了解聖堂特色，再運用COSPACES這個虛擬環境加入編程，將聖堂的特色以自己創造的方式作互動介紹，最後發佈及邀請嘉賓參與虛擬的場景，帶上VR 眼鏡隨時隨地去了解香港的歷史古蹟-「嘉諾撒仁愛女修會聖堂」及「聖母無原罪主教座堂。」

1.2. 學習成果

透過是次計劃，進行虛擬實境學習，可以得到以下的學習效果：

● 在沉浸式虛擬校園中學習與交流

透過虛擬實境，重現整個建築物的環境及特色，讓學生隨時隨地可在虛擬聖堂中學習，例如：拜苦路，而不受環境天氣影響。另可促進跨地域學習，可邀請國內姐妹學校的同學一同感受聖堂的氣氛，甚至進一步可邀請國外同學一起參與創建聖堂的互動，讓同學了解天主教聖堂的文化。

● 互動及個人化的學習

透過虛擬環境的學習，學生可以在安全及自由度大的學習空間嘗試發揮創意，可以透過虛擬角色的參與、發問問題和進行互動。她們可以發揮創意，在真實的虛擬環境中，加設人物或物件，例如：蠟燭、跪拜的同學等，並就自身對聖堂的認知，透過重建再造，利用已有知識和經驗來建構新知識的過程，當將自身認知表達出來時，這是一種積極的建構活動，可以幫助更深刻地理解學習內容。通過表達和互動，能更清楚地組織自己的想法，並進一步鞏固知識。透過編程，可製作更多不同的有趣互動，鼓勵她們探索多種的解決方案，從而提升創造力和創新精神。

● 體驗式學習及即時反饋

學生通過模擬真實情景來"體驗" 學習,令她回想起在聖堂的禮儀,更專注於當中的氣氛,專注地細看每一件聖物及聖像。如在真實環境中，她們可能礙於高度只能仰望，現在能從不同角度去觀察，放大注視。編程的互動更可提供即時反饋，除了與受眾的互動外，亦可更快地糾正錯誤，讓人正確認識聖堂內的知識。

● 歷史傳承與記載

透過是次計劃的訪談、觀察及修女們的講解，這歷史建築物的真實史料得以保存，用影像真實地還原在這虛擬世界中，學生傳承著舊有聖堂文化，又與時並進加點新意。這些都是虛擬世界另闢出來的，以另一面貌還原其珍貴獨有的史料。

2.結語

虛擬實境（VR）技術的應用確實為教育帶來了革命性的變革。從單純觀察罐頭片段-「戶外教學」到「創客」的轉變，是一個非常有前瞻性的教育模式。透過讓學生親自參與、設計和執行VR項目，不僅能讓他們更好地理解VR技術，還能培養他們的創造力和問題解決能力。透過這種方式，學生可以從被動的觀察者轉變為積極的創造者，這也符合現代教育強調的自主學習和實踐精神。與此同時，他們也能夠學習到編程等技術技能，這將對他們未來的學習和職業發展產生深遠的影響。此外，這種創客教育模式還能激發學生的興趣和好奇心，讓他們更加投入到學習過程中。實際上，這種虛擬實境和編程技術的結合，正是未來教育的一大趨勢。

香港教育大學
The Education University of Hong Kong

Artificial Intelligence and Digital Competency Education Centre
人工智能及數碼能力教育中心

Co-organiser:

MetaACES 2025

Website URL:

https://www.eduhk.hk/metaaces2025/

Email:

metaaces2025@eduhk.hk

Supporting Organisations:

Education Bureau
The Government of the Hong Kong Special Administrative Region of the People's Republic of China

AiTLE

香港電腦教育學會
The Hong Kong Association for Computer Education