THE INTERNATIONAL REVIEW
OF RESEARCH IN OPEN AND
DISTRIBUTED LEARNING

Vol 25
No 3

August 2024

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

Editorial – Volume 25, Issue 3

Special Issue: Artificial Intelligence in Open and Distributed Learning: Does It Facilitate or Hinder Teaching and Learning?

Ehsan Namaziandost[1,*] and Afsheen Rezai[2]
[1]Department of General Courses, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran & Department of English *Language Teaching, Ahvaz Branch*, *Islamic*, *Azad University*, *Ahvaz*, *Iran; [2]Teaching English and Linguistics Department, University of Ayatollah Ozma Borujerdi, Borujerd City, Iran; *Corresponding author*

Artificial intelligence (AI) is a rapidly evolving field with the potential to revolutionize various aspects of education, especially in open and distributed learning, including distance education, hybrid learning, and blended learning. AI can transform curriculum design, content delivery, assessment, feedback, learner support, and learning analytics (Chen et al., 2020). AI offers personalized and adaptive learning paths based on learners' preferences, needs, goals, and performance, enhancing their educational experience (Holmes et al., 2023). It also provides timely feedback and guidance, fostering engagement and motivation. AI creates interactive and immersive learning environments, such as games, simulations, and virtual reality, sparking learners' interest and involvement. It promotes social and collaborative learning by facilitating communication and cooperation among learners, instructors, and resources (Holmes et al., 2023).

Additionally, AI optimizes various tasks like content creation, grading, assessment, and learning analytics, improving the quality and efficiency of education processes. However, AI in education also raises significant challenges and risks. Ethical, legal, social, pedagogical, and technical issues need consideration (Chen et al., 2022; Ouyang & Jiao, 2021). For example, maintaining AI applications' quality, validity, reliability, and fairness remains crucial. Protecting learners' and instructors' privacy, security, and autonomy in AI-mediated learning contexts is also vital (Holmes et al., 2023).

Moreover, fostering critical thinking, creativity, and human values in AI-enhanced learning experiences is crucial. Lastly, addressing the digital divide and potential marginalization of learners and instructors without access to or skills in AI technologies is paramount (Holmes et al., 2023). Overall, AI has immense potential in education, but its responsible and informed implementation is necessary to ensure its benefits are maximized while mitigating potential risks.

Given the above points, this special issue aims to explore the opportunities and challenges of AI in open and distributed learning, including distance education, hybrid learning, and blended learning, from multiple perspectives. We invited original research articles that address the following topics but were not limited to:

- theoretical and conceptual frameworks for understanding and evaluating AI in open and distributed learning;
- empirical studies on designing, developing, implementing, and evaluating AI applications like ChatGPT in open and distributed learning;
- best practices and case studies on integrating AI into open and distributed learning curricula, pedagogies, and policies;
- critical analyses and reflections on AI's ethical, legal, social, pedagogical, and technical implications in open and distributed learning; and
- future trends for AI in open and distributed learning.

This special issue features 20 papers, each peer reviewed by at least two experts in the field.

The first article, "AI application (ChatGPT) and Saudi Arabian primary school students' autonomy in online classes: Exploring students and teachers' perceptions" by **Ali Rashed Ibraheam Almohesh,** investigated the impact of an AI-powered application, namely ChatGPT, on the autonomy of Saudi Arabian primary students participating in online classes. The research used a mixed-method design involving 250 Saudi Arabian primary students from six primary schools in Riyadh, Saudi Arabia. The findings revealed that ChatGPT significantly affected the participants' perceptions of autonomy and its dimensions. The results also indicated that AI-powered applications contributed to the students' autonomy in 10 different ways. Participants also mentioned that AI-powered apps might have some negative consequences.

The second paper, entitled "Threats and opportunities of students' use of AI-integrated technology (ChatGPT) in online higher education: Saudi Arabian educational technologists' perspectives," written by **Prof. Mesfer Mihmas Mesfer Aldawsari** and **Dr. Nouf Rashed Ibrahim Almohish** uncovered the perspectives of 20 educational technologists from four Saudi Arabian universities regarding the integration of AI-powered technology, particularly ChatGPT, into online higher education. The study adopted a qualitative research method that relied on the principles of theoretical sampling to select participants and conducted in-depth interviews to collect their insights. Twenty Saudi Arabian educational technologists volunteered to take part in the research. The results uncovered a rich range of insights into the challenges and opportunities associated with students using AI-integrated technology in online higher education. Additionally, eight threats highlighted concerns about data security, privacy, and potential risks associated with AI technology in educational institutions.

The third article, by **Dan Wang**, is titled "Teacher- versus AI-generated (Poe application) corrective feedback and the language learners' writing anxiety, complexity, fluency, and accuracy." This study investigated the effects of corrective feedback (CF) on language learners' writing anxiety, writing complexity, fluency, and accuracy and compares the effectiveness of feedback from human teachers with an AI-driven application called Poe. Using a quasi-experimental design with pretest and posttest measures involving 75 participants, the results revealed the significant effects of teacher and AI-generated feedback on learners' writing anxiety, accuracy, and fluency. Interestingly, the group that received AI-generated feedback performed better than the group that received teacher feedback or no AI support.

The fourth article is by **Hong Duan** and **Wei Zhao**, titled "The effects of educational artificial intelligence-powered applications on teachers' perceived autonomy, professional development for online teaching, and digital burnout." They explored the repercussions of AI-empowered technologies on teachers' autonomous behavior, digital burnout, and professional development. Using a sample of 320 high school teachers in China, the results indicated a discernible positive impact of AI-integrated technology intervention on teachers' professional development and autonomous behaviors. Incorporating AI-enhanced tools aided teachers' professional growth and bolstered their independent and self-directed instructional practices. Moreover, the study revealed that using AI-integrated technology significantly reduced teachers' susceptibility to digital burnout, signifying a potential alleviation of stressors associated with technology-mediated teaching.

The fifth article by **Ting Xiao, Sisi Yi,** and **Shamim Akhter** disclosed the interplay between self-esteem (S-E), cognitive-emotion regulation (CER), academic enjoyment (AE), and language success (LS) in artificial intelligence (AI)-supported online language learning. Three hundred eighty-nine English as a Foreign Language (EFL) learners in China participated. The results highlighted the vital function of online courses assisted by AI in enhancing students' CER and AE learning. This implies that students with a robust sense of self-efficacy can effectively regulate their cognitive and affective processes in AI-supported language learning.

In article 6, **Zhiqun Ouyang, Yujun Jiang,** and **Huying Liu** checked how language learners' willingness and engagement to communicate in English as a Foreign Language (EFL) classrooms are affected by Duolingo. The study was conducted on 80 first-year language learners from the Foreign Language Department of Hunan International Economics University in China. Using a quasi-experimental method, the findings confirmed the effects on learner engagement, which showed significant gains in affective, cognitive, and behavioral domains, indicating Duolingo's beneficial impact on engagement in general. Furthermore, the results confirmed Duolingo's contribution to improved language attitudes, engagement, and communicative confidence.

Next, in the seventh paper entitled "The auxiliary role of artificial intelligence applications in mitigating the linguistic, psychological, and educational challenges of teaching and learning Chinese language by non-Chinese students," **Jingfang Xia, Yao Ge, Zijun Shen,** and **Mudasir Rahman Najar** delved into the auxiliary role of AI-empowered applications in mitigating the educational, linguistic, and psychological challenges which none-Chinese learners face while learning Chinese/Mandarin language. Qualitative research was employed, and 20 Chinese language teachers were selected through theoretical sampling. The findings revealed that AI-empowered educational applications help language learners overcome the commonly reported educational, psychological, and linguistic challenges that non-Chinese learners and Mandarin teachers might encounter. Findings verify the effectiveness of AI-empowered applications like ChatGPT, Poe, Brainly, etc, in helping teachers and learners of the Chinese language learn grammar, structure, idioms, and cultural issues of the Chinese language.

Dongmin Ma, Human Akram, and **Hua Chen** in the eighth paper, checked the potential variations across Chinese and international students (from diverse countries across the world) in terms of attitudes (AU) and their behavioral intentions (BI) towards AI use. The data were collected from 689 valid cases from diverse schools of a Chinese University through a survey approach employing questionnaires. The results showed a substantial discrepancy between Chinese and international students' prevalence, attitudes, and behavioral intentions towards AI use. Findings further revealed a more robust Perceived Ease of Use

(PEOU) effect on AU and BI among international students compared to their Chinese counterparts. The findings also suggested that cultural backgrounds and prior technological exposure play intricate roles in shaping perceptions of AI technology.

Using a mixed-methods study based on the technology acceptance model (TAM), **Hanwei Wu, Yunsong Wang,** and **Yongliang Wang** investigated the determinants of behavioral intention to use AI among 464 Chinese EFL college learners in the ninth article. The results showed that perceived ease of use significantly and positively predicted perceived usefulness and attitude toward AI. Moreover, attitude toward AI significantly and positively predicted behavioral intention to use AI. However, contrary to the TAM assumptions, perceived usefulness did not significantly predict either attitude toward AI or behavioral intention to use AI. In addition, mediation analyses suggested that perceived ease of use significantly and positively impacted students' behavioral intention to use AI through their attitude toward AI rather than through perceived usefulness. Lastly, semi-structured interviews with 15 learners provided a nuanced understanding of the statistical patterns.

In paper 10, **Jingyu Xiao**, **Goudarz Alibakhshi, Alireza Zamanpour, Mohamad Amin Zarei, Shapour Sherafat,** and **Seyyed-Fouad Behzadpoor** investigated the structural relationship among Iranian undergraduate students' AI literacy, academic well-being, and educational attainment. Through a convenience sampling approach, 400 undergraduate students from virtual universities equipped with LMS platforms and facilities were selected. The results demonstrated that the hypothetical model enjoyed acceptable psychometrics (divergent and convergent validity, internal consistency, and composite reliability). Results also showed that the general model had goodness of fit. The direct effect of AI on academic well-being and educational attainment was confirmed. Moreover, findings also indicated that AI literacy in India significantly affects educational attainment by measuring variables of academic well-being.

In paper 11, **Sha Gao** checked the effect of AI applications on enhancing undergraduate students' academic emotions and test anxiety. Using a convenience sampling approach, data were collected from 160 undergraduate students majoring in different fields of study who were divided into control and experimental groups. The findings showed that using AI-empowered applications significantly reduced the students' test anxiety and negative academic emotions but enhanced the students' positive academic emotions. Students can use ChatGPT as an auxiliary instrument to overcome negative emotions and improve their educational attainment.

In article 12, **Gürhan Durak, Serkan Çankaya, Damla Özdemir,** and **Seda Can** aimed to present a comprehensive bibliometric analysis of 1726 academic studies from among those indexed by the Web of Science database between 2013 and 2023 to provide a general framework for the concept of artificial intelligence in education (AIEd). Several bibliometric analysis techniques were applied, and the motivations behind each analysis's execution and method of producing findings were documented. The findings showed that the number of studies on AIEd has increased significantly, with the USA and China being the most common countries of origin. Institutions in the USA stand out from those around the world. Pioneering journals in education have also emerged as prominent in AIEd. On the other hand, collaboration between authors is limited. The study was supplemented with keyword analysis to reveal thematic AIEd concepts and to reflect changing trends.

In the 13th paper, **Tahereh Heydarnejad** and **Fidel Çakmak** investigated the links between teacher immunity (TI), work passion (WP), job satisfaction (JS), occupational success (OS), and psychological well-being (PW-B) in the context of AI-assisted online linguistic learning. Three hundred eighty-nine Iranian teachers of English as a Foreign Language (EFL) were given the Language Teacher Immunity Instrument, the Work Passion Scale, the Job Satisfaction Questionnaire, the Occupational Well-Being Scale, and the Psychological Well-Being at Work Scale. The findings emphasized the crucial role that TI and WP play in providing a balance in their JS, OS, and PW-B while applying AI in their language instruction.

Article 14 was the effort of **Ying He.** This study intended to picture the effects of employing automated writing evaluation (AWE) in fostering learners' writing skills, motivation to write, enjoyment in writing, and academic buoyancy in open and distributed English as a foreign language (EFL) learning. Eighty-six intermediate EFL students from China took part in this research. The participants in the experimental group (n=44) receive instruction and feedback only from their teachers, while the control group (n=42) is exposed to their teachers' instruction and AWE. The data analysis results indicated that the experimental group participants outperformed their peers in the control group in learners' motivation to write, enjoyment in writing, academic buoyancy, and academic success in writing.

In the 15th paper, **Ferdi Çelik, Ceylan Yangın Ersanlı,** and **Goshnag Arslanbay** investigated the impact of ChatGPT-simplified authentic texts on university students' reading comprehension, inferencing, and reading anxiety levels. One hundred five undergraduate EFL students engaged in original and ChatGPT-simplified text readings, serving as their controls. The findings revealed a significant improvement in reading comprehension and inferencing scores following ChatGPT intervention. However, no significant change in reading anxiety levels was observed. The study suggests that ChatGPT-simplification positively influences reading comprehension and inferencing, but its impact on reading anxiety remains inconclusive.

The 16th article by **Fatih Karataş and Erkan Yüce** investigated the experiences and perceptions of 141 preservice teachers engaged with AI, mainly through ChatGPT, over a three-week implementation on Zoom to understand its influence on their evolving professional identities and instructional methodologies. Employing Strauss and Corbin's methodological approach of open, axial, and selective coding to analyze reflective narratives, the study unveils significant themes that underscore the dual nature of AI in education. Key findings revealed ChatGPT's role in enhancing educational effectiveness and accessibility while raising ethical concerns regarding academic integrity and balanced usage. Specifically, ChatGPT was found to empower personalized learning and streamline procedures, yet challenges involving information accuracy and data security remained.

In article 17, **His-Hsun Yang** proposed a hypothetical model combining the Unified Theory of Acceptance and Use of Technology (UTAUT) with Self-Determination Theory (SDT) to explore design professionals' behavioral intentions toward using AI tools. Surveying design professionals in regions influenced by Confucian culture, using Chinese speaking, and analyzing 565 valid cases with AMOS supported the structural model hypothesis. The model explains 52.1% of the variance in behavioral intention to use (BIU), proving its effectiveness in explaining these variances. The results further validate the greater importance of performance expectancy (PE) over effort expectancy (EE) in influencing BIU. Additionally, it has been shown that the impact on intrinsic motivation (IM) and extrinsic motivation (EM) can either be amplified or diminished by anxiety about JR. For individuals experiencing higher levels of JR anxiety, there is a marked increase in IM. They may perceive adopting AI tools as an opportunity to enhance their skills and job security. Conversely, this anxiety also significantly boosts EM, as the potential for improved efficiency and productivity with AI use becomes a compelling incentive.

The 18th article by **Selay Arkün-Kocadere** and **Şeyma Çağlar Özhan** compared the impact of human and AI-generated instructors in video lectures on video engagement and academic performance. In addition, they examined students' opinions on both types of videos. A convergent-parallel approach mixed method was used in this study. A total of 108 undergraduate students, 48 in the experimental group, 52 in the control group, and 8 in the focus group interview participated. The findings of the experimental part revealed that learners' video engagement was higher in the course with the human instructor compared to the course with the AI-generated instructor. However, the instructor type did not significantly affect academic performance. The results based on the qualitative part showed that students thought the AI-generated instructor caused distraction, discomfort, and disconnectedness. However, when the video lesson topic is exciting, or students focus on the video with the intention of learning, these situations can be ignored. In conclusion, even in today's conditions, there is no difference in performance between human and AI-generated instructors.

The 19th article by **Odiel Estrada-Molina, Juanjo Mena,** and **Alexander López-Padrón** intended to determine the trends, the applied computational techniques, and the areas of educational use of deep learning in open learning through a systematic review. Among the main results, it is worth noting that the scientific literature focuses on the following areas: (1) predicting student dropout, (2) automatic grading of short answers, and (3) recommendation of MOOC courses. It was concluded that pedagogical challenges include the effective personalization of content for different learning styles and the need to address possible inherent biases in the datasets (socio-demographics, traces, competencies, learning objectives, etc.) used for training. Regarding deep learning, the authors observed an increase in pre-trained models, the development of more efficient architectures, and the growing use of interpretability techniques.

The last article, titled "ChatGPT in ESL writing classrooms: Potentials and implications" by **Karim Ibrahim** and **Robert Kirkpatrick** used a systematic review design to synthesize available research on the educational potentials of ChatGPT as an instructional assistant, outline the implications of these potentials for L2 writing instruction, and discuss their practical implications. Based on a meta-analysis of 42 research articles, the findings demonstrate that ChatGPT can enhance L2 writing instruction by boosting learners' motivation, automating instructional tasks, and offering instantaneous, personalized feedback.

References

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *1*, 100002. https://doi.org/10.1016/j.caeai.2020.100002

Holmes, W., Bialik, M., & Fadel, C. (2023). Artificial intelligence in education. In C. Stuckelberger & P. Duggal (Eds.), *Data ethics: Building trust: How digital technologies can serve humanity* (pp. 621-653). Globethics Publications. https://doi.org/10.58863/20.500.12424%2F4273108

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, *2*, 100020. https://doi.org/10.1016/j.caeai.2021.100020

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

AI Application (ChatGPT) and Saudi Arabian Primary School Students' Autonomy in Online Classes: Exploring Students and Teachers' Perceptions

Dr. Ali Rashed Ibraheam Almohesh
Department of Arabic Language, College of Education in Al-Kharj, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia

Abstract

In education, the integration of artificial intelligence (AI) has presented opportunities to transform the dynamics of online learning. This study investigated the impact of an AI-powered application, namely ChatGPT, on the autonomy of Saudi Arabian primary students participating in online classes. It also explored how the implementation of Chat GPT influenced Saudi Arabian primary students' autonomy. In this mixed-methods study, a quasi-experimental design assessed the impact of ChatGPT on learner autonomy among 250 Saudi Arabian primary students from six primary schools in Riyadh, Saudi Arabia. The quantitative analysis employed descriptive statistics and t-tests, while the qualitative data underwent interpretative phenomenological analysis. To ensure coding reliability, 20% of the codes were independently reviewed by an external coder, with a 94% inter-coder agreement coefficient reached through consensus. Findings revealed that ChatGPT significantly affected the participants' perceptions of autonomy and its different dimensions. Qualitative data showed that AI-powered applications contributed to the students' autonomy in 10 different ways. Participants also mentioned that AI-powered apps might have some negative consequences. This study has theoretical implications for redefining learner autonomy in the digital age and calls for the exploration of many facets of autonomy. Practical applications from this study include strategic integration of AI into online education, data security, and the need for orientation programs.

Keywords: AI-powered applications, ChatGPT, students' autonomy, online classes, students' perceptions

AI Application (ChatGPT) and Saudi Arabian Primary School Students' Autonomy in Online Classes: Exploring Students and Teachers' Perceptions

The use of artificial intelligence (AI) in education is having a significant impact on online learning in Saudi Arabia, influencing personalized learning, automating tasks for educators, and providing insights into student performance (Khan et al., 2022; Popenici & Kerr, 2017; Seo et al., 2021). AI tutoring systems such as those developed by Hwang et al. (2020) were examined. Roll et al. (2018) and VanLehn (2011) aimed to provide personalized advice and support by adapting content to students' individual learning patterns. AI teaching assistants have simplified educators' tasks by addressing routine questions from students in online forums. Meanwhile, AI analytics have provided insights into student performance through clickstream data analysis (Chaudhry et al., 2023; Holstein et al., 2018; Pradana et al., 2023).

Despite the potential benefits of AI in education, concerns may arise among students and teachers in Saudi Arabia. Students may perceive the collection and analysis of their data by AI systems as an invasion of privacy, reminiscent of incidents such as the Facebook-Cambridge Analytica data scandal. Concerns about data or algorithmic bias from AI agents have been viewed as discriminatory (Murphy, 2019; Rudolph et al., 2023). Educators have expressed concerns that an over-reliance on AI could impact students' ability to learn independently, think critically, and solve problems creatively (Wogu et al., 2018). Therefore, it is crucial to examine students' and teachers' perceptions of the impact of AI on online learning in the Saudi Arabian context (Seo et al., 2021).

Within the AI in Education (AIEd) community, there has been an active effort to explore the role of AI systems in shaping online education, extending beyond traditional educational settings (Roll & Wylie, 2016). A systematic review of AIEd literature from 2007 to 2018 highlighted a significant gap, emphasizing the need for more rigorous examination and contemplation of the ethical ramifications associated with AI systems in the dynamic space of learner-instructor interactions (Zawacki-Richter et al., 2019). Popenici and Kerr (2017) focused on the broader impact of AI on the learning and teaching landscape and revealed potential conflicts such as privacy concerns, shifts in power dynamics, and issues of excessive control. These findings have underscored the urgency for continued research efforts aimed at identifying existing gaps, unresolved issues, or potential barriers to the full transformative potential of AI systems within educational contexts.

ChatGPT has gained attention for its ability to provide detailed written responses based on extensive databases, but there have been concerns about its factual accuracy (Ray, 2023). A pilot study using ChatGPT for academic papers found that the AI chatbot produced coherent and informative content, suggesting a potential focus on enhancing students' creativity and critical thinking in education (Zhai, 2022). Thoughtfully used, ChatGPT could offer language teachers an opportunity to enrich language instruction and create engaging language learning experiences for their students.

In contemporary language education, learner autonomy has assumed a pivotal role, especially in the era of digital advancements. Autonomous learners have proactively shouldered the responsibility for their educational goals, exercising informed choices regarding what, how, and when to learn, all while proficiently managing their learning processes within online environments (Benson, 2007; Kang & Im, 2013). This autonomy has not only allowed students to discern their strengths and weaknesses but also facilitated the adaptation of learning strategies and active engagement in the learning process. Furthermore, it has endowed students with indispensable lifelong learning skills, fostering independence in managing their time and pursuing knowledge.

Research on AI systems in online education has highlighted a gap in understanding how both students and instructors perceive and interact with these technologies (Wogu et al., 2018). A comprehensive understanding of learner perceptions and experiences is crucial for the ethical and effective implementation of AI systems in online education. Therefore, research into learner perceptions and experiences can inform strategies that harness the potential of AI while addressing privacy concerns and preserving learner autonomy (Kang & Im, 2013).

In the context of Saudi Arabia, the use of artificial intelligence (AI) in education has witnessed a surge, impacting various facets of online learning. The implementation of AI systems holds the promise of personalized learning experiences, task automation for educators, and valuable insights into student performance. However, within this landscape, there is a critical need to understand how these AI applications are perceived by both students and instructors in the Saudi Arabian educational setting. Given the cultural nuances and specificities of the Saudi Arabian context, such research becomes imperative to ensure that the integration of AI aligns with the local educational ethos and addresses any unique concerns that may arise.

The rationale for this study was based on the potential gaps and unexplored territories within the Saudi Arabian educational framework concerning AI in online learning. While there is growing enthusiasm about the benefits of AI, including its potential to enhance personalized learning and streamline educational tasks, there has been limited research on how individuals within the Saudi Arabian educational system perceive and interact with these technologies. Moreover, the cultural and contextual factors specific to Saudi Arabia may introduce distinctive dynamics in the learner-instructor relationship when AI is introduced. This study aimed to bridge this gap by delving into the perceptions, concerns, and experiences of students and instructors, offering insights that can inform the ethical and effective implementation of AI in online education within the Saudi Arabian context. Understanding these nuances is pivotal for ensuring that AI systems align with the cultural values, privacy expectations, and educational goals unique to Saudi Arabia, ultimately fostering an environment where AI complements and enhances the learning experience without compromising learner autonomy. This study investigated the effects of an AI-powered application, ChatGPT, on Saudi Arabian primary school students' autonomy in online classes. It also explored students' and teachers' perceptions of the AI-powered application. The study was framed by the following research questions:

1. Do AI-powered applications (e.g., ChatGPT) have significant impact on Saudi Arabian primary school students' autonomy in online classes?

2. What are the Saudi Arabian students and teachers' perceptions of using AI-powered applications in online classes?

Literature Review

The development of chatbots, starting with Weizenbaum's ELIZA in the 1960s and progressing through entities such as ALICE, Cleverbot, and integration into messaging apps such as Facebook Messenger, has been marked by advances in mimicking human-like appearances, as highlighted by Ayedoun et al. (2019) and Huang et al. (2018). Modern chatbots have used sophisticated techniques such as natural language processing and neural machine translation (Smutny & Schreiberova, 2020). There has been a growing trend to integrate chatbots into second and foreign language learning (L2 and FL), which has attracted the attention of researchers such as Wang et al. (2021). Huang et al. (2017) studied GenieTutor and the Mondly chatbot, and highlighted their significant role in English as a foreign language (EFL) learning contexts, as they provided a range of benefits and transformed the learning experience for students.

GenieTutor, developed by Huang et al. (2017) represented a significant advance in the application of chatbots in language learning. It provided students with an interactive and personalized learning environment characterized by features such as unlimited patience, instant responses, and a tailored focus on specific topics. Similarly, the Mondly chatbot illustrated the effectiveness of chatbots in EFL learning. Its unique capabilities helped reduce learner anxiety and promoted a supportive learning atmosphere. The benefits that chatbots have brought to language learning go beyond mere convenience. Fryer et al. (2020) highlighted how unlimited patience was a crucial aspect of these AI-controlled language companions. In traditional learning environments, human teachers may be limited by time and resources, making it difficult to respond to the individual pace of each student. However, chatbots have overcome this limitation by offering learners the luxury of unlimited patience. This has not only allowed students to grasp concepts at their own pace, but also ensured a personalized learning experience that adapts to their individual needs.

Additionally, the use of chatbots helped reduce learner anxiety. Fryer et al. (2020) emphasized how the absence of human intervention in the learning process can create a stress-free environment for students. Traditional language learning environments can be anxiety-inducing, especially for beginners who are wary of making mistakes. Non-judgmental and consistently supportive, chatbots have created a more relaxed and enjoyable atmosphere and encouraged learners to take risks, experiment with language, and ultimately improve their language skills. AI systems have played a crucial role in shaping the learning environment. Various AI systems have offered diverse possible uses. For example, AI teaching assistants have improved communication (Rusmiyanto et al., 2023) and AI assessment systems (Perin & Lauterbach, 2018) have optimized grade communication. Integrating AI into online education has also included continuous feedback systems (Luckin, 2017), virtual avatars for collaboration (Heidicker et al., 2017), and AI facial analytics (Aslan et al., 2019) to improve teacher presence and thus strengthen technology-enhanced learning environments.

Despite its positive impact, the integration of commercial AI systems such as Proctorio for proctoring during exams has introduced complexities in the interaction between learners and teachers, raising concerns about test anxiety (Bajaj & Li, 2020). Similarly, the application of Squirrel AI aimed at adaptive learning has raised concerns that it could potentially limit students' creative learning processes (Beard, 2020). Research has suggested positive effects of chatbots on critical thinking skills, as well as effectiveness, especially in speaking tasks (El Shazly, 2021; Kooli, 2023). However, concerns have been raised about novelty effects in language learning (Fryer et al., 2017), along with criticism of mechanical behavior and the lack of essential communication components (Smutny & Schreiberova, 2020). In particular, empirical studies on the influence of chatbots on L2 and FL learning are still incomplete (Kooli, 2023). Smutny and Schreiberova (2020) proposed research to provide guidelines for integrating chatbots into teaching methods and analyze interactions between learners and chatbots.

Learner Autonomy in Online Education

Learner autonomy in education refers to a student's capacity to take responsibility for their learning and actively engage in the learning environment. This includes (a) making decisions, (b) setting goals, (c) monitoring progress, (d) self-assessment, (e) choosing effective learning strategies, (f) collaborating with others, (g) seeking guidance from peers and educators, and (h) reflecting on learning experiences. The concept encompasses both cognitive elements like awareness, perception, motivation, and reflection, as well as behavioral aspects, including specific learning actions and strategies (Benson, 2007).

The digital age has introduced new opportunities for learner autonomy in education. Online learning environments have provided students with the ability to manage their learning independently, offering flexibility in setting schedules, organizing resources, and taking charge of their educational journey (Dang, 2010, 2012). Learner autonomy, vital in online learning, has promoted active learning and preparing students for lifelong, self-directed learning beyond the classroom (Davis et al., 2019). With the support of technology, students have had unprecedented access to self-study, synchronous and asynchronous interaction with instructors, and collaborative learning experiences (Hutapea, 2019; Tran & Duong, 2020).

Online learning modes, including synchronous and asynchronous learning, offer varying levels of direct interaction with instructors and peers. Asynchronous learning, facilitated through platforms like Moodle, e-mail, and discussion forums, has allowed students to access educational resources at their convenience, promoting self-regulation and self-motivation (Zhong, 2018). In contrast, synchronous learning through videoconferencing tools like Zoom or Google Meet has enabled real-time interaction despite geographical distances, further enhancing learner autonomy (Dashtestani, 2020).

To promote learner autonomy in online education, teachers play a crucial role. Autonomy is not a fixed state, but an ongoing process that students achieve through specific conditions (Benson, 2007). Educators should provide guidance and scaffolding to facilitate students' development of autonomy in online learning (Lai, 2019). This has included teaching students self-regulation strategies, promoting motivation and engagement, emphasizing active learning, promoting metacognition through self-assessment, and encouraging self-directed learning (Almusharraf, 2020). Peer assessment, collaborative group work, and knowledge sharing opportunities also contribute to learners' autonomous decision-making and problem-solving (Lai, 2019). Overall, learner autonomy in online education has leveraged technology and educational strategies to empower students to take ownership of their learning, manage their time effectively, and develop the skills necessary for lifelong self-directed learning and successful collaboration (Borg & Alshumaimeri, 2019). Online learning environments have provided students with multiple opportunities to become autonomous, self-regulated, and confident learners who are able to make informed decisions and actively participate in their educational journeys.

Methodology

Sample and Procedure

In line with the research objectives, we used a mixed-methods research design. For the quantitative phase, we used a quasi-experimental research design (i.e., pretest/posttest, control and experimental groups) to assess the effect of ChatGPT on students' learner autonomy and its different aspects. However, for the qualitative phase, we employed a qualitative phenomenological research method to delve deeply into the lived experiences of the Grade 6 Saudi Arabian primary students' use of AI-powered applications such as ChatGPT. To ensure control for external variables such as large differences and teacher effects, we selected six intact, sixth grade classes from six primary schools in Riadh province, Saudi Arabia. Three intact classes consisting of 75 students in total were assigned to the experimental group, and three classes consisting of 75 students in total were assigned to the experimental group. The students in the experimental group were invited to a two-hour workshop on the Chat GBT application, and they were taught how it might be useful in education, while the control group received instruction through conventional teaching methods, based on the education system of the country. The participants for the qualitative phase were 14 students and five primary school teachers who were engaged with AI-empowered technology (ChatGPT) in teaching and learning.

A learner autonomy questionnaire developed by Little (1996) was employed for data collection. This questionnaire consisted of a five-point Likert scale, ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*), and was designed to address our first research question, which aimed to evaluate students' levels of learner autonomy. The questionnaire items were categorized into five domains: (a) self-direction (eight items); (b) self-instruction (four items); (c) self-access (three items); (d) motivation (seven items); and (e) collaboration (six items). The questionnaire was translated to Arabic. The reliability of the instrument was estimated using Cronbach's alpha, and the results showed that the internal consistency of the questionnaire and its dimensions exceeded 0.83. Furthermore, an interview checklist consisting of 10 open-ended questions followed by prompts was used to explore the participant's perceptions, feelings, and cognitions of the advantages and challenges of using AI-powered apps such as ChatGPT. The interview content was confirmed by three university professors interested in educational studies.

After completing 10 teaching units, all six classes were asked to fill out the learning autonomy questionnaire to assess the impact of teaching methods and AI applications on learner autonomy. Thirty students from Group A, which used AI applications, were then interviewed individually to gain deeper insights into their experiences. It is noteworthy that after the survey of the 14th student, data saturation was reached, which indicated that no further important information was to be expected.

Data Analysis

The quantitative data in this study were analyzed using descriptive statistics and independent sample *t*-tests. At the same time, the qualitative data was subjected to a multi-stage process. First, the qualitative data were transcribed and then an interpretative phenomenological analysis (IPA) was carried out. The IPA involved several phases, including immersion in the text, identifying themes, grouping themes, exploring interrelationships, and summarizing with supporting examples. To facilitate the coding, categorization, and topic development processes, the MAXQDA 2022 software was used. Using a thematic approach, codes, categories, and themes were extracted directly from the qualitative data. To increase the reliability of the coding process, 20% of the codes (n = 24) were independently checked by an external coder with experience in thematic analysis. During this independent review, disagreements between the primary coder and external coder occurred twice, resulting in an intercoder agreement coefficient of 94%. To resolve these discrepancies, a common consensus was reached through discussion and mutual agreement. The aim of this rigorous process was to ensure the accuracy and consistency of the qualitative data analysis in this study.

Results

Impact of AI on Students' Autonomy

Table 1 summarizes data comparing the intact classes' scores on different dimensions of learner autonomy before the treatment.

Table 1

Control and Experimental Group Scores on Learner Autonomy Pretest

Group	Collaboration		Motivation		Self-access		Self-instruction		Self-direction		***t***	***df***	***p***
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	0.89	248	0.62
Control	3.2	0.85	3.3	0.83	2.90	0.79	2.65	0.67	3.1	0.56			
Experimental	3.3	0.79	3.26	0.91	2.83	0.82	2.70	0.83	3.20	0.67			

As indicated in Table 1, no statistically significant differences were observed between the control group concerning collaboration. Likewise, there were no significant differences between the control group and the experimental group regarding motivation. Similarly, no statistically significant differences were found between the control group and the experimental group in terms of self-access. Additionally, there were no significant distinctions between the control group and the experimental group concerning self-instruction. Finally, the results indicated no significant difference between the control group and the experimental group in mean scores on self-direction. Consequently, the two groups demonstrated homogeneity at the commencement of the study. To assess whether AI-powered applications contributed to students' autonomy, t-tests were conducted on the groups' posttest scores, and the results are presented in Table 2.

Table 2

Control and Experimental Group Scores on Learner Autonomy Posttest

Group	Collaboration		Motivation		Self-access		Self-instruction		Self-direction		*t*	*df*	*p*
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	0.89	248	< 0.001
Control	3.4	0.83	3.29	0.87	3.1	0.80	2.71	0.83	3.23	0.72	12.23	248	
Experimental	4.11	1.1	4.21	0.84	4.2	0.83	3.92	1.1	4.21	0.89			

As depicted in Table 2, there were notable and statistically significant differences between the control group and the experimental group (in terms of collaboration, motivation, self-access, self-instruction and self-direction. Consequently, the two groups exhibited divergent perceptions of learner autonomy at the conclusion of the treatment, highlighting the positive contribution of AI-powered applications to students' autonomy.

Students' and Teachers' Perceptions

The interviews with 14 students and five teachers were transcribed, and 35 open themes were extracted, which were reduced to 10 central themes. Each central theme, along with its constituents, has been explained and exemplified in the following sections.

Autonomy in Learning

The first extracted theme consisted of self-directed learning, independence, and control over learning. This theme reflected students' appreciation for the ability to shape their learning experiences independently, allowing them to set their own pace and explore topics of personal interest. For example, Student 1 emphasized the significance of AI tools in providing autonomy, noting that "these AI tools give me the autonomy to learn at my own pace and explore topics of interest." Teacher 2 echoed a similar sentiment, stating "students feel more in control of their learning journey; they can decide what and when to study." This finding underscored the common perspective among students regarding the autonomy that AI-powered applications offer in their educational journeys.

Personalized Learning

The second central theme encompassed tailored content, adaptive feedback, and individual progress. Students appreciated AI's ability to cater to their specific learning needs. For instance, Student 9 highlighted the personalized nature of AI applications, noting that they "adapt to my level, offering a personalized learning experience." Teacher 1 reinforced this view, stating "the feedback the students get from these apps is specific to their needs, which enhances their learning." This demonstrated the consensus among students on the value of personalized learning experiences facilitated by AI-powered applications.

Time Management

The third central theme encompassed efficient use of time, flexibility, and scheduling. Students found AI tools helpful in optimizing their study schedules. For example, Student 4 highlighted the efficiency of AI tools. "Using AI tools helps me manage my study time more efficiently, and I can learn when it suits me." Student 6 echoed this sentiment, saying, "I appreciate the flexibility to study anytime, especially when I have a busy schedule" (S6). This reflects the consensus among students on the positive impact of AI-powered applications on time management in their learning journeys.

Motivation

The fourth central theme comprised engagement, gamification, and progress tracking. Students derived motivation from AI applications, including gamified elements and progress monitoring. For instance, Student 6 noted the impact of gamification. "The gamified elements in these apps make learning fun and keep me motivated." Teacher 3 also added that "seeing the students' progress and achievements motivates them to continue using these tools." These testimonies underscored the role of AI-powered applications in maintaining student motivation and engagement.

Learning Enhancement

The fifth central theme included complementing traditional education, skill development, and deeper understanding. Students found AI applications augmented their classroom learning and deepen their understanding of subjects. For instance, Student 12 expressed how AI applications complemented their learning. "These AI applications complement my classroom learning and help me grasp concepts more deeply." Teacher 2 also highlighted the benefit of these tools, saying that "I've developed language skills faster using these tools alongside my classes" (Teacher 2). These accounts underscored the role of AI-powered applications in enhancing and reinforcing students' educational experiences.

Autonomy Versus Guidance

The sixth central theme dealt with balancing autonomy with the need for human guidance. Students valued the independence AI offered but also recognized the importance of human interaction. Student 7 emphasized the need for occasional guidance. "While I value autonomy, sometimes I wish for human guidance to clarify doubts." Student 11 shared their approach to striking a balance. "I find a balance by using AI tools for self-paced learning and consulting teachers when needed." These perspectives highlighted the importance of harmonizing autonomous learning with human support when necessary.

Data Privacy and Ethics

The seventh central theme encompassed data security concerns, ethical implications, and privacy awareness. Students expressed concerns about how their data was used and were mindful of the ethical aspects of AI in education. Student 4 voiced concerns about data use. "I'm concerned about how my data is used, and I want to ensure my privacy while using AI apps." Teacher 3 also emphasized the importance of ethical awareness. "The ethical implications of AI in education are important; I want to be aware of these issues." The findings within this theme underscored students' vigilance regarding data privacy and ethical considerations in AI-powered educational tools.

Skill Transferability

The eighth central theme involved the applicability of knowledge and skills gained from AI apps to real-life situations. Students pondered the practical value of the skills they acquired. For example, Student 14 expressed curiosity about skill applicability. "I wonder if the skills I gain from these apps will be useful in real-life situations." Similarly, Teacher 5 articulated hopes for skill application. "I hope the students can apply the knowledge and skills acquired from AI apps to their future career." These viewpoints highlighted participants' considerations regarding the real-world utility of skills acquired through AI applications.

Technological Dependence

The ninth central theme encompassed concerns about reliance on AI, potential reductions in critical thinking, and overreliance on technology. Students expressed apprehensions about becoming overly dependent on AI. For instance, Student 12 said "I worry that I might become too dependent on AI and lose my critical thinking skills." Student 3 also stated that "I'm cautious not to rely solely on AI; I still want to think critically and solve problems." Similarly, Teacher 2 noted that if students were highly dependent on the use of technology they might not think critically. These perspectives highlighted awareness of the need to balance technology use with critical thinking and problem-solving skills.

Educational Access

The final central theme was concerned with the potential for AI tools to enhance educational access for diverse learners and address learning barriers. Participants recognized the value of AI in promoting inclusivity. For example, Teacher 3 highlighted the potential of AI tools noting that they "can make education more accessible for people with different learning needs." Teacher 1 expressed appreciation for the inclusivity AI offers and that "it can bridge educational gaps and empower more learners." These insights underscored the positive impact of AI in fostering educational inclusivity and accessibility.

Discussion

In the digital age, the integration of AI into online education has revolutionized the way students and instructors interact. AI-powered applications have enhanced various aspects of the online learning experience, from streamlining communication between instructors and students to providing personalized learning content and continuous feedback. In this study, we delved into the intricate relationship between AI-powered applications and learner autonomy in the online education context, with a specific focus on collaboration, motivation, self-access, self-instruction, and self-direction.

The quantitative data revealed noteworthy disparities in autonomy between the control group and the experimental group who were exposed to AI-powered applications. These differences manifested across multiple dimensions, including collaboration, motivation, self-access, self-instruction, and self-direction. The experimental group, enriched with the use of AI-powered applications, exhibited significantly higher levels of learner autonomy in each of these domains. This indicated a positive contribution of AI technology to enhancing learner autonomy within online education.

To comprehensively understand and substantiate these findings, it is crucial to contextualize them within the broader landscape of existing research on the intersection of AI and online education, particularly as it relates to learner autonomy. A multitude of studies have illustrated the positive impact of AI systems on diverse facets of online education, from facilitating communication between instructors and students to delivering personalized learning experiences and feedback mechanisms. Goel and Polepeddi (2018) autonomously responded to student introductions, posted weekly announcements, and addressed routine queries, effectively enhancing collaboration within the online learning environment. Furthermore, the work of Ross et al. (2018) offered insight into online adaptive quizzes that provided personalized learning content tailored to each student's specific needs. This personalization not only motivated students but also fostered a deeper sense of engagement, aligning with the findings on motivation in our study. Heidegger et al. (2017) explored the innovative use of virtual avatars in immersive virtual environments to facilitate collaboration among physically separated users. These avatars significantly contributed to a heightened sense of presence and effective collaboration, aligning with our findings on collaboration.

Additionally, Aslan et al. (2019) devised AI facial analytics to enhance instructors' presence as coaches in technology-mediated learning environments, thereby boosting self-direction and autonomy. Moreover, Luckin (2017) demonstrated AI systems that provided continuous feedback on students' learning processes and their progress toward learning goals. This continuous feedback mechanism was instrumental in motivating students and guiding their self-instruction, in line with our study's findings. Furthermore, Tran and Duong (2020) emphasized the importance of learner autonomy in online education, with the support of technology, offering unprecedented access to self-study, asynchronous and synchronous interaction with instructors, and collaborative learning experiences. Their work aligns with our focus on autonomy in online education.

Finally, Richardson et al. (2020) highlighted the role of instructors in promoting learner autonomy in online learning, emphasizing the importance of self-regulation strategies, motivation, engagement, metacognition, and self-directed learning. Their findings were consistent with our study's emphasis on the significance of autonomy in online education and the role of instructors in fostering it. However, it is important to acknowledge that not all studies aligned with our findings. Bergmans et al. (2021) discussed the implementation of Proctorio, a system designed to prevent cheating by monitoring students during exams and raised concerns about test-taking anxiety and potential challenges in collaboration. Similarly, Beard (2020) highlighted concerns regarding the potential restriction of creative learning in the context of Squirrel AI, suggesting potential issues with motivation.

These variations underscored the importance of considering the specific context, implementation, and characteristics of AI-powered applications, as these factors can influence outcomes. Moreover, they emphasized the need for further research to gain a more nuanced understanding of the relationship between AI technology and learner autonomy. In conclusion, our findings shed light on the substantial impact of AI-powered applications on learner autonomy within the online education landscape. By examining the contributions of AI technology to collaboration, motivation, self-access, self-instruction, and self-direction, we have gained valuable insights into the complex dynamics of online education. While existing research provides a strong foundation, it is crucial to delve deeper into the unique features and functionalities of AI systems to maximize their potential to enhance learner autonomy, ultimately enriching the online learning experience for students and educators.

Implications

The study has several theoretical implications. First, it contributes to the growing literature on the integration of artificial intelligence into educational environments and illuminates its impact on student autonomy—a crucial aspect of effective learning. The findings have expanded our understanding of the nuanced dynamics among technology, pedagogy, and student responsibility in the unique context of Saudi primary education. Additionally, the study can inform educators, policymakers, and curriculum designers about the potential benefits and challenges associated with AI applications in promoting student autonomy, thereby guiding the development of future educational strategies and interventions. Furthermore, examining students' and teachers' perceptions provides a comprehensive perspective that adds depth to the theoretical framework and provides valuable insights into the social and cultural dimensions that influence the implementation of AI in Saudi Arabian classrooms.

This study extended the theoretical foundation of learner autonomy by emphasizing the influential role of AI-powered applications in online education. Autonomy is no longer solely dependent on human agency but can be significantly shaped by technology. This paradigm shift broadens the conceptualization of how learners can develop and exert autonomy in digital learning environments, calling for a reevaluation of established autonomy theories. The study also introduced a nuanced perspective by highlighting that learner autonomy encompasses various facets, including collaboration, motivation, self-access, self-instruction, and self-direction. This diversified view calls upon researchers to explore a broader spectrum of factors when investigating learner autonomy in the online education context. This expansion enhances our comprehension of the multifaceted nature of autonomy in contemporary learning environments. Also, the findings offered a significant contribution to the theoretical notion that AI technology can serve as a mediating factor to enhance the quality of education and foster learner autonomy. This proposition paves the way for the development of more comprehensive models that elucidate the role of AI in mediating online learning experiences, with implications for broader educational theories. Moreover, the study presents a theoretical underpinning for the complex relationship between autonomy and guidance in online education. This underlines the need for a thoughtful reconciliation between autonomous learning and the necessity for human interaction and support within digital learning landscapes, a topic that warrants continued theoretical exploration.

Educational institutions can strategically leverage the study's findings further to integrate AI-powered applications into their online learning environments. These applications should not only facilitate self-directed learning but also actively promote collaboration, motivation, self-access, self-instruction, and self-direction. Such integration stands to enhance the overall learning experience. In addition, instructors engaged in online education should be equipped with appropriate training and support to effectively harness the potential of AI systems in promoting learner autonomy. These initiatives should extend to using AI tools to encourage self-directed learning, foster collaboration, and boost student motivation. In addition, educational institutions and course developers can benefit from revisiting and adapting their curricula to align with the principles of learner autonomy and AI integration. Curricular design should strike a harmonious balance between autonomous learning experiences and guided instruction, thereby empowering students to navigate their educational journeys.

Educational institutions and technology developers should prioritize data security, transparency, and ethical practices to mitigate potential concerns and ensure responsible AI use. AI applications can be further refined to offer personalized feedback and adapt to the unique learning needs of each student. This practical enhancement has the potential to significantly contribute to improving motivation, self-access, and self-instruction, as highlighted in the study.

In addition, to facilitate students in maximizing the benefits of AI-powered applications, educational institutions can initiate orientation programs. These programs can provide students with guidance on effective use of AI tools, enabling them to make informed decisions about their learning paths and seek human support when required. Collaboration among educators, technologists, and researchers is paramount for designing and implementing AI-powered applications that align with the principles of learner autonomy. Interdisciplinary teamwork can yield holistic solutions that enhance the online learning experience by ensuring that AI systems are effectively integrated into the pedagogical framework. Finally, the study underscores the need for ongoing research and development efforts in the field of AI in education. Institutions and technology developers should continue to innovate and enhance AI applications to better support learner autonomy in the online learning realm. Continuous improvement and innovation are essential in this ever-evolving landscape.

Despite the merits of this study, there were some limitations including potential cultural biases in survey responses and a focus on a specific educational level. Further studies can explore broader educational contexts and validate findings across diverse demographics. Examining long-term impacts of AI implementation in Saudi Arabian online education and assessing the effectiveness of cultural adaptations in AI systems would provide deeper insights. Researchers may also consider investigating the evolving role of AI in fostering learner autonomy while considering additional cultural dimensions in the Saudi Arabian context as well as other contexts.

Funding

This study was supported via funding from Prince Sattam Bin Abdulaziz University Project Number (PSAU 2024 /R/1445).

References

Almusharraf, N. (2020). Teachers' perspectives on promoting learner autonomy for vocabulary development: A case study. *Cogent Education, 7*(1), 1823154. https://doi.org/10.1080/2331186X.2020.1823154

Aslan, S., Alyuz, N., Tanriover, C., Mete, S. E., Okur, E., D'Mello, S. K., & Arslan Esme, A. (2019, May). Investigating the impact of a real-time, multimodal student engagement analytics technology in authentic classrooms. *Proceedings of the 2019 Chi Conference on Human Factors in Computing Systems* (pp. 1–12). https://dl.acm.org/doi/10.1145/3290605.3300534

Ayedoun, E., Hayashi, Y., & Seta, K. (2019). Adding communicative and affective strategies to an embodied conversational agent to enhance second language learners' willingness to communicate. *International Journal of Artificial Intelligence in Education, 29*(1), 29–57. https://doi.org/10.1007/s40593-018-0171-6.

Bajaj, M. & Li, J. (2020, April 4). Students, faculty express concerns about online exam invigilation amidst COVID-19 outbreak. The Ubyssey. https://ubyssey.ca/news/Students-express-concerns-about-online-exams/

Beard, A. (2020, March 19). *Can computers ever replace the classroom?* The Guardian. https://www.theguardian.com/technology/2020/mar/19/can-computers-ever-replace the-classroom.

Benson, P. (2007). Autonomy in language teaching and learning. *Language Teaching, 40*(1), 21–40. https://doi.org/10.1017/ S0261444806003958

Bergmans, L., Bouali, N., Luttikhuis, M., & Rensink, A. (2021). On the efficacy of online Proctoring using Proctorio. In *Proceedings of the 13th International Conference on Computer Supported Education - Volume 1*. (pp. 279-290). https://doi.org/10.5220/0010399602790290

Borg, S., & Alshumaimeri, Y. (2019). Language learner autonomy in a tertiary context: Teachers' beliefs and practices. *Language Teaching Research*, *23*(1), 9–38. https://doi.org/10.1177/1362168817725759

Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT: A case study. *Cogent Education, 10*(1), 2210461. https://doi.org/10.1080/2331186X.2023.2210461

Dang, T. T. (2010). Learner autonomy in EFL studies in Vietnam: A discussion from a sociocultural perspective. *English Language Teaching, 3*(2), 3–9. https://doi.org/10.5539/elt.v3n2p3

Dang, T. T. (2012). Learner autonomy: A synthesis of theory and practice. *The Internet Journal of Language, Culture and Society, 35*(1), 52–67. https://aaref.com.au/wp-content/uploads/2018/05/35-08.pdf

Dashtestani, R. (2020). Online English for academic purposes instruction in the context of Iran: Exploring the instructor element. *Teaching English with Technology, 20*(5), 23–37, http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.desklight-f557ad7d-dbb6-4211-9307-1fa091f43998

Davis, N. L., Gough, M., & Taylor, L. L. (2019). Online teaching: Advantages, obstacles and tools for getting it right. *Journal of Teaching in Travel & Tourism, 19*(3), 256–263. https://doi.org/10.1080/15313220.2019.1612313

El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, *38*(3), e12667. https://doi.org/10.1111/exsy.12667

Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior, 75*, 461–468. https://doi.org/10.1016/j.chb.2017.05.045

Fryer, L. K., Coniam, D., & Lăpușneanu, D. (2018). Bots for language learning now: current and future directions. *Language, Learning and Technology 24*(2), 8-22. http://hdl.handle.net/10125/44719

Goel, A. K., & Polepeddi, L. (2018). Jill Watson: A Virtual Teaching Assistant for Online Education. In C. Dede, J. Richards, & B. Saxberg (Eds.), *Learning engineering for online education* (pp. 120–143). https://doi.org/10.4324/9781351186193-7

Heidicker, P., Langbehn, E., & Steinicke, F. (2017). Influence of avatar appearance on presence in social VR. *2017 IEEE Symposium on 3D User Interfaces* (pp. 233–234). Institute of Electrical and Electronic Engineers. https://doi.org/10.1109/3DUI.2017.7893357

Holstein, K., Hong, G., Tegene, M., McLaren, B. M., & Aleven, V. (2018). The classroom as a dashboard: Co-designing wearable cognitive augmentation for K-12 teachers. *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge* (pp. 79–88). https://doi.org/10.1145/3170358.3170377

Huang, J.-X., Lee, K.-S., Kwon, O.-W., & Kim, Y.-K. (2017). A chatbot for a dialogue-based second language learning system. In K. Borthwick, L. Bradley & S. Thouësny (Eds.), CALL in a climate of change: Adapting to turbulent global conditions – short papers from EUROCALL 2017 (pp. 151–156). https://doi.org/10.14705/rpnet.2017.eurocall2017.705

Huang, J.-X., Kwon, O.-W., Lee, K.-S., & Kim, Y.-K. (2018). Improve the chatbot performance for the DB-CALL system using a hybrid method and a domain corpus. In P. Taalas, J. Jalkanen, L. Bradley & S. Thouësny (Eds.), Future-proof CALL: Language learning as exploration and encounters – short papers from EUROCALL 2018 (pp. 100–105). https://doi.org/10.14705/rpnet.2018.26.820

Hutapea, N. M. (2019). Improving senior high school students learning autonomy through generative learning. *Journal of Educational Sciences, 3*(1), 84–95. https://jes.ejournal.unri.ac.id/index.php/JES/article/view/6979/6184

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, 100001. https://www.sciencedirect.com/science/article/pii/S2666920X20300011

Kang, M., & Im, T. (2013). Factors of learner-instructor interaction which predict perceived learning outcomes in online learning environment. *Journal of Computer Assisted Learning, 29*(3), 292–301. https://doi.org/10.1111/jcal.12005

Khan, S., Zaman, S. I., & Rais, M. (2022). Measuring student satisfaction through overall quality at business schools: A structural equation modeling. *South Asian Journal of Social Review 1*(2), 34–55. https://doi.org/10.57044/sajsr.2022.1.2.2210

Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability*, *15*(7), 5614. https://doi.org/10.3390/su15075614.

Lai, C. (2019). Learning beliefs and autonomous language learning with technology beyond the classroom, *Language Awareness, 28*(4), 1–19. https://doi.org/10.1080/09658416.2019.1675679

Little, D. (1996). Freedom to learn and compulsion to interact: Promoting learner autonomy through the use of information systems and information technologies. In R. Pemberton, S. L. Edward Li, W. F. Winnie, & H. D. Pierson (Eds.), *Taking control: Autonomy in language learning* (pp. 203-218). Hong Kong University Press.

Luckin, R. (2017). Towards artificial intelligence-based assessment systems. *Nature Human Behavior, 1*(3), 1–28. https://doi.org/10.1038/s41562-016-0028

Murphy, K. P. (2019). *Machine learning: A probabilistic perspective*. The MIT Press. https://mitpress.mit.edu/9780262018029/machine-learning/

Perin, D., & Lauterbach, M. (2018). Assessing text-based writing of low-skilled college students. *International Journal of Artificial Intelligence in Education, 28*(1), 56–78. https://doi.org/10.1007/s40593-016-0122-z

Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning, 12*(1), 22. https://doi.org/10.1186/s41039-017-0062-8

Pradana, M., Elisa, H. P., & Syarifuddin, S. (2023). Discussing ChatGPT in education: A literature review and bibliometric analysis. *Cogent Education, 10*(2), 2243134. https://doi.org/10.1080/2331186X.2023. 2243134

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems, 3*(1), 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Richardson, J. W., Lingat, J. E. M., Hollis, E., College, R., & Pritchard, M. (2020). Shifting teaching and learning in online learning spaces: An investigation of a faculty online teaching and learning initiative. *Online Learning Journal, 24*(1), 67–91. https://doi.org/10. 24059/olj.v24i1.1629

Roll, I., Russell, D. M., & Gašević, D. (2018). Learning at scale. *International Journal of Artificial Intelligence in Education, 28*(4), 471–477. https://doi.org/10.1007/s40593-018-0170-7

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education, 26*(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3

Ross, B., Chase, A.-M., Robbie, D., Oates, G., & Absalom, Y. (2018). Adaptive quizzes to increase motivation, engagement and learning outcomes in a first year accounting unit. *International Journal of Educational Technology in Higher Education, 15*(1). https://doi.org/10.1186/s41239-018-0113-2

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching, 6*(1). https://doi.org/10.37074/jalt.2023.6.1.9

Rusmiyanto, R., Huriati, N., Fitriani, N., Kusumaning Tyas, N., Rofi'i, A., & Nurmalia Sari, M. (2023). The role of artificial intelligence (AI) in developing English language learner's communication skills. *Journal on Education*, *6*(1), 750–757. https://doi.org/10.31004/joe.v6i1.2990

Seo, K., Dodson, S., Harandi, N. M., Roberson, N., Fels, S., & Roll, I. (2021). Active learning with online video: The impact of learning context on engagement. *Computers & Education*, *165*, 104–132. https://doi.org/10.1016/j.compedu.2021.104132

Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, *151(5)*, 103862. https://doi.org/10.1016/j.compedu.2020.103862

Tran, T. Q., & Duong, T. M. (2020). EFL learners' perceptions of factors influencing learner autonomy development. *Kasetsart Journal of Social Sciences, 41*(1), 194–199. https://so04.tci-thaijo.org/index.php/kjss/article/view/231622

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), 197–221. https://doi.org/10.1080/00461520.2011.611369

Wang, Y., Zhao, R., Li, Z., Wang, H., Yang, Q., & Yang, D. (2021). Building a dialogue agent for language learners. *Proceedings of the 12th International Conference on Educational Data Mining* (pp. 300–309).

Wogu, I. A. P., Misra, S., Olu-Owolabi, E. F., Assibong, P. A., Udoh, O. D., Ogiri, S. O., & Damasevicius, R. (2018). Artificial intelligence, artificial teachers and the fate of learners in the 21st century education sector: Implications for theory and practice. *International Journal of Pure and Applied Mathematics, 119*(16), 2245–2259. https://acadpubl.eu/hub/2018-119-16/2/232.pdf

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, X. (2022, December 27). ChatGPT user experience: Implications for education. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4312418

Zhong, Q. M. (2018). The evolution of learner autonomy in online environments: A case study in a New Zealand context. *Studies in Self-Access Learning Journal, 9*, 71–85. https://doi.org/10.37237/090106

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

Threats and Opportunities of Students' Use Of AI-Integrated Technology (ChatGPT) in Online Higher Education: Saudi Arabian Educational Technologists' Perspectives

Prof. Mesfer Mihmas Mesfer Aldawsari and Dr. Nouf Rashed Ibrahim Almohish
Department of Arabic Language, College of Education in Al-Kharj, Prince Sattam bin Abdulaziz University, Saudi Arabia

Abstract

This research study explored the perspectives of 20 educational technologists from four Saudi Arabian universities regarding the integration of AI-powered technology, particularly ChatGPT, into online higher education. The study used a qualitative research method that relied on the principles of theoretical sampling to select participants and conducted in-depth interviews to collect their insights. The approach taken for data analysis was thematic analysis, which uncovered a rich range of insights on both the challenges and opportunities associated with students' use of AI-integrated technology in the context of online higher education. Ten significant challenges emerged that shed light on the complexities and intricacies of integrating AI-powered technology into educational environments. These challenges included issues related to technological infrastructure, pedagogical adaptation, and the need for comprehensive training programs to empower both teachers and learners. Additionally, eight threats were examined that highlighted concerns about data security, privacy, and potential risks associated with AI technology in educational institutions. This study not only provided a comprehensive overview of the current landscape of AI-integrated technology in Saudi Arabian higher education, but also provided valuable insights for education stakeholders, technologists, and policy makers. It underscored the necessity of proactive measures to mitigate challenges and threats while harnessing the opportunities presented by AI technology to enhance the quality and effectiveness of online higher education.

Keywords: AI-integrated technology, ChatGPT, higher education, online learning, threats, challenges, education technologists

Threats and Opportunities of Students' Use Of AI-Integrated Technology (ChatGPT) in Online Higher Education: Saudi Arabian Educational Technologists' Perspectives

The dominant perspective regarding artificial intelligence (AI) technologies has largely revolved around understanding these systems as a collection of processes and their corresponding responses, emphasizing qualities such as autonomy, adaptability, and interactivity (Ahmad et al., 2020; Baker et al., 2019; Chai, Lin, et al., 2020; Chai, Wang, & Xu, 2020; Dai et al., 2020; Dignum, 2021; Druga et al., 2019). These characteristics are considered fundamental technological focuses that researchers have argued should be integral to AI systems. Although autonomy, adaptability, and interactivity are considered extremely important, they may not cover all the essential criteria for an effective K–12 education. Specifically, these criteria are about skills that are taught by human educators, such as self-efficacy, technical skills, and socialization skills. Samuel (2021), building on Dignum's (2021) notion, emphasized that AI technologies should not only replicate human actions, but also mimic expressions of human intelligence, cognition and logic. This highlighted the need to refine features that determine effective AI in education. The recent challenges in education due to the COVID-19 pandemic provided a unique opportunity to examine the demands on stakeholders, including educators, students, and parents (Fourtané, 2020; Ghallab, 2019; Samuel, 2021; Samuel et al., 2022, 2020).

Prior research in decision support systems, particularly in knowledge support systems and decision support systems (DSS), has focused on their interaction with humans and their impact on decision-making and information dissemination. Kasper (1996) outlined the design theory for DSS, emphasizing optimization for effectiveness. Sankar et al. (1995) highlighted the importance of adaptability in DSS interfaces for favorable decision-making, while Gonzalez and Kasper (1997) found that increased interactivity was associated with enhanced decision-making. Ulfert et al. (2022) observed that higher autonomy in DSS reduced information overload but could impact technostress and intention to use.

The boundary between artificial intelligence and decision support systems is not clear. López-Fernández et al. (2011) suggested that knowledge support systems and decision support systems were extensions of AI research. Phillips-Wren (2012) discussed AI algorithms that provided so-called intelligence to these systems, often using techniques such as neural networks and machine learning. Turban (1995) referred to intelligent systems as expert decision support systems. However, the definition of truly AI-powered tools for specific educational goals has remained unclear. Different areas such as finance and healthcare have different needs. In finance, AI tools can work autonomously (Pelaez et al., 2022). In healthcare, AI often acts as a decision support system subject to human supervision (Panch et al., 2018). In education, the interaction is dynamic, as AI interacts with K–12 learners. Miller (2019) explored human-AI interaction, particularly in robotics, where AI learned from human behavior, establishing a learning loop. Identifying core attributes and applications of AI in education requires careful scrutiny, given the high subjectivity of outcomes contingent on response quality.

The integration of AI in online higher education, specifically through platforms like ChatGPT, has introduced a paradigm shift in the way students access and interact with educational content. While there is growing interest in the opportunities and challenges associated with students' use of AI-integrated technology in online higher education, this study focused on examining these aspects from the unique perspective of Saudi Arabian faculties. This emphasis was crucial because of their pivotal role in shaping the educational experiences of students in Saudi Arabia, a country known for its robust online education ecosystem.

This study looked at the increasing adoption of AI-integrated technology such as ChatGPT in Saudi Arabian online education, driven by demographic changes, technological advancements, and global events. It focused on faculty perspectives and provided a faculty-centric view of AI integration, which has been a gap in the existing research. The cultural context, threats, and opportunities of AI, as well as the unique characteristics of ChatGPT, were the key focus of the study. This study sought to understand how Saudi Arabian faculties perceived the opportunities and risks of ChatGPT in online higher education, and to contribute valuable insights to the discourse on AI in education. Specifically, the following questions were raised:

1. What are the opportunities for students' use of AI-integrated technology (i.e., ChatGPT) while attending online higher education, from the perspectives of educational technologists in Saudi Arabia?

2. What are the threats of students' use of AI-integrated technology (i.e., ChatGPT) while attending online higher education, from the perspectives of educational technologists in Saudi Arabia?

Literature Review

The integration of AI into education stems from fundamental work on decision support and knowledge support systems that has evolved since 1989 with the founding of the *Journal of Artificial Intelligence in Education* (Holstein et al., 2018; Leelawong & Biswas, 2008; Williamson & Eynon, 2020). Notable advances, including the development of Betty's Brain, intelligent tutoring systems, and AI infrastructure, have aimed to improve students' learning experiences. With AI potentially taking on a teaching role, there is a need for research to explore teacher-student dynamics, particularly in K–12 education.

Leelawong and Biswas (2008) examined the effects of student learning facilitated by Betty's Brain, highlighting increased interactivity and feedback, and demonstrating higher learning gains compared to traditional intelligent tutoring systems (ITS). As an extension of ITS research, Holstein et al. (2018) examined combining ITS with real-time analytics for improved teacher monitoring and demonstrated better learning outcomes (Holstein et al., 2018; Leelawong & Biswas, 2008). This positive synergy between AI in education systems and human teaching has highlighted the potential benefits while recognizing the need for further research, particularly in areas related to socialization and psychological challenges in K–12 environments. To optimize the use of AI tools in large classrooms, it is critical to address concerns about scalability and teacher intervention.

While younger students learn academically, they also develop crucial social skills through interaction with peers and teachers. It is essential for students to establish a level of social and emotional comfort with technology, perceiving it as an extension of the teacher, fostering trust and connection (Frenzel et al., 2009). Regrettably, the social and emotional dimensions have received insufficient attention from researchers (Karnouskos, 2022). Intrinsic factors like sociability, enjoyment, and adaptability significantly influence AI acceptance, especially when AI is delivered via robots (De Graaf & Allouch, 2013). Students' behavior and learning outcomes are profoundly shaped by their interactions, attitudes, and emotions towards technology (Karnouskos, 2022). Some companies have explored emotional robots to serve therapeutic roles akin to support animals (Karnouskos, 2022).

Online education, as the second area of this study, has garnered substantial attention within the scholarly literature that explores technology-based learning. Recent developments, notably accelerated by the COVID-19 pandemic, have prompted a thorough examination and investigation of the transition to online learning and the attendant challenges.

A significant body of literature in this domain has centered on online learning environments tailored for adult learners, particularly within the context of college programs. Universities, due to their relatively greater financial resources compared to elementary and secondary schools, have been better equipped to integrate technology into their educational offerings. Entire markets have emerged around online education, with universities that exclusively offer online programs and the advent of massive open online courses such as Coursera and Udemy (Wilson & Gruzd, 2014). Online learning environments have greatly expanded to cover various domains, including coding, business courses, and arts and music classes. Increasingly, platforms like YouTube have been used for educational purposes, often without cost to users, as content creators earn through advertising. The unpredictable and dynamic nature of these technologies has led to a less standardized landscape. The term online learning, a more specific classification within the broader framework of distance learning, has traced its origins to early correspondence schools but is distinguished by its reliance on Internet-based technologies (Kentnor, 2015). Functioning as a pedagogical method, online learning aims to facilitate the transfer of knowledge from an expert, typically a teacher, to a knowledge-seeker, typically a student. The literature in information systems has contributed valuable insights into effective information communication, with theories like media richness and media synchronicity (Dennis et al., 2008) offering a collective understanding of the relationships among information richness, technology suitability, and the nature of tasks. Dennis et al. (2008) specifically categorizes the essential functions of teachers and students, emphasizing conveyance tasks involving knowledge impartation and convergence tasks where students synthesize solutions based on provided information.

In education, key stakeholders, including teachers and students, manifest distinct goals, assessment criteria, technology responsiveness, and reactions to technological stimuli. Research in education, particularly within the context of online learning, has evolved over time to address evolving needs. Initial focuses on design issues and learning characteristics during the 1990s have expanded to encompass intricate aspects, including the examination of learning communities, advanced instructional design methods, and innovative pedagogical approaches (Martin et al., 2020; Neumann & Herodotou, 2020).

Undergraduate students aspire to acquire a wide range of competencies, including vocational skills, and they exhibit a genuine curiosity for knowledge (Reed et al., 2022). These students need a unique set of skills in an online learning environment to thrive in a constantly evolving business landscape, including remote peer engagement and proficiency in using digital technologies for knowledge acquisition (Roper, 2007). Considering this diversity in goals and learning methodologies, the role of AI in education should be highly adaptable and sensitive to the distinct needs and developmental stages of younger learners.

Methodology

Sample and Procedure

The participants in this study were 20 Saudi Arabian educational technologists. These individuals were specifically recruited for their expertise and knowledge in the field of educational technology in Saudi Arabia. The selection of participants was based on a purposive sample with the aim of including individuals who could provide valuable insights into the threats and opportunities associated with students' use of AI-integrated technology in online higher education. The researchers contacted potential participants through professional networks, educational institutions, and relevant organizations in Saudi Arabia. A qualitative research approach was used to explore Saudi educational technologists' perspectives on the threats and opportunities of students' use of AI-integrated technology in online higher education. This approach enabled an in-depth discussion of the participants' viewpoints, experiences, and insights. Data was collected from participants through individual, semi-structured interviews, conducted either in person or via online platforms such as video conferencing tools, depending on participants' preference and availability. The researchers followed an interview guide that included a series of open-ended questions designed to explore various aspects related to the topic. The aim of the questions was to obtain participants' opinions on the potential threats and opportunities of AI-integrated technology in online higher education, as well as their experiences, observations, and recommendations. The interviews were audio-recorded with the participants' consent to ensure accurate data capture. The researchers also took detailed notes during the interviews to supplement the audio recordings. The interviews were conducted in Mandarin Saudi Arabi, the participants' native language, to ensure clear communication and a comfortable environment for sharing their perspectives**.**

Data Analysis

The data obtained from the interviews were transcribed and translated into English for analysis, while ensuring that the original meaning and intent of the participants were retained. A thematic analysis approach was used to identify and analyze recurring patterns, themes, and key findings in the data. The analysis process included several steps. First, the researchers read and re-read the transcripts to familiarize themselves with the data. The first codes were then generated to label and categorize meaningful sections of text. These codes were reviewed and refined through an iterative process, and connections among codes were explored to identify overarching themes. The themes were further refined and organized into a coherent framework that captured the participants' perspectives on the threats and opportunities of students' use of AI-integrated technology in online higher education. The data analysis process was conducted systematically to ensure the trustworthiness and validity of the findings. The researchers maintained an audit trail, documenting decisions made throughout the analysis process to enhance transparency and replicability. Additionally, member checking was performed by sharing the preliminary findings with a subset of participants to validate the interpretations and ensure alignment with their perspectives.

Results

Opportunities in Student Use of AI-Integrated Technology

The first research question explored the opportunities of AI-integrated technology (i.e., ChatGPT) for higher education from educational technologists' perspectives. Interviews with 20 educational technologists were thematically analyzed to reveal and refine 10 main themes, each consisting of two or more sub-categories. Each theme is explained and exemplified below.

Enhanced Personalization

AI-integrated technology, such as ChatGPT, can provide customized learning experiences to individual students. This enhanced personalization involves adapting educational materials, learning paths, and instructional methods to cater to specific needs, learning styles, and pace of each student. This approach aims to optimize the learning process by tailoring it to the unique requirements and preferences of students, ultimately improving their engagement and outcomes. The following quotation exemplified the extracted theme:

> The use of AI-integrated technology enables tailored content delivery and enables educators to provide students with learning materials that match their specific interests and learning styles. This personalized approach allows students to engage deeply with the topic and enhances their overall learning experience. (Professor 2)

Increased Access to Information

A large collection of resources and educational materials have become more readily available through digital platforms. As a result, learners can study a variety of academic subjects and participate in ongoing learning. Using technology, students can access information beyond traditional classroom resources, such as online libraries and diverse learning materials, which broadens their knowledge and promotes a culture of lifelong learning. As Professor 3 argued:

> The digital age has transformed education by giving students unprecedented access to information. Extensive online libraries offer a wealth of resources that students can access, allowing them to delve deeply into interesting topics and learn in a self-directed manner. This democratization of knowledge promotes a culture of continuous learning and empowers students to become lifelong learners.

Efficient Feedback Mechanisms

AI technology has provided efficient feedback mechanisms to give students timely and helpful feedback. By highlighting areas of strength and areas that need more work, this feedback helps students understand concepts and perform better. AI gives teachers the ability to provide students with personalized feedback to help them monitor their progress, pinpoint areas for growth, and modify their learning approaches. This is made possible by real-time assessment and comprehensive performance analytics. As Professor 8 pointed out:

> Detailed performance analysis through AI technology provide valuable insights into student learning progress. By analyzing data on their performance, engagement, and learning patterns, educators can gain a comprehensive understanding of students' strengths and areas that need improvement. This allows for personalized feedback that guides students toward targeted learning strategies and ultimately promotes their academic growth and success.

24/7 Accessibility

Through 24/7 access, the availability of AI-enhanced systems enable students to interact with educational content whenever they choose, thereby offering flexibility and convenience. Students can access educational resources and engage in learning activities in accordance with their unique schedules and time zones thanks to flexible learning hours and global access. This accessibility creates educational opportunities for people all over the world and gives students the power to take charge of their education. According to Professor 10:

> AI-powered systems that provide 24/7 accessibility have changed the learning landscape. With flexible learning times, students can study educational content as they wish and thus balance their studies with other commitments. This flexibility promotes a learning orientation.

Professor 7 noted that AI systems offer an "approach that enables personalized learning experiences and takes into account the different needs and learning styles of students."

Global Collaboration

We have seen opportunities for international collaboration facilitated by online education with AI integration. Global collaboration allows students to engage with peers and educators from diverse cultural backgrounds, fostering cross-cultural communication and providing a platform for diverse learning experiences. Through virtual classrooms and collaborative projects, students can broaden their perspectives, develop intercultural skills, and gain a deeper understanding of global issues. Professor 12 illustrated this theme:

> Diverse learning experiences facilitated by online education with AI integration enrich the educational journey for students. By interacting with educators and fellow students from diverse cultural backgrounds, students gain exposure to different perspectives, approaches, and problem-solving techniques. This exposure enhances their critical thinking abilities, and adaptability, and prepares them to thrive in an interconnected and interdependent world.

Interactive Learning

AI technology can provide engaging and interactive learning experiences for students. By incorporating elements such as gamified learning and immersive simulations, AI enhances student engagement and motivation. These interactive approaches make the learning process more enjoyable, encourage active participation, and facilitate a deeper understanding of the subject matter.

> Immersive simulations powered by AI technology provide students with hands-on learning experiences. These simulations simulate real-world scenarios, allowing students to apply their knowledge, make decisions, and observe the consequences of their actions in a safe and controlled environment. This interactive approach enhances student understanding, critical thinking, and problem-solving skills. (Professor 13)

Cost-Effective Education

Increased affordability and reduced financial barriers have been enabled by use of online AI-integrated platforms. These platforms offer cost- effective education options that have the potential to lower tuition costs and provide savings on commuting and accommodation expenses. By leveraging technology, students can access quality education at a lower cost, making education more accessible and inclusive. As Professor 15 stated:

> The adoption of online AI-integrated platforms in education brings significant cost savings for students. By eliminating the need for commuting and accommodation expenses associated with attending physical campuses, students can access education from the comfort of their homes. This cost-effectiveness reduces financial burdens and expands educational opportunities for a wider range of students.

Remote Learning

AI-powered online platforms have enabled learners to access quality educational opportunities regardless of geographical location, including various remote and underserved areas. AI technology facilitates the delivery of educational content, interactive learning experiences, and collaboration, enabling students to engage in remote learning effectively. This statement from Professor 18 exemplified this theme:

> Remote learning facilitated by AI technology has the potential to bridge the educational divide in underserved areas. By leveraging online platforms, students in remote and underserved locations can access comprehensive educational resources, engage in interactive learning experiences, and collaborate with peers and educators from around the world. This inclusivity in education empowers students who would otherwise face limited educational options.

Data-Driven Insights

AI-generated data has been used to inform teaching methods and improve student learning outcomes. By analyzing data on student progress and curriculum performance, faculties and institutions can gain valuable insights into students' needs and make informed decisions to tailor their teaching methods and improve the curriculum. This data-driven approach enhances the effectiveness of education by aligning it with the specific requirements of students. For instance, Professor 19 stated that "AI-generated data helps faculties tailor teaching methods to meet students' needs. By analyzing student progress, engagement, and learning patterns, educators can provide targeted support and personalize instruction for improved outcomes."

Preparation for the Future

Exposing students to AI-integrated technology equips them with the skills and adaptability needed for a rapidly evolving job market. By focusing on the development of 21st-century skills and technological literacy, education has prepared students to navigate and succeed in a world influenced by AI and emerging technologies. As Professor 17 stated:

> Exposure to AI-integrated technology in education cultivates 21st-century skills in students. Skills such as critical thinking, problem-solving, collaboration, creativity, and adaptability are essential for success in a rapidly changing job market. By engaging with AI and technology, students develop these skills, positioning themselves as adaptable and valuable contributors in the future workforce.

Threats from Students' Use of AI-Integrated Technology

The second research question explored the threats of AI-integrated technology (i.e., ChatGPT) for higher education from educational technologists' perspectives. Interviews were thematically analyzed, revealing eight main themes each consisting of two or more sub-categories. These themes are explained and exemplified below.

Privacy Concerns

The use of AI-integrated technology in education has raises concerns about the privacy of students' data and personal information. As educational institutions collect and analyze student data, it has become crucial to address data security and potential issues related to invasive surveillance. Safeguarding the privacy of students is essential for maintaining trust and ensuring the ethical use of AI technology in educational settings. Professor 12 stated:

> Privacy concerns arise with the use of AI-integrated technology in education, specifically regarding data security. Educational institutions must prioritize robust data protection measures to safeguard students' personal information. This includes adopting encryption protocols, secure storage systems, and strict access controls to prevent unauthorized access or data breaches.

Quality of Content

The proliferation of AI-generated content in education has raised concerns about the variable quality of educational materials, which can impact students' overall learning experience. It is important to address issues related to the accuracy of information and the potential for plagiarism in AI-generated content to ensure that students receive reliable and original educational resources. This statement from Professor 4 exemplified this theme:

> The accuracy of information in AI-generated content is a significant concern in education. While AI can automate content creation, there is a need for careful monitoring and verification to ensure the accuracy and reliability of educational materials. Educators should play an active role in reviewing and curating AI-generated content to maintain high standards of quality.

Digital Dependence

Concerns have been raised that overreliance on AI technology for learning can potentially impact students' critical thinking and problem-solving skills. Reduced analytical thinking and limited creativity may arise from excessive dependence on AI tools and automation in the educational process. This theme was supported by the following statement from Professor 11:

> The overreliance on AI technology in education may lead to reduced analytical thinking among students. When AI tools provide ready-made answers and solutions, students may become accustomed to relying on them rather than engaging in critical thinking and independent problem-solving. It is crucial to strike a balance by encouraging students to think analytically and develop their problem-solving skills alongside the use of AI technology.

Teacher Redundancy

The widespread use of AI as a replacement for human educators has raised concerns about potential reductions in the number of teaching positions, resulting in job displacement and loss of human interaction, an important aspect of the educational experience. As Professor 14 stated:

> The widespread adoption of AI in education may lead to job displacement among educators. As AI technology advances, there is a possibility that certain tasks traditionally performed by teachers could be automated. It is important to find a balance between leveraging AI for efficiency and preserving the valuable role of human educators in providing personalized instruction, mentorship, and guidance.

Technology Gaps

Unequal access to AI-integrated technology and the Internet can exacerbate educational disparities, leading to concerns about the digital divide and limited technological resources in education. This finding is consistent with the statement from Professor 9 that "unequal access to AI-integrated technology exacerbates educational disparities. The digital divide creates inequalities in educational opportunities, limiting students' ability to benefit from AI tools. Bridging this divide is essential for ensuring equitable access to AI-integrated education." Similarly, Professor 15 stated that "limited technological resources contribute to the technology gap in education. Schools with budget constraints struggle to provide AI technology, widening disparities. Allocating resources and promoting equitable access to technology is crucial for addressing this issue."

Social Isolation

The overemphasis on online education in AI-integrated learning may lead to reduced social interaction, which can have implications for students' mental health. Concerns have arisen regarding feelings of loneliness and potential mental health challenges associated with limited social connections. As Professor 14 explained.

> The overemphasis on online education can contribute to feelings of loneliness among students. The absence of face-to-face interactions and limited opportunities for socialization may lead to a sense

of isolation. It is crucial to create spaces and activities that foster social connections and promote a sense of community, even in AI-integrated educational settings.

Ethical Dilemmas

The use of AI technology in education has raised ethical questions, including concerns about algorithmic bias and the ethical use of AI. It is important to address these in order to ensure fairness and ethical decision-making in AI-integrated educational systems.

> Algorithmic bias is a significant ethical concern in AI-integrated education. AI algorithms can inadvertently perpetuate biases in areas such as grading, admissions, or resource allocation, which may result in unfair treatment or discrimination. It is crucial to develop and implement transparent and unbiased algorithms, regularly evaluate their performance, and address any instances of bias to ensure equitable educational opportunities for all students. (Professor 19)

Human-AI Interaction Challenges

The integration of AI technology in education can present challenges for students and faculty members, including difficulties in adapting to AI technology, which may result in inefficiencies and frustration. Two relevant subcategories include user resistance and training requirements.

> User resistance is a common challenge when introducing AI technology in education. Students and faculty members may initially struggle to adapt to new AI tools and processes, leading to resistance and reluctance to fully engage with the technology. It is important to address this resistance by providing clear explanations of the benefits, offering training and support, and fostering a positive learning environment that encourages exploration and experimentation. (Professor 17)

Discussion

Qualitative data analysis revealed 10 key opportunities for AI-powered technologies like ChatGPT in higher education. However, these opportunities come with their own challenges and nuances that require careful consideration. One of the most promising prospects is the potential for greater personalization in education. AI can adapt learning materials and methods to the individual needs of students, revolutionizing the educational experience. Nevertheless, we must consider the lessons learned from national programs to ensure fairness, accountability, and transparency, and to prevent bias and discrimination (Ahmad et al., 2020).

Another beneficial opportunity lies in the improved access to information enabled by AI integration. However, this wealth of online resources has also raised questions about the credibility and quality of the content. Therefore, as Mayer (2014) emphasized, it is essential to provide students with the skills to critically evaluate and recognize reliable information. Efficient feedback mechanisms enabled by AI provide real-time assessment and feedback to improve the learning process. Nevertheless, the ethical use of student data must remain a priority, in line with the findings of Ghallab (2019).

In addition, 24/7 availability through AI technology offers flexibility and convenience. However, a balance must be struck as excessive accessibility can lead to burnout and a so-called always-on culture (Mullikin, 2020). Global collaboration in education enabled by AI can also enrich the learning experience. Nevertheless, it must overcome challenges related to intercultural communication and cultural sensitivity (Gonzalez & Kasper, 1997). Incorporating the development of intercultural competencies into the educational process (Chai, Lin, et al., 2020) is a crucial step. AI-driven interactive learning is a valuable tool that can increase student engagement and motivation. However, thorough research is needed to assess the impact on learning outcomes (Holstein et al., 2018). It is important to ensure that gamification serves the larger educational goals. Low-cost education enabled by AI is a huge opportunity. Nevertheless, careful consideration of financing models and long-term sustainability is required. The challenge is to provide quality education at a lower cost without compromising on quality.

Although AI has clearly played a role in facilitating distance learning, it also needs to address the digital divide and guarantee fair access to technology (Karnouskos, 2022). One of the main objectives is to close the gap between urban and rural areas (Samuel, 2021). The potential for data-driven insights to improve instruction means they should be applied carefully and with a view toward enhancing the learning process for students (Samuel et al. 2022). Educational objectives must be in line with data analytics. Finally, there is an urgent need to prepare students for the labor market of the future. But in addition to technical skills, this preparation should cover social, emotional, and ethical competencies as well (Mishra & Koehler, 2006).

The goal is to create well-rounded people who are prepared for a complicated world—which is undoubtedly a complex challenge. In conclusion, there are many benefits to integrating AI into education; however, for this to happen, ethical, pedagogical, and practical issues must be considered. Proactively tackling these obstacles is necessary to fully realize the benefits of artificial intelligence in higher education, minimize possible hazards, and guarantee optimal learning results.

The second research question, which delved into the perceived threats of AI-integrated technology like ChatGPT in higher education from the perspective of educational technologists, unearthed a complex landscape of concerns. Thematic analysis of interviews with 20 educational technologists revealed eight central themes, each comprising multiple sub-categories, offering a comprehensive picture of the multifaceted challenges that surround AI's role in education.

Privacy concerns emerged as a prominent theme, emphasizing the paramount importance of safeguarding students' personal data and privacy. With educational institutions increasingly collecting and analyzing student data, robust measures like encryption protocols, secure storage systems, and strict access controls become essential to prevent data breaches and unauthorized access. Ahmad et al. (2020) articulated, "privacy concerns arise with the use of AI-integrated technology in education, specifically regarding data security" (p. 4). To address these concerns, transparency, explicit consent, and clear guidelines for data collection and usage must be established. Balancing between leveraging AI for educational purposes and preserving students' privacy rights is essential (Baker et al, 2019).

The second theme related to perceived threats, namely quality of content, was concerned with the authenticity and dependability of educational materials produced by AI. To maintain high standards of quality, concerns regarding the integrity of the information in such content highlighted the necessity for meticulous monitoring and verification. Educators should take an active role in reviewing and selecting AI-generated content. This theme also raised the possibility of plagiarism resulting from AI's ability to generate text. This issue emphasizes the importance of upholding academic integrity and using plagiarism detection software (Mayer, 2014).

The topic of digital dependency highlighted the risks of over-reliance on AI technology for learning. Overdependence can lead to reduced analytical thinking and hinder students' creativity. Encouraging students to think analytically and develop problem-solving skills alongside AI is essential (Holstein et al., 2018). Teacher redundancy, the fourth theme from our second research question, raised concerns about AI's possible replacement of human educators. While AI can automate specific tasks, it cannot replace the unique value that human educators bring to education, such as personalized instruction, mentoring, and guidance. Preserving these irreplaceable human qualities is critical to complement AI technology effectively.

Technology gaps, the fifth theme among the threats raised by participants, highlighted the inequalities that different levels of access to AI-integrated technology and the Internet can cause in education. Bridging these differences will be critical to ensuring that all students have equal access to AI tools (Wilson & Gruzd, 2014). Social isolation, the sixth theme, encompassed concerns arising from an overemphasis on online education through AI-integrated learning. Such an emphasis may reduce student social interactions and lead to mental health impacts. To address these concerns, it is essential to create spaces and activities that promote social connections and a sense of community (Roper, 2007).

Ethical dilemmas formed the seventh theme, encompassing concerns regarding algorithmic bias and the ethical use of AI in education. Algorithmic bias can lead to unfair treatment or discrimination in grading, admissions, or resource allocation. To counteract these biases, transparent and unbiased algorithms and regular performance evaluations are crucial (Ghallab, 2019).

Human-AI interaction challenges encompassed the issues faced by students and faculty members when adapting to AI technology. These challenges have often resulted in resistance and inefficiencies. Overcoming user resistance involves providing clear explanations and training, as well as cultivating a positive learning environment. Proper training through ongoing professional development programs is vital for effective human-AI interaction (Martin et al., 2020).

Conclusions and Implications

In summary, the integration of AI in education presents numerous opportunities but also demands proactive addressing of ethical, pedagogical, and practical considerations to unlock its full potential and ensure the best possible learning outcomes. Key opportunities include (a) enhanced personalization, (b) increased access to information, (c) efficient feedback mechanisms, (d) 24/7 accessibility, (e) global collaboration, (f) interactive learning, (g) cost-effective education, (h) bridging the digital divide, (i) leveraging data insights, and (j) preparing students for the future job market. However, these opportunities come with challenges, such as (a) ensuring fairness, (b) maintaining content quality, (c) ethical use of student data, (d) preventing burnout, (e) considering cross-cultural communication, (f) aligning data analytics with educational goals, and (g) developing well-rounded individuals.

Exploring threats associated with AI-integrated technology in higher education has unearthed a complex landscape of concerns. Privacy concerns emphasize the need to safeguard student data and privacy through encryption and strict access controls. Ensuring transparency and explicit consent is crucial. Quality of content raises worries about the accuracy and potential plagiarism in AI-generated educational materials, highlighting the importance of promoting academic integrity. Digital dependence underscores the risk of overreliance on AI, which may hinder analytical thinking and creativity, necessitating a balance. Teacher redundancy raises concerns about the possible replacement of human educators, emphasizing the unique value they bring to education. Technology gaps underscore inequalities in access, highlighting the need to bridge disparities. Social isolation reveals concerns about reduced social interaction and potential mental health implications, emphasizing the importance of fostering a sense of community. Ethical dilemmas encompass issues of algorithmic bias and ethical AI use, necessitating transparent and unbiased algorithms and ethical guidelines. Finally, human-AI interaction challenges encompass the difficulties faced by students and faculty members when adapting to AI technology and emphasize the need for proper training and a positive learning environment.

Funding

This study was supported via funding from Prince Sattam Bin Abdulaziz University Project Number (PSAU 2024 /R/1445).

References

Ahmad, M. A., Teredesai, A., & Eckert, C. (2020). Fairness, accountability, transparency in AI at scale: Lessons from national programs. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 690–690). https://doi.org/10.1145/3351095.3375690

Baker, T., Smith, L., & Anissa, N. (2019). *Educ-AI-Tion rebooted? Exploring the future of artificial intelligence in schools and colleges*. Nesta. https://www.nesta.org.uk/report/education-rebooted

Chai, C. S., Lin, P. Y., Jong, M. S. Y., Dai, Y., Chiu, T. K., & Huang, B. (2020, August). Factors influencing students' behavioral intention to continue artificial intelligence learning. *2020 International Symposium on Educational Technology* (pp. 147–150). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ISET49818.2020.00040

Chai, C. S., Wang, X., & Xu, C. (2020). An extended theory of planned behavior for the modelling of Chinesesecondary school students' intention to learn. *Artificial Intelligence: Mathematics, 8*(11), 2089. https://doi.org/10.3390/math8112089

Dai, Y., Chai, C. S., Lin, P. Y., Jong, M. S. Y., Guo, Y., & Qin, J. (2020). Promoting students' well-being by developing their readiness for the artificial intelligence age. *Sustainability, 12*(16), 6597. https://doi.org/10.3390/su12166597

De Graaf, M. M., & Allouch, S. B. (2013). Exploring influencing variables for the acceptance of social robots. *Robotics and Autonomous Systems, 61*(12). https://doi.org/10.1016/j.robot.2013.07.007

Dennis, A. R., Fuller, R. M., & Valacich, J. S. (2008). Media, tasks, and communication processes: A theory of media synchronicity. *Management Information Systems Quarterly,32*(2), *575*–600. https://doi.org/10.2307/25148857

Dignum, V. (2021). The role and challenges of education for responsible AI. *London Review of Education. 19*(1), 1–11. https://doi.org/10.14324/LRE.19.1.01

Druga, S., Vu, S. T., Likhith, E., & Qiu, T. (2019). Inclusive AI literacy for kids around the world. *Proceedings of FabLearn 2019* (pp. 104–111). https://doi.org/10.1145/3311890.3311904

Fourtané, S. (2020, August 27). Ethics of AI: Benefits and risks of artificial intelligence systems. *Interesting Engineering*. https://interestingengineering.com/innovation/ethics-of-ai-benefits-and-risks-of-artificial-intelligence-systems

Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology, 101*(3), 705–716. https://doi.org/10.1037/a0014695

Ghallab, M. (2019). Responsible AI: Requirements and challenges. *AI Perspectives, 1*(1), 1–7. https://doi.org/10.1186/s42467-019-0003-z

Gonzalez, C., & Kasper, G. M. (1997). Animation in user interfaces designed for decision support systems: The effects of image abstraction, transition, and interactivity on decision quality. *Decision Sciences, 28*(4), 793–823. https://doi.org/10.1111/j.1540-5915.1997.tb01332.x

Holstein, K., McLaren, B. M., & Aleven, V. (2018). Student learning benefits of a mixed reality teacher awareness tool in AI-enhanced classrooms. *International Conference on Artificial Intelligence in Education* (pp. 154–168). Springer. https://doi.org/10.1007/978-3 319-93843-1_12

Howlin, C., & Lynch, D. (2014, November). A framework for the delivery of personalized adaptive content. In *2014 International Conference on Web and Open Access to Learning* (pp. 1–5). Institute of Electrical and Electronics Engineers. doi:10.1109/ICWOAL.2014.7009203

Karnouskos, S. (2022). Symbiosis with artificial intelligence via the prism of law, robots, and society. *Artificial Intelligence Law, 30*(1), 93–115. https://doi.org/10.1007/s10506-021-09289-1

Kasper, G. M. (1996). A theory of decision support system design for user calibration. *Information Systems Research*, *7*(2), 215–232. https://doi.org/10.1287/isre.7.2.215

Kentnor, H. E. (2015). Distance education and the evolution of online learning in the United States. *Curriculum and Teaching Dialogue*, *17*(1), 21–34. https://digitalcommons.du.edu/law_facpub/24/

Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: The Betty's Brain system. *International Journal of Artificial Intelligence in Education*, *18*(3), 181–208. https://content.iospress.com/articles/international-journal-of-artificial-intelligence-in-education/jai18-3-02

López-Fernández, H., Fdez-Riverola, F., Reboiro-Jato, M., Glez-Peña, D., & Méndez, J. R. (2011). Using CBR as design methodology for developing adaptable decision support systems. In J. Chiang (Ed.), *Efficient decision support systems: Practice and challenges from current to future* (pp. 124–144). IntechOpen. https://doi.org/10.5772/16923

Martin, F., Sun, T., & Westine, C. D. (2020). A systematic review of research on online teaching and learning from 2009 to 2018. *Computers & Education*, *159*(4), 104009. https://doi.org/10.1016/j.compedu.2020.104009

Mayer, R. E. (2014). Cognitive theory of multimedia learning. In *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge University Press.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, *267*(1), 1–38. https://doi.org/10.1016/j.artint.2018.07.007

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Mullikin, J. (2020). *Focusing on feedback in distance learning*. Edutopia. Retrieved August 29, 2022, from https://www.edutopia.org/article/focusing-feedback-distancelearning

Neumann, M. M., & Herodotou, C. (2020). Evaluating YouTube videos for young children. *Education and Information Technologies*, *25*(4), 4459–4475. https://doi.org/10.1007/s10639-020-10183-7

Panch, T., Szolovits, P., & Atun, R. (2018). Artificial intelligence, machine learning and health systems. *Journal of Global Health, 8*(2), 1-8. https://doi.org/10.7189/jogh.08.020303

Pelaez, A., Jacobson, A., Trias, K., & Winston, E. (2022). The Turing teacher: Identifying core attributes for AI learning in K–12. *Frontier Artificial Intelligence*. *5*(1), 1031450. https://doi.org/10.3389/frai.2022.1031450

Phillips-Wren, G. (2012). AI tools in decision making support systems: A review. *International Journal of Artificial Intelligence Tools, 21*(2). https://doi.org/10.1142/S0218213012400052

Reed, D. E., Kaplita, E. C., McKenzie, D. A., & Jones, R. A. (2022). Student experiences and changing science interest when transitioning from K–12 to college. *Education Sciences*, *12*(7), 496. https://doi.org/10.3390/educsci12070496

Roper, A. R. (2007). How students develop online learning skills. *EDUCAUSE Quarterly*, *30*(1), 62–65. https://er.educause.edu/articles/2007/2/how-students-develop-online-learning-skills

Samuel, J. (2021). *A call for proactive policies for informatics and artificial intelligence technologies*. Scholars Strategy Network. Retrieved October 2, 2022, from https://scholars.org/contribution/call-proactive-policies-informatics-and

Samuel, J., Kashyap, R., Samuel, Y., & Pelaez, A. (2022). Adaptive cognitive fit: Artificial intelligence augmented management of information facets and representations. *International Journal of Management, 65*(4), 102505. https://doi.org/10.1016/j.ijinfomgt.2022.102505

Samuel, J., Navazali, G. G., Rahman, M., Esawi, E., & Samuel, Y. (2020). COVID-19 public sentiment insights and machine learning for Tweets classification. *Information*, *11*(6), 314–330. https://doi.org/10.3390/info11060314

Sankar, C. S., Ford, F. N., & Bauer, M. (1995). A DSS user interface model to provide consistency and adaptability. *Decision Support Systems,13*(1), 93–104. https://doi.org/10.1016/0167-9236(93)E0033-A

Turban, E. (1995). *Decision support and expert systems management support systems*. Prentice-Hall.

Ulfert, A. S., Antoni, C. H., & Ellwart, T. (2022). The role of agent autonomy in using decision support systems at work. *Computers in Human Behavior*, *126,* 106987. https://doi.org/10.1016/j.chb.2021.106987

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology, 45*(3), 223–235. https://doi.org/10.1080/17439884.2020.1798995

Wilson, L., & Gruzd, A. (2014). MOOCs: International information and education phenomenon? *Bulletin of the American Society for Information Science and Technology, 40*(5), 35–40. https://doi.org/10.1002/bult.2014.1720400510

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

Teacher- Versus AI-Generated (Poe Application) Corrective Feedback and Language Learners' Writing Anxiety, Complexity, Fluency, and Accuracy

Dan Wang
School of Foreign Studies, East China University of Political Science and Law, ShangHai, China

Abstract

This study examines the effects of corrective feedback (CF) on language learners' writing anxiety, writing complexity, fluency, and accuracy, and compares the effectiveness of feedback from human teachers with an AI-driven application called Poe. The study included three intact classes, each with 25 language learners. Using a quasi-experimental design with pretest and posttest measures, one class received feedback from the teacher, one from the Poe application, and the third received no response to their writing. Data were generated though tests and a writing anxiety scale developed for the study. Data analysis, conducted using one-way ANOVA tests, revealed significant effects of teacher and AI-generated feedback on learners' writing anxiety, accuracy, and fluency. Interestingly, the group that received AI-generated feedback performed better than the group that received teacher feedback or no AI support. Additionally, learners in the AI-generated feedback group experienced a more significant reduction in writing anxiety than their peers. These results highlight the remarkable impact of AI-generated CF on improving writing outcomes and alleviating anxiety in undergraduate language learners at East China University of Political Science and Law . The study demonstrates the benefits of integrating AI applications into language learning contexts, particularly by promoting a supportive environment for students to develop writing skills. Educators, researchers, and developers can use these findings to inform pedagogical practices and technological interventions to optimize the language learning experience in primary school settings. This research highlights the effectiveness of AI-driven applications in language teaching. It highlights the importance of considering learners' psychological well-being, particularly anxiety levels, when developing effective language learning interventions.

Keywords: artificial intelligence, AI, corrective feedback, writing anxiety, complexity, fluency, accuracy

Introduction

Providing written feedback to correct errors made by language learners has long been a fundamental practice in teaching writing, capturing considerable attention in writing research in Second Language (L2). Given the multifaceted and complex nature of writing, feedback encompasses a broad array of responses, offering insights into the accuracy, successful communication, and content of learners' expressions or discourse (Li & Vuono, 2019; Thi & Nikolov, 2021). Pedagogically, feedback is a crucial link between assessment and teaching, providing appropriate information about the language learners' correct performance and guidance to achieve the target learning goals. Consequently, considerable focus has been directed toward understanding the significant contribution of CF to the language learners' writing performance. Written Corrective Feedback (WCF) plays a pivotal role in enhancing learners' writing performance, accuracy, fluency, organizational skills, and task achievement (Karim & Nassaji, 2018, 2019; Leeman, 2010; Lim & Renandya, 2020; Liu & Brown, 2015; Liu & Huang, 2020; Luo & Liu, 2017).

Numerous studies have investigated the impact of written CF in the writing ability of foreign language learners, comparing various types of feedback (Han & Hyland, 2015; Truscott, 2010; Truscott & Hsu, 2008). The findings from these studies have been synthesized through quantitative and qualitative systematic reviews, commonly known as meta-analyses. Earlier meta-analyses, including the work of Russell and Spada (2006), underscored the significant contribution of CF to the development of grammatical knowledge in language learners.

In subsequent analyses conducted by Hyland and Hyland (2006), the emphasis was on the diversity evident in student populations, writing genres, feedback practices, and research designs. Notably, Truscott (2010) posited a negative impact of error correction on students' ability to accurately report information, drawing this conclusion from an examination of twelve published studies. In later meta-analyses conducted by researchers (e.g., Biber et al., 2011; Kang & Han 2015; Liu & Brown, 2015; Sia & Cheung, 2017), valuable perspectives were offered regarding the efficacy of written CF. These analyses considered individual differences among learners and addressed methodological limitations in the existing literature.

Recent meta-analyses by Lim and Renandya (2020) presented evidence supporting the potential of written WCF to enhance L2 writing skills of English as a Foreign Language (EFL) learners, with a specific focus on improving grammatical accuracy. However, ongoing debates persist, addressing questions about the extent of the benefits derived from WCF and the sustained efficacy of various feedback treatments, particularly in comparing implicit and explicit approaches. The existing body of literature on CF in language learning predominantly revolves around traditional teacher-generated feedback, with limited exploration into the consequences of feedback generated by AI. Notably, there is a restricted investigation into the impact of AI-generated feedback, primarily through applications, highlighting a gap in research in this area. While studies have examined the effectiveness of feedback on aspects such as accuracy, complexity, and anxiety, there is a noticeable gap in the research regarding a direct comparison between teacher-generated and AI-generated CF.

One specific instance of an AI-powered application is the Personalized Online Experience (Poe). This application integrates AI technologies to tailor online interactions based on individual user behavior, preferences, and historical data. The Poe application optimizes content recommendations, user interfaces, and overall interaction design by continuously learning from user engagement. Through machine-learning algorithms, Poe evolves to anticipate user needs, delivering a more personalized and efficient online experience. This enhances user satisfaction and exemplifies the potential of AI-powered applications to revolutionize how we interact with digital platforms, creating a more intelligent and user-centric digital landscape.

Understanding these two feedback sources' potential differences and implications is essential for informing pedagogical practices and optimizing language learning experiences. The rationale for this study stems from the increasing integration of AI technologies in language education and the need to evaluate their efficacy compared to traditional teaching methods. With the emergence of AI applications such as Poe, which claim to provide nuanced and personalized corrective feedback, it is crucial to assess their impact on language learners. This study investigates how learners respond to feedback from AI applications compared to feedback from human teachers, specifically regarding writing anxiety, writing complexity, and accuracy. The rationale behind this comparative analysis lies in the potential benefits and drawbacks of AI-generated feedback, which may differ from the interpersonal and contextual aspects associated with teacher-generated feedback. By addressing this gap, the research seeks to contribute valuable insights into the evolving landscape of language education. More specifically, this study attempts to answer the following questions:

1. Do AI-generated (through Poe application) and teacher corrective feedbacks equally reduce the EFL learners' writing anxiety?

2. Do AI-generated (through Poe application) and teacher corrective feedbacks equally foster EFL learners' writing fluency?

3. Do AI-generated (through Poe application) and teacher corrective feedbacks equally foster the EFL learners' writing accuracy?

4. Do AI-generated (through Poe application) and teacher corrective feedbacks equally foster the EFL learners' writing complexity?

Literature Review

Empirical considerations do not solely drive the research on WCF; it is also underpinned by theoretical frameworks that highlight its potential contributions to L2 development (Polio, 2012). Skill-acquisition theories, such as DeKeyser's (2007), underscore the importance of practice and explicit instruction in developing accuracy, a concept aligned with WCF's role in aiding learners to store and retrieve declarative knowledge. The theoretical foundations of the noticing hypothesis (Schmidt, 2012) and the interaction hypothesis (Long, 1980) further support WCF by helping learners identify gaps in interlanguage by providing the needed evidence.

In L2 writing, the distinction between corrective and non-corrective feedback, focusing on form and content, is emphasized (Luo & Liu, 2017; Zhang, 2021). Corrective feedback targets language learning by providing negative evidence to enhance accuracy, while non-corrective feedback addresses broader aspects such as content, organization, and linguistic performance. The role of WCF in L2 writing goes beyond traditional written commentary feedback, with feedback strategies ranging from direct to indirect and metalinguistic (Ellis, 2009a, 2009b, 2009c). The choice between comprehensive and focused feedback treatments has been explored, with recent studies highlighting the benefits of focused feedback. However, some still argue for a comprehensive approach (Benson & DeKeyser, 2018).

Considering the significance of writing tasks in L2 writing, factors such as task types and complexity play a vital role (Liu & Huang, 2020). Both unfocused and focused writing tasks are used for evaluating the language learners' writing proficiency. The impact of different writing task genres on language use, each with distinct communicative and functional requirements, contributes to a deeper understanding of how diverse writing tasks influence linguistic performance, encompassing aspects such as accuracy and complexity (Polio & Yoon, 2018).

Research into WCF is empirically driven and rooted in theoretical frameworks exploring its potential contributions to L2 development (Zhang, 2021). Skill-acquisition theories, like DeKeyser's (2007), posit that accuracy results from practice, explicit instruction, and extensive practice, prerequisites for transforming declarative knowledge into procedural knowledge. The concept of written CF is in harmony with these concepts, with the objective of aiding learners in storing and retrieving declarative knowledge, specifically explicit knowledge related to the target language. Stressing the importance of both practice and feedback, Evans et al. (2014) and Hartshorn and Evans (2015) highlighted automatization's crucial role within the skill acquisition theory framework.

In L2 writing, scholars highlight the role of WCF in fostering students' writing skills and abilities. The crucial distinction between corrective and non-corrective feedback, which focuses on form and content, is essential (Long, 1980; Luo & Liu, 2017). Corrective feedback aims at negative evidence to promote learning the target language, specifically addressing accuracy. Conversely, non-corrective feedback provides commentary on broader aspects, including organization, linguistic performance, and format. Exploring different types of WCF beyond traditional written commentary feedback is gaining interest.

Empirical investigations into WCF examine its facilitative role through the comparison of feedback strategies against no-feedback conditions (Ellis, 2009a, 2009b, 2009c; Kurzer, 2018) and assess the relative effectiveness of various feedback strategies (Riazantseva, 2012). Feedback interventions delineate between comprehensive and focused approaches, determining the extent of WCF provided to students. While earlier studies leaned towards comprehensive error correction, recent research underscores the advantages of focused feedback. Scholars (e.g., Benson & DeKeyser, 2018; Stefanou & Révész, 2015) propose that correcting the errors in a focused way yields more significant benefits than addressing all errors indiscriminately. However, certain studies advocate for comprehensive feedback that addresses a range of errors rather than concentrating on a specific sort of error (Bonilla López et al., 2018).

Ellis (2009c) and Robinson (2011) advocated for focusing on meaning in tasks, classifying them as unfocused or focused, based on general language use or specific linguistic features in L2 writing, in which both unfocused tasks and focused writing tasks serve as assessments of learners' proficiency. They also believe that recognizing how task demands influence L2 writing is paramount because tasks establish contexts which provide opportunities to uptake CF. While the impact of cognitive demands imposed by tasks on learners' accuracy has received limited attention in WCF research, empirical investigations into various genres of writing tasks reveal distinct communicative and functional requirements.

Studies on AI

AI constitutes a domain with a rich historical and philosophical background (Bozkurt et al., 2023; Cao, 2023). Its evolution has raised fundamental inquiries about machine cognition and the capacity for independent creativity beyond programmed instructions (Kurzweil, 2014; Winterson, 2022, pp. 9–32). These inquiries led to the adoption of the concept of AI technologies (Benavides et al., 2020; Bozkurt et al, 2023; Winterson, 2022, pp. 9–32).

Despite these strides, AI has become so profoundly integrated into daily life that there are expectations of an era where human and artificial intelligence converge (Kurzweil, 2014). The ubiquitous influence of AI extends to communication and advisory roles across various professions, including media, accounting, and copywriting (Bozkurt et al., 2021). Since the inception of computerized AI, educators have expressed concerns about the potential obsolescence of their roles (Goksel & Bozkurt, 2019; Selwyn et al., 2023). More recently, apprehensions have emerged regarding the possibility of students completing assignments or responding to questions with undetectable AI assistance, raising concerns about academic integrity and the authenticity of students' work (Diebold, 2023; Luan et al., 2020; Ouyang et al., 2022).

The systematic exploration of AI technologies in educational settings predominantly revolves around their application for forecasting learner outcomes and behaviors which establish adaptive learning environments, improve academic performance, and enhance overall learning achievements and experiences (Chu et al., 2022; Zawacki-Richter et al., 2019). Recent examinations of literature in K–12 education reveal a broadening range of applications for artificial intelligence in education (AIED), including collaborative learning, modeling approaches, and visualization. This signifies a shift beyond conventional pedagogical methods, as noted by Humble and Mozelius (2022) and Zawacki-Richter et al. (2019). However, the multifaceted adoption of AI in education necessitates a comprehensive understanding of its potential implications within the broader social, cultural, pedagogical, and organizational contexts. Despite the potential advantages of integrating AI into education, numerous persistent challenges and ethical considerations warrant attention. These challenges encompass attitudes toward AI, educators' proficiency in effectively using technological tools, ethical concerns, and various technological hurdles (Sharma et al., 2019).

Ethical considerations form a cornerstone in the discourse surrounding AI in education. Scholars argue for a comprehensive examination of ethical concerns, including privacy issues and data ownership, before the widespread adoption of AI technologies (Humble & Mozelius, 2022). The potential influence of major ed-tech organizations over educational institutions raises additional ethical questions, particularly regarding privacy and corporate control, as these organizations may have access to student and staff data for corporate gains (Bozkurt et al., 2021).

Moreover, the existing literature on AI in education often consists of descriptive studies, indicating a need for a robust theoretical foundation to propel the field forward (Chen et al., 2020). Establishing a theoretical framework is essential for advancing our understanding of the implications of AI in educational settings and guiding future research and implementation strategies. Therefore, a concerted effort to address these challenges and ethical considerations is imperative to harness responsibly AI's full potential in education.

In a focused examination of the AI-powered chatbot ChatGPT from an educational standpoint, Tlili et al. (2023) supported the use of ChatGPT in education. They advocated for a new teaching philosophy to effectively integrate AI-powered technologies into education, emphasizing the importance of responsible, humanized chatbots and the development of digital literacy competencies (Ng et al., 2021). Concerns about academic integrity have prompted discussions on AI tools' ethical and responsible use in education (Cox, 2021), with calls for updated policies and strategies. Researchers, instructors, and policymakers are cautioned to proactively address potential disruptions caused by integrating AI technologies (Tang et al., 2021). Additionally, it is acknowledged that novel assessment formats emphasizing creativity and critical thinking, areas where AI cannot entirely replace human judgment, may be essential (Dogan et al., 2023).

Methodology

Sample and Procedure

The study involved three intact classes of language learners enrolled in online courses at the School of Foreign Studies, East China University of Political Science and Law, China. Each class consisted of 25 members. As both researcher and instructor, I recruited participants from these classes, all taking a writing course as part of their language curriculum. In order to homogenize the language learners based on writing accuracy, fluency, and complexity, a writing test comprising three tasks was administered to the entire pool of participants. Following the initial evaluation, participants were kept on for the treatment; however, the final analysis only included those whose writing test results were within the range of +/- 1 standard deviation (SD) from the mean. The purpose of implementing this criterion was to guarantee a study group that was relatively homogeneous. The analysis comprised 75 language learners' pretest and posttest results, representing those who satisfied the predetermined requirements. All participants were native speakers of Mandarin, with English being their second language. The study focused on this specific population to explore the impact of the proposed treatment on the writing skills of Chinese English language learners at the University of X. In the initial phase of the study, three intact classes underwent a pre-assessment involving the administration of the Writing Anxiety Scale and a Writing Test. The classes were randomly assigned to three groups for the subsequent intervention. The first group received corrective feedback from the teacher, focusing on addressing issues in their written work. In the second group, students were trained to use the Poe application. This involved submitting their writings to Poe and requesting revisions, edits, and paraphrasing, specifically emphasizing grammar, accuracy, coherence, and complexity. The third group received no corrective written feedback during this intervention period.

Weekly assignments that matched the course material were given to participants, with the understanding that their submissions would be evaluated for correctness, organization, and content. Punctuation, spelling, and recognizable grammatical errors that might obstruct clear communication were all evaluated. Assignments from the start, middle, and end of the semester were chosen for comparison in order to assess any changes in the intricacy, accuracy, and fluency of participants' work over the course of the semester. For this study, three assignments were selected from each participant, for a total of 225 assignments analyzed. All groups engaged in their specific interventions over the course of 14 sessions. The same Writing Anxiety Scale and Writing Test were used for the post-assessment in all three intact classes after this intervention. After the data was gathered, it was analyzed to look for variations in the groups' performance and anxiety levels when writing.

Data Analysis

A T-unit is characterized as one main clause along with any subordinate clauses that are attached to or embedded within it. The classification of clauses involves the distinction between dependent and independent clauses, with an independent clause being self-sufficient. In contrast, a dependent clause, which includes adverbial, nominal, and adjectival clauses, comprises a finite verb and a subject (Wolfe-Quintero,1998). Following the frameworks proposed by Storch (2009), the evaluation of complexity was conducted through the examination of the ratios of clauses per T-unit (C/T) and dependent clauses per T-unit (DC/T).

The assessment of accuracy took into account the proportion of error-free T-units (EFT/T), the proportion of error-free clauses (EFC/C), and the total number of errors per total number of words (E/W). Errors were categorized into syntactic errors (e.g., word order, incomplete sentences), morphological errors (e.g., tense, agreement, use of articles), and errors in word choice. Notably, spelling and mechanical errors such as punctuation were excluded from consideration. Fluency metrics included the total number of words (W), the count of T-units, and the length of T-units measured in words per T-unit (W/T).

An additional coder was engaged to ensure coding reliability. The inter-coder reliability achieved high scores of .92 for T-unit identification and .97 for clause identification. Regarding the identification of error-free clauses and T-units, the reliability scores were .91 and .93, respectively. The data analysis involved employing a one-way analysis of variance (ANOVA) to compare the scores of the three groups on writing accuracy, fluency, complexity, and writing anxiety tests. This method was chosen to determine whether there were any statistically significant differences among the means of the three independent groups. Before delving into ANOVA, key assumptions, including normality and homogeneity of variances, were thoroughly examined to ensure the reliability of the subsequent results. The null hypothesis assumed no significant differences existed between the group means, while the alternative hypothesis posited that at least one group mean differed.

Post hoc tests, Bonferroni, were then employed to pinpoint specific group differences if the ANOVA results were significant. This ensured that effect size measures, such as eta-squared or omega-squared, could be computed to offer insights into the practical significance of the observed differences. The same one-way ANOVA procedure was applied to posttest scores, allowing for an examination of changes or improvements within each group over the intervention period.

Results

Research Question 1

The first research question delved into how different corrective written feedbacks—teacher, AI generated, and no correction—affect language learners' anxiety levels in writing. The data analysis yielded noteworthy results, presented in tables 1 and 2.

Table 1

ANOVA for Groups' Scores on Writing Anxiety

Variable	Correction type	*N*	*M*	*F*	*p*	η^2
Writing anxiety	AI generated	25	1.60	53.54	.001	0.70
	Teacher	25	2.10			
	No correction	25	2.70			

Table 1 shows the mean scores for writing anxiety differed significantly across the three groups. The AI-generated group (Poe group) exhibited the lowest anxiety levels, with a mean score of 1.60, followed by the teacher group at 2.10 and the no-correction group at 2.70. The effect size (η^2) of 0.70 suggests a significant impact. The findings in Table 2 reveal a significant reduction in writing anxiety among learners who received AI-generated CF compared to those receiving teacher-generated feedback or no correction (p = .001).

Table 2

Bonferroni for Comparisons Between the Groups' Writing Anxiety

Dependent variable	(I) Correction type	(J) Correction type	Mean difference
Writing anxiety	AI generated	Teacher	-.50
		No correction	-1.6
	Teacher	No correction	-.60

Note. I = x; J = x.
p = .001

Research Question 2

The second research question centered on a comparison of the writing fluency of students in the three intact classes. Results are presented in tables 3 and 4.

Table 3

ANOVA for Groups' Scores on Writing Fluency

Variable	Correction type	*N*	*M*	*F*	*p*	η^2
EFC/C	AI generated	25	0.8	3.56	.001	0.75
	Teacher	25	0.6			
	No correction	25	0.4			
EFT/T	AI generated	25	0.84	4.21	.001	0.61
	Teacher	25	0.64			
	No correction	25	0.40			
E/W	AI generated	25	0.85	6.25	.001	0.51
	Teacher	25	0.62			
	No correction	25	0.41			

Note. EFC/C= proportion of error-free clauses, FFT/T= error-free T-units, E/W= the total number of errors per total number of words.

Concerning fluency, a one-way analysis of variance (ANOVA) revealed noteworthy distinctions in the proportion of error-free T-units (F (2, 72) = 3.56, p < .05, η^2 = 0.71). Upon a more detailed investigation into the proportion of error-free clauses, significant differences emerged among distinct groups (F (2, 72) = 4.12, p < .05, η^2 = 0.61). The results presented in Table 4 from pairwise comparisons underscored that both AI-generated and teacher-generated CF exhibited significantly higher proportions of error-free clauses compared to the no-correction group. Furthermore, the total number of errors per total number of words exhibited notable differences between groups (F (2, 72) = 6.25, p < .05, η^2 = .51). Subsequent pairwise comparisons elucidated that the error rate within the AI-generated feedback group was lower than that within the teacher-generated feedback group. Additionally, the error rate in the teacher-generated group was lower than in the no-correction group. As a result, these findings suggest an improvement in syntactic accuracy across the various feedback groups.

Table 4

Bonferroni for Comparisons Between the Groups' Writing Fluency

Dependent variable	(I) Correction type	(J) Correction type	Mean difference (I-J)
EFC/C	AI generated	Teacher	0.20
		No correction	0.40
	Teacher generated	No correction	0.20
EFT/T	AI generated	Teacher	0.24
		No correction	0.44
	Teacher	No correction	0.20
E/W	AI generated	Teacher	0.23
		No correction	0.34
	Teacher	No correction	0.21

Note. EFC/C= proportion of error-free clauses, FFT/T= error-free T-units, E/W= the total number of errors per total number of words
$p = .001$

Research Question 3

The third research question concerned the writing complexity of students in the three intact classes. Results are presented in tables 5 and 6.

Table 5

ANOVA for Groups' Scores on Writing Complexity

Variable	Correction type	*N*	*M*	*F*	*p*	η^2
CT	AI generated	25	2.5	5.62	.001	0.62
	Teacher	25	2.00			
	No correction	25	1.60			
DCT	AI generated	25	1.20	4.95	.001	0.62
	Teacher	25	0.80			
	No correction	25	0.60			

Note. CT = the ratios of clauses per T-unit; DCT = dependent clauses per T-unit

Table 5 presents discernible variations in the ratio of clauses per T-unit ($F(2, 75) = 5.62$, $p < .05$, $\eta^2 = 0.62$) and the ratio of dependent clauses per T-unit ($F(2, 72) = 4.95$, $p < .05$, $\eta^2 = 0.62$) among the three cohorts. A more in-depth examination, facilitated by a post hoc analysis (as outlined in Table 6), provides additional evidence affirming the superior performance of AI-generated CF over teacher-generated CF. Furthermore, the teacher-generated feedback, in comparison, demonstrates higher effectiveness when contrasted with the no-correction group. These findings shed light on the nuanced impact of different feedback approaches on the syntactic structure, offering valuable insights into the intricate dynamics of language learning and correction methods.

Table 6

Bonferroni for Comparisons Between the Groups' Writing Complexity

Dependent variable	(I) Correction type	(J) Correction type	Mean difference (I-J)
CT	AI generated	Teacher	0.50
		No correction	0.90
	Teacher generated	No correction	0.40
DCT	AI generated	Teacher	0.60
		No correction	0.40
	Teacher	No correction	0.20

Note. CT = the ratios of clauses per T-unit; DCT = dependent clauses per T-unit
$p = .001$

As seen in Table 6, the mean differences in scores were computed for each combination of correction methods, revealing significant differences across all comparisons ($p < .001$). In the CT, AI-generated correction yielded the highest mean difference of 0.50 compared to teacher-generated with no correction (0.40). In the DCT, AI-generated correction also resulted in the highest mean difference of 0.60, followed by teacher-generated correction at 0.20 and no correction at 0.40. These findings suggest that AI-generated correction is more effective than teacher-generated or no correction in improving learners' performance in writing complexity.

Research Question 4

The fourth research question concerned comparing the writing fluency of students in the three intact classes. Results are presented in tables 7 and 8.

Table 7

ANOVA for Groups' Scores on Writing Fluency

Variable	Correction type	*N*	*M*	*F*	*p*	η^2
Total of T-units	AI generated	25	31	80.149	.001	0.56
	Teacher	25	28			
	No correction	25	24			
Total of words	AI generated	27	750	95.38	.000	0.61
	Teacher	25	640			
	No correction	35	590			
W/T	AI generated	27	19.50	51.64	.000	0.51
	Teacher	25	17.30			
	No correction	35	15.10			

Note. W/T= the length of T-units measured in words per T-unit

The findings of the study reveal significant differences across the three types of corrections (AI generated, teacher generated, and no correction) regarding various linguistic variables. The total number of T-units produced by participants under the AI-generated correction condition was significantly higher (M = 31) compared to the teacher correction condition (M = 28) and the no-correction condition (M = 24). This difference was statistically significant (F = 80.149, $p < .001$), indicating a substantial impact of the correction method on the overall syntactic structure. Similarly, the total number of words in the AI-generated correction condition (M = 750) surpassed those in the teacher correction condition (M = 640) and the no-correction condition (M = 590), with a significant overall difference (F = 95.38, $p < .001$). This suggests that AI-generated corrections influenced the participants to produce more words in their writing.

The words per T-unit (W/T) ratio significantly differed among the three conditions. Participants in the AI-generated correction condition had a higher W/T ratio (M = 19.50) compared to the teacher correction condition (M = 17.30) and the no-correction condition (M = 15.10). This difference was statistically significant (F = 51.64, $p < .001$), indicating that AI-generated corrections influenced the number of words and the distribution of words within T-units. Results of the post hoc test (Table 8) also verified that the differences between the AI-generated and teacher-generated CF on three aspects of writing fluency were statistically significant (p = .001), favoring the AI-generated feedback group. The writing fluency of the no-correction group was significantly lower than the writing fluency of the teacher-generated corrective feedback (p = .001).

Table 8

Bonferroni for Comparisons Between the Groups' Writing Fluency

Dependent variable	(I) Correction type	(J) Correction type	Mean difference (I-J)
Total of T-units	AI generated	Teacher	3.00
		No correction	7.00
	Teacher generated	No correction	4.00
Total of words	AI generated	Teacher	110
		No correction	150
	Teacher	No correction	50
W/T	AI generated	Teacher	2.20
		No correction	4.40
	Teacher	No correction	2.20

Note. p = .001

Discussion

Incorporating AI into educational settings has become a focal point of scholarly inquiry, with researchers delving into its potential to augment learning outcomes. This discourse consolidates insights derived from a quasi-experimental study examining the influence of AI-generated CF on writing anxiety, fluency, accuracy, and complexity among EFL learners. The study systematically compares the efficacy of AI-generated feedback against feedback provided by teachers. The results are contextualized within the broader literature on AI in education, the digital transformation in higher education, and the overarching domain of second language acquisition**.** The results indicate that both AI-generated and teacher-provided feedback significantly affected the language learners' writing accuracy, fluency, and complexity.

Interestingly, AI-generated feedback proves to be more effective than teacher-generated feedback. This aligns with the broader discourse on the efficacy of AI in education, as discussed by Bozkurt et al. (2021) and Chu et al. (2022). These studies emphasized the transformative potential of AI in enhancing educational practices and suggested that AI could provide personalized and timely feedback, addressing individual learning needs.

A noteworthy outcome of the study is the reduction in learners' writing anxiety facilitated by teacher and AI-generated feedback. This finding resonates with the work of Ellis (2009a), who highlighted the importance of feedback in creating a supportive learning environment and reducing learners' anxiety. The study contributes to the growing body of research that recognizes the emotional aspects of language learning and emphasizes the role of technology, including AI, in fostering a positive learning experience (Han & Hyland, 2015).

The comparison between AI- and teacher-generated feedback draws attention to the unique advantages of AI, as evidenced by the study's results. The AI system used in the research, the Poe application, outperformed human teachers in enhancing writing skills and reducing anxiety. This aligns with the findings of Bonilla López et al. (2018) who investigated the differential effects of feedback forms in second-language writing. The discussion here underscores the potential of AI to provide consistent and objective feedback, addressing some limitations associated with human feedback, such as variability and subjectivity.

The theoretical underpinnings of the study draw support from skill acquisition theory (DeKeyser, 2007) and the task complexity framework (Robinson, 2011). Skill acquisition theory underscores the importance of practice and feedback in language learning, aligning with the study's focus on corrective feedback's impact on writing skills. Insights from the task complexity framework shed light on how cognitive demands embedded within writing tasks influence learners' language development, providing a valuable perspective for interpreting outcomes related to complexity.

Beyond its theoretical contributions, this study enriches the ongoing discourse on AI in education. In line with trends discussed by Bozkurt et al. (2023) and Chen et al. (2020), it emphasizes the need for a nuanced understanding of AI's role in education, considering both practical applications and theoretical implications. Bozkurt et al.'s (2023) systematic review and exploration of speculative futures for ChatGPT contributed to the broader dialogue on responsibly integrating generative AI into education. Therefore, this study not only provides insights into language learning dynamics but also aligns with and extends the broader conversation on the integration of AI into educational contexts.

Implications and Conclusions

The findings of this study hold practical implications for language educators and policymakers. Integrating AI-generated feedback systems, such as the Poe application, into language classrooms could enhance the quality and efficiency of feedback provision. However, as discussed by Kurzweil (2014) and Tlili et al. (2023), ethical considerations should guide the responsible implementation of AI in education. Teachers may need to adapt their roles to incorporate AI as a supportive tool rather than a replacement. In conclusion, the quasi-experimental study on the impact of AI-generated corrective feedback on EFL learners' writing skills contributes valuable insights to the evolving landscape of AI in education. The effectiveness of AI in fostering writing accuracy, fluency, and complexity, and reducing anxiety positions it as a promising tool for language learning. However, the responsible integration of AI into educational practices requires a thoughtful and ethical approach. This discussion bridges the study's findings with existing literature, providing a comprehensive understanding of the implications for language education in the digital transformation era.

Limitations and Suggestions

While the study presents compelling insights into the efficacy of AI-generated corrective feedback, certain limitations warrant consideration. First, while valuable for initial exploration, the quasi-experimental design may lack the robustness of a randomized controlled trial. The study's focus on one specific AI application, Poe, raises questions about the generalizability of findings to other AI platforms. Moreover, the study primarily gauges short-term impacts, leaving the long-term effects of AI-generated feedback on language acquisition unexplored. Additionally, the absence of qualitative data may limit a nuanced understanding of learners' perceptions and experiences with AI feedback. Future research could employ mixed-methods approaches, incorporating qualitative insights to complement quantitative findings. Furthermore, the study needs to delve deeper into the sociocultural aspects influencing the reception of AI in diverse educational contexts, an avenue ripe for exploration. Addressing these limitations will enhance the robustness and applicability of research on AI in language education.

Funding

This work was sponsored by the Chinese Fund for the Humanities and Social Sciences (Chinese Academic Translation Project of the National Social Science Foundation of China) (22WZSB037).

References

Benavides, L. M. C., Tamayo Arias, J. A., Arango Serna, M. D., Branch Bedoya, J. W., & Burgos, D. (2020). Digital transformation in higher education institutions: A systematic literature review. *Sensors, 20*(11), Article 3291. https://doi.org/10.3390/s20113291

Benson, S., & DeKeyser, R. (2018). Effects of written corrective feedback and language aptitude on verb tense accuracy. *Language Teaching Research, 23*(6), 702–726. https://doi.org/10.1177/1362168818770921

Biber, D., Nekrasova, T., & Horn, B. (2011). The effectiveness of feedback for L1-English and L2-writing development: A meta-analysis. *ETS Research Report Series*, *2011*(1), i–99.https://doi.org/10.1002/j.2333-8504.2011.tb02241.x

Bonilla López, M., Van Steendam, E., Speelman, D., & Buyse, K. (2018). The differential effects of comprehensive feedback forms in the second language writing class. *Language Learning, 68*(3), 813–850. https://doi.org/10.1111/lang.12295

Bozkurt, A., Karadeniz, K., Baneres, D., Guerrero-Roldán, A. E., & Rodríguez, M. E. (2021). Artificial intelligence and reflections from educational landscape: A review of AI studies in half a century. *Sustainability, 13*(2), Article 800. https://doi.org/10.3390/su13020800

Bozkurt, A., Xiao, J., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., Farrow, R., Bond, M., Nerantzi, C., Honeychurch, S., Bali, M., Dron, J., Mir, K., Stewart, B., Costello, E., Mason, J., Stracke, C. M., Romero-Hall, E., Koutropoulos, A., … Jandrić, P. (2023). Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education, 18*(1), 53–130. https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/709

Cao, L. (2023). Trans-AI/DS: Transformative, transdisciplinary and translational artificial intelligence and data science. *International Journal of Data Science and Analytics,15*(1), 119–132. https://doi.org/10.1007/s41060-023-00383-y

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*(2020), Article 100002. https://doi.org/10.1016/j.caeai.2020.100002

Chu, H.-C., Hwang, G.-H., Tu, Y.-F., & Yang, K.-H. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology, 38*(3), 22–42. https://ajet.org.au/index.php/AJET/article/view/7526

Cox, A. M. (2021). Exploring the impact of artificial intelligence and robots on higher education through literature-based design fictions. *International Journal of Educational Technology in Higher Education, 18*(2021), Article 3. https://doi.org/10.1186/s41239-020-00237-8

DeKeyser, R. (2007). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 94–112). Routledge.

Diebold, G. (2023, January 17). *Higher education will have to adapt to generative AI—And that's a good thing*. Center for Data Innovation. https://datainnovation.org/2023/01/higher-education-will-have-to-adapt-to-generative-ai-and-thats-a-good-thing/

Dogan, M. E., Goru Dogan, T., & Bozkurt, A. (2023). The use of artificial intelligence (AI) in online learning and distance education processes: A systematic review of empirical studies. *Applied Sciences, 13*(5), Article 3056. https://doi.org/10.3390/app13053056

Ellis, R. (2009a). A typology of written corrective feedback types. *ELT Journal, 63*(2), 97–107. https://doi.org/10.1093/elt/ccn023

Ellis, R. (2009b). Corrective feedback and teacher development. *L2 Journal: An electronic refereed journal for foreign and second language educators, 1*(1), 2–18. https://doi.org/10.5070/l2.v1i1.9054

Ellis, R. (2009c). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics, 19*(3), 221–246. https://doi.org/10.1111/j.1473-4192.2009.00231.x

Evans, N. W., Hartshorn, K. J., Cox, T. L., & Martin de Jel, T. (2014). Measuring written linguistic accuracy with weighted clause ratios: A question of validity. *Journal of Second Language Writing, 24*(1), 33–50. https://doi.org/10.1016/j.jslw.2014.02.005

Goksel, N., & Bozkurt, A. (2019). Artificial intelligence in education: Current insights and future perspectives. In S. Sisman-Ugur & G. Kurubacak (Eds.), *Handbook of research on learning in the age of transhumanism* (pp. 224–236). IGI Global. https://doi.org/10.4018/978-1-5225-8431-5.ch014

Han, Y., & Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal of Second Language Writing, 30*(1), 31–44. https://www.doi.org/10.1016/j.jslw.2015.08.002

Hartshorn, K. J., & Evans, N. W. (2015). The effects of dynamic written corrective feedback: A 30-week study. *Journal of Response to Writing, 1*(2), 6–34.

Humble, N., & Mozelius, P. (2022). The threat, hype, and promise of artificial intelligence in education. *Discover Artificial Intelligence, 2*(2022), Article 22. https://doi.org/10.1007/s44163-022-00039-z

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching, 39*(2), 83–101. https://doi.org/10.1017/S0261444806003399

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal, 99*(1), 1–18. https://doi.org/10.1111/modl.12189

Karim, K., & Nassaji, H. (2018). The revision and transfer effects of direct and indirect comprehensive corrective feedback on ESL students' writing. *Language Teaching Research, 24*(4), 519–539. https://doi.org/10.1177/1362168818802469

Karim, K., & Nassaji, H. (2019). The effects of written corrective feedback: A critical synthesis of past and present research. *Instructed Second Language Acquisition, 3*(1), 28–52. https://doi.org/10.1558/isla.37949

Kurzer, K. (2018). Dynamic written corrective feedback in developmental multilingual writing classes. *TESOL Quarterly, 52*(1), 5–33. https://doi.org/10.1002/tesq.366

Kurzweil, R. (2014). The singularity is near. In R. L. Sandler (Ed.), *Ethics and emerging technologies* (pp. 393–406). Palgrave Macmillan UK. https://doi.org/10.1057/9781137349088_26

Leeman, J. (2010). Feedback in L2 learning: Responding to errors during practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 111–138). Cambridge University Press. https://doi.org/10.1017/cbo9780511667275.007

Li, S., & Vuono, A. (2019). Twenty-five years of research on oral and written corrective feedback in *System*. *System, 84*(1), 93–109. https://doi.org/10.1016/j.system.2019.05.006

Lim, S. C., & Renandya, W. A. (2020). Efficacy of written corrective feedback in writing instruction: A meta-analysis. *TESL-EJ, 24*(3), 1–26. https://tesl-ej.org/wordpress/issues/volume24/ej95/ej95a3/

Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing, 30*(1), 66–81. https://doi.org/10.1016/j.jslw.2015.08.011

Liu, Y., & Huang, J. (2020). The quality assurance of a national English writing assessment: Policy implications for quality improvement. *Studies in Educational Evaluation, 67*(2), Article 100941. https://doi.org/10.1016/j.stueduc.2020.100941

Long, M. H. (1980). *Input, interaction, and second language acquisition*. University of California.

Luo, Y., & Liu, Y. (2017). Comparison between peer feedback and automated feedback in college English writing: A case study. *Open Journal of Modern Linguistics, 7*(04), 197–215. https://doi.org/10.4236/ojml.2017.74015

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J. H., Ogata, H., Baltes, J., Guerra, R., Li, P., & Tsai, C.-C. (2020). Challenges and future directions of big data and artificial intelligence in education. *Frontiers in Psychology*, *11*, Article 580820. https://doi.org/10.3389/fpsyg.2020.580820

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*(2021), Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies, 27*(6), 7893–7925. https://doi.org/10.1007/s10639-022-10925-9

Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing, 21*(4), 375–389. https://doi.org/10.1016/j.jslw.2012.09.004

Polio, C., & Yoon, H. J. (2018). The reliability and validity of automated tools for examining variation in syntactic complexity across genres. *International Journal of Applied Linguistics (United Kingdom), 28*(1), 165–188. https://doi.org/10.1111/ijal.12200

Riazantseva, A. (2012). Outcome measure of L2 writing as a mediator of the effects of corrective feedback on students' ability to write accurately. *System, 40*(3), 421–430. https://doi.org/10.1016/j.system.2012.07.005

Robinson, P. (2011). Second language task complexity, the cognition hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 3–38). John Benjamins.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–164). John Benjamins.

Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. *Perspectives on Individual Characteristics and Foreign Language Education*, *6*(27), 27-49.

Selwyn, N., Hillman, T., Bergviken-Rensfeldt, A., & Perrotta, C. (2023). Making sense of the digital automation of education. *Postdigital Science and Education, 5*(1), 1–14. https://doi.org/10.1007/s42438-022-00362-9

Sharma, R. C., Kawachi, P., & Bozkurt, A. (2019). The landscape of artificial intelligence in open, online and distance education: Promises and concerns. *Asian Journal of Distance Education, 14*(2), 1–2. https://doi.org/10.5281/zenodo.3730631

Sia, P. F. D., & Cheung, Y. L. (2017). Written corrective feedback in writing instruction: A qualitative synthesis of recent research. *Issues in Language Studies, 6*(1), 61–80. https://www.ils.unimas.my/images/pdf/v6n1/ILS_Vol6No1_Sia.pdf

Stefanou, C., & Révész, A. (2015). Direct written corrective feedback, learner differences, and the acquisition of second language article use for generic and specific plural reference. *The Modern Language Journal, 99*(2), 263–282. https://doi.org/10.1111/modl.12212

Storch, N. (2009). The impact of studying in a second language (L2) medium university on the development of L2 writing. *Journal of Second Language Writing*, *18*(2), 103–118. https://api.semanticscholar.org/CorpusID:62520708

Tang, K.-Y., Chang, C.-Y., & Hwang, G.-J. (2021). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, *31*(4), 2134–2152. https://doi.org/10.1080/10494820.2021.1875001 Thi, K, N., & Nikolov, M. (2021). Feedback treatments, writing tasks, and accuracy measures: A critical review of research on written corrective feedback. TESL-E, 25(3), 1-25. https://www.tesl-ej.org/pdf/ej99/a16.pdf

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*(1), Article 15. https://doi.org/10.1186/s40561-023-00237-x

Truscott, J. (2010). Further thoughts on Anthony Bruton's critique of the correction debate. *System, 38*(4), 626–633. https://doi.org/10.1016/j.system.2010.10.003

Truscott, J., & Hsu, A. Y.-P. (2008). Error correction, revision, and learning. *Journal of Second Language Writing, 17*(4), 292–305. https://doi.org/10.1016/j.jslw.2008.05.003

Winterson, J. (2022). *12 Bytes: How artificial intelligence will change the way we live and love*. Penguin Random House.

Wolfe-Quintero, K. (1998). The connection between verbs and argument structures: Native speaker production of the double object dative. *Applied Psycholinguistics*, *19*(2), 225–257. https://doi.org/10.1017/S0142716400010055

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zhang, T. (2021). The effect of highly focused versus mid-focused written corrective feedback on EFL learners' explicit and implicit knowledge development. *System, 99*(2), Article 102493. https://doi.org/10.1016/j.system.2021.102493

International Review of Research in Open and Distributed Learning
Volume 25, Number 3

August – 2024

The Effects of Educational Artificial Intelligence-Powered Applications on Teachers' Perceived Autonomy, Professional Development for Online Teaching, and Digital Burnout

Hong Duan [1,2] and Wei Zhao[1,*]

*[1]School of Information Science and Technology, Northeast Normal University, Changchun, Jilin,130117, China and Normal College, Shihezi University, Shihezi, Xinjiang, 832003, China; [2]Normal College, Shihezi University, Shihezi, Xinjiang, 832003, China; * Corresponding author*

Abstract

The transformative impact of advancements in educational technology, particularly those powered by artificial intelligence (AI), on the landscape of education and the teaching profession has been substantial. This study explores the repercussions of AI-powered technologies on teachers' autonomous behavior, digital burnout, and professional development. The study involved a cohort of 320 high school teachers in China segregated into control and experimental groups. The experimental group received instructions on AI-integrated applications and how they might be used in education. However, the teachers assigned to the control group did not receive information on the use of AI educational applications. Three distinct questionnaires probing autonomous behaviors, digital burnout, and online professional development were administered, and the ensuing data were analyzed using independent sample *t*-tests. The findings elucidate a discernible positive impact of AI-integrated technology intervention on teachers' professional development and autonomous behaviors. The incorporation of AI-enhanced tools facilitated an augmentation in teachers' professional growth and bolstered their independent and self-directed instructional practices. Notably, using AI-integrated technology significantly reduced teachers' susceptibility to digital burnout, signifying a potential alleviation of stressors associated with technology-mediated teaching. This research provides valuable insights into the multifaceted effects of AI-powered technologies on educators, shedding light on enhancing professional competencies and mitigating digital burnout. The implications extend beyond the confines of this study, resonating with the broader discourse on leveraging technology to augment the teaching profession and optimize the learning environment.

Keywords: artificial intelligence, AI, AI applications, teacher autonomy, online professional development, digital burnout

Introduction

The advent of the information age has witnessed rapid technological developments, with digitalization playing a pivotal role across various aspects of our lives. The onset of the COVID-19 pandemic further accelerated technology adoption globally, necessitating preventive measures such as enclosure practices. Consequently, technology use has become ubiquitous, shaping a digital lifestyle that significantly influences professional efficiency and sociopsychological well-being (Erten & Özdemir, 2020). The pandemic prompted an unprecedented surge in technology usage in online shopping, the professional life, communication, social media, and news consumption. As a result, in some places, there were individuals who spent some of their waking hours online during quarantine (Sharma et al., 2020). However, this extensive reliance on technology, particularly in professional and social spheres, has brought about adverse effects, including stress, fatigue, reduced performance, and burnout. Burnout, defined as a professional deformation affecting health (Şengün, 2021), encompasses emotional and desensitization syndromes associated with continuous interaction with people (Maslach et al., 2001). Occupations requiring constant interaction, such as teaching, are especially prone to burnout, with factors such as student discipline issues, overcrowded classrooms, communication challenges, professional dissatisfaction, unfair administrative practices, and low income contributing to the stress and burnout experienced by teachers (Serter, 2021).

One of the areas that might be affected by technological innovation is teachers' online professional development. Professional development (PD) as defined by Richards and Farrell (2005) is fundamentally perceived as a shared trajectory of growth for educators, constituting a process geared towards long-term objectives that improve teachers' understanding of the art of teaching and their roles as educators. Comprising a myriad of activities, PD is strategically positioned to enhance the overall quality of teaching. Its diverse manifestations range from informal, individual endeavors, such as reading professional literature, to formal and large-scale initiatives offered by institutions and entities such as ministries of education (Borg, 2018).

Furthermore, the evolution of the information era has introduced a distinctive form of burnout known as digital burnout, which has become increasingly prevalent. The COVID-19 pandemic has compelled teachers to rely heavily on technology, especially in the transition to distance education. The immersion in digital tools, extending beyond traditional working hours, has given rise to digital burnout, marked by stress, fatigue, desensitization, reduced attention, and physical and mental health issues (Erten & Özdemir, 2020). Amidst these technological shifts, the role of AI in education has gained prominence.

AI in education is a transformative concept that describes leveraging advanced technologies to enhance learning processes, adapt educational content to individual needs, and revolutionize the way students and educators engage with information. By employing machine-learning algorithms, AI systems can analyze vast amounts of data to identify patterns, personalize learning experiences, and offer real-time feedback to students (Bozkurt et al., 2021). This technology also facilitates the development of intelligent tutoring systems that cater to diverse learning styles, providing targeted support for students facing challenges. Additionally, AI-powered educational tools contribute to the creation of interactive and immersive learning environments, fostering critical thinking, problem-solving skills, and preparing students for the demands of the rapidly evolving digital era (Bozkurt et al., 2023). While AI in education holds great promise, ethical considerations, data privacy, and the need for responsible implementation are crucial aspects to ensure a positive impact and equitable access to enhanced educational experiences for all learners. AI-powered applications in education have become instrumental in reshaping teaching methodologies, challenging traditional paradigms, and offering new opportunities for both educators and learners. However, as technology infiltrates educational landscapes, exploring its impact on teachers' autonomy, PD, and the emergence of digital burnout has become imperative. This study delves into the intricate relationship between educational AI-powered applications and teachers' experiences, investigating the mediating role of teachers' technical, pedagogical, and content knowledge. While the importance of AI in education is widely acknowledged (Bozkurt et al., 2021, 2023; Cao, 2023; Tang et al., 2021; Tlili et al., 2023), there exists a notable gap in understanding how this technology has influenced teachers' perceptions of autonomy, PD, and digital burnout, highlighting the need for comprehensive exploration and analysis in this evolving educational landscape.

Literature Review

Contribution of AI to Education

AI is a transformative tool that encompasses machines which are able to do tasks traditionally carried out by humans (Salas-Pilco et al., 2022;). Its usage is experiencing unprecedented growth, fundamentally altering various features and dimensions of human life (Cao, 2023). In education, recent years have witnessed the effective integration of AI and learning analytics (Salas-Pilco et al., 2022). Education, spanning both school and higher education, is a multifaceted domain that includes teacher education and is recognized as integral to shaping the future. Incorporating AI in education, particularly teacher education, reshape various fields (Salas-Pilco et al., 2022).

Early use of the term AI marked the inception of a field that would eventually compete with human intellect (Zawacki-Richter et al., 2019). AI, often referred to as machine intelligence, mimics human intelligence and has become ingrained in daily life activities such as online shopping, Internet browsing, and GPS-based navigation (Benavides et al., 2020; Bozkurt et al., 2023). In education, AI holds significant potential for task automation, teacher support, and addressing classroom weaknesses.

The introduction of AI in education, particularly in teacher education, can enhance the quality of education and transform traditional teaching methods (Jamal, 2023). AI can provide teachers access to various tools and resources for PD. AI-powered assessment tools offer real-time feedback on student performance, enabling teachers to adjust their strategies accordingly. The comparison between search results on Google and ChatGPT for "ways to improve teaching skills" illustrates AI's potential to enhance teachers' skills (Goksel & Bozkurt, 2019; Jamal, 2023; Nataraj, 2022).

AI facilitates personalized learning by tailoring resources and experiences to individual students. It can identify learning styles, interests, and abilities, allowing teachers to create customized lessons. An AI-driven personalized learning framework scrutinizes real-time student performance data, presenting customized educational materials, adapting learning parameters, and furnishing feedback (Qadir, 2022). Likewise, Jamal (2023) contended that a noteworthy hurdle in teacher education lies in establishing a robust subject matter foundation. AI confronts this challenge by furnishing educators with superior educational resources, encompassing online lectures, educational videos, and electronic books. Through the analysis of teacher performance data, AI discerns knowledge deficiencies, facilitating the formulation of precise PD initiatives tailored to address particular requirements (Qadir, 2022). Moreover, AI assists teachers in identifying students' learning styles by analyzing data on their interactions with online learning systems. This information informs instructional strategies tailored to individual learning styles (Mhlanga, 2023; Qadir, 2022; Thurzo et al., 2023).

More significantly, AI offers adaptive learning experiences, adjusting the difficulty and complexity of content to match individual learning paces and abilities. Personalized learning through AI helps students master subjects and improve learning outcomes (Qadir, 2022). Finally, it supports continuous PD by providing feedback on teacher performance and recommending tailored PD opportunities based on specific needs (Chu et al., 2022).

Teachers' Professional Development

In recent decades, there has been a growing emphasis on the significance of exceptional teaching and targeted PD designed to enhance student learning outcomes, particularly in the context of Teaching English as a Foreign Language (TEFL; Powell & Bodur, 2019). Improving teacher quality is a fundamental requirement for augmenting the overall quality of education (Borg, 2018). Nevertheless, challenges regarding PD have emerged within the TEFL domain, stemming from a lack of understanding regarding PD planning, unawareness of diverse PD types and their quality, and a disregard for the perspectives of teachers on PD activities (Cirocki & Farrell, 2019).

It is well established that the COVID-19 pandemic has exerted substantial impacts on the field of education. Consequently, educators must now equip themselves to address the suspension of in-person classes due to unforeseen emergencies (Moorhouse, 2020). Fortunately, technological advancements have facilitated online instruction for teaching second or foreign languages, obviating the need for physical presence (Shin & Kang, 2018). Therefore, teachers are compelled to migrate towards online platforms to sustain student engagement in the learning process. This transition significantly amplifies teachers' workloads, involving not only the relocation of instructional materials to an online learning environment but also the incorporation of requisite applications (Allen et al., 2020).

The foundational principle of teacher training is to advance educators' personal and professional development over their teaching careers, emphasizing continuous learning from self, peers, and experiences (Sandoval-Cruz et al., 2022). Conceptualizing training as a "system" necessitates a committed political stance by all relevant authorities, spanning candidate selection, support for newly qualified teachers, and ongoing resource provision (Loughran, 2019). Continuous training is indispensable for in-service and novice teacher development, viewed as an individual and collective journey across diverse contexts (Marcelo et al., 2023).

The efficacy of the transition to online teaching is fundamentally contingent upon teachers' support and active participation (Adnan & Kainat, 2020). The shift from traditional face-to-face instruction to online learning necessitates educators to assume new roles and acquire additional competencies. Teachers must possess the knowledge and skills to engage in online teaching methodologies adeptly. PD is critical in facilitating online educators' mastery of innovative pedagogies, adaptation to novel roles, acquisition of essential competencies, and reconstructing their professional identity within the online learning milieu (Adnan & Kainat, 2020).

In light of the unexpected challenges introduced by the COVID-19 pandemic, the realm of language learning has been compelled to fully embrace online modalities. A myriad of challenges has arisen for students, teachers, and parents within the domain of online language learning. Teachers, in particular, find themselves compelled to augment their qualifications and equip themselves with the requisite knowledge and skills to enhance the quality of instructional delivery within the online learning milieu. This has generated a demand for research that delves into PD concerning the integration of technology in language learning, as articulated by Atmojo and Nugroho (2020). Despite the wealth of research and best practices in the domain of online language learning, references that explicate the preparation of language educators for online instructional environments and delineate the essential competencies for effective teaching within this context are conspicuously limited (Compton, 2009). This scarcity is notable despite the recent proliferation of online learning initiatives and the burgeoning interest in online teacher professional development (OTPD). There exists a discernible gap in understanding teachers' perceptions of online teaching, prompting the need for further exploration and comprehension in this regard.

Based on PD, teacher training prioritizes enhancing teachers' capacity to navigate the challenges of teaching practice and assumes the role of teachers as researchers and knowledge producers (Monereo & Caride, 2022). This perspective mandates research training as part of their initiation and development (Wylleman et al., 2009). The process hinges on political, economic, institutional, and personal efforts to secure resources, recognizing that reflexive practitioners contribute significantly to transforming educational environments (Alibakhshi & Dehvari, 2015).

Teachers should be active protagonists in their training, not mere recipients, as their commitment to learning is integral to PD and educational improvement (Marcelo et al., 2023). This bold reconstruction aligns with the teacher-subject concept, emphasizing permanent PD, necessitating stimuli and incentives for teachers to adapt, change, and innovate throughout their careers (Marcelo et al., 2023).

Technological advancements have facilitated the emergence of novel and varied forms of PD (Parsons et al., 2019). Technological progress has allowed language educators worldwide to pursue PD credentials and academic degrees through online platforms (Shin & Kang, 2018). Online teacher professional development (TPD) is a promising avenue for augmenting teachers' knowledge, skills, and competencies, offering flexible, cost-effective, and extensive alternatives across a broad spectrum of topics. Online TPD encompasses various formats, including courses, seminars, workshops, discussions, and resources, delivered synchronously, asynchronously, or in a blended fashion through platforms such as websites, blogs, wikis, podcasts, and social media. However, more than access to online TPD is needed to ensure optimal outcomes; effective use of technology requires attention to design and implementation principles rather than merely considering it as a delivery medium (Powell & Bodur, 2019).

Teachers' Digital Burnout

Technology usage has surged during the pandemic in various areas, such as work, communication, online shopping, social media, and news consumption. Sharma et al. (2020) noted that more individuals spent "nearly all" of their waking hours online during the COVID-19 quarantine (para. 2). However, this heavy reliance on technology, both in professional and social spheres, has resulted in negative consequences, including stress, fatigue, decreased performance, and burnout. Burnout, defined by the World Health Organization in 2019, is a professional deformation affecting individuals' health (Şengün, 2021).

Burnout is characterized by a loss of power and a lack of effort (Maslach & Leiter, 2016). Maslach and Jackson (1981) described it as emotional burnout and desensitization syndrome, often arising from working with people. Pines and Aronson (1988) viewed burnout as a physical, emotional, and mental breakdown, leading to a loss of capacity, energy, idealism, and purpose, accompanied by feelings of pessimism, despair, and entrapment. Studies have suggested that burnout is common among professions involving continuous interaction with people, such as teaching (Maslach et al., 2001; Oplatka, 2002). Teachers, in particular, may experience burnout due to factors such as student discipline problems, overcrowded classrooms, communication challenges with parents, professional dissatisfaction, unfair administrators, and low income (Serter, 2021).

Burnout can manifest in various ways, encompassing physical and psychological aspects (Deliorman Bakoğlu et al., 2009). In the current information era, a new form of burnout, digital burnout, has become more prevalent.

The COVID-19 pandemic has compelled teachers to rely heavily on technology, especially with the shift to distance education. Teachers have immersed themselves in digital tools, extending their traditional working hours to adapt to the new distance education system. This constant exposure to digital tools 24/7 has given rise to digital burnout, characterized by stress, fatigue, desensitization, decreased attention, and physical and mental health issues (Erten & Özdemir, 2020).

Teachers' Perceived Autonomy

Enhancing the quality of education is the fundamental mission of higher education and an essential prerequisite for constructing a robust educational foundation for a nation (Ruiz-Alfonso et al., 2021). Higher education is the primary force in nurturing talents essential for societal advancement. In the current era of rapid information technology development, possessing deep learning capabilities signifies the capacity for innovation, creativity, and sustainable development—a critical skill demanded in the context of contemporary societal and era advancements (Esteban-Guitart & Gee, 2020).

Moreover, higher education teaching primarily focuses on undergraduates, whose principal responsibility is to learn how to learn. This goes beyond mere surface-level comprehension and rote memorization of knowledge. Instead, the emphasis is on cultivating a profound understanding of knowledge, critically engaging with new information, mastering concepts through practical activities, honing critical thinking skills, and fostering learning and innovation abilities (Zhang et al., 2022). Consequently, students' learning styles play a pivotal role in evaluating teaching quality, prompting increasing scholarly attention to transition students from shallow to deep learning (Sølvik & Glenna, 2021). Teachers' autonomy support is widely recognized as a crucial external factor influencing college students' deep learning (Kaplan, 2018; Zhao & Qin, 2021). Autonomy support from teachers entails emotional validation, support, and encouragement for students' autonomous decisions and free choices (Ryan et al., 2016). It has been observed that teachers' autonomy support not only fosters positive teacher-student relationships but also encourages deep learning styles contributing to the enhancement of students' self-efficacy.

In sum, critical analysis shows that the number of studies on the use of AI in fostering the variables related to teacher development is scant. Therefore, it is necessary to see whether and how innovation in AI and related educational applications might affect teachers' cognitive and affective variables such as perceived autonomy, PD for online teaching, and digital burnout.

Research Questions

This study investigated the impacts of AI-powered educational applications on high school teachers' autonomy, digital burnout, and online PD. More specifically, the following research questions were addressed:

1. Does high school teachers' use of AI-powered educational applications foster their perceived autonomy?

2. Does high school teachers' use of AI-powered educational applications affect their professional development for online teaching?

3. Does high school teachers' use of AI-powered educational applications affect their digital burnout?

Methodology

Sample and Procedure

In this study, 350 high school teachers from a city in China were initially identified and invited to participate in an online survey. Of this pool, 330 accepted the invitation and actively engaged in the study by completing questionnaires. From this cohort, 330 high school teachers (160 females and 170 males) were randomly selected to ensure a representative sample for further investigation. From among these, 120 teachers were then chosen to participate in a specialized online workshop focusing on the application of AI in education. The workshop, spanning two three-hour sessions, provided insights into various AI applications such as ChatGPT, Poe, Duolingo, and others. The selected teachers were supported by their respective schools and actively encouraged to integrate AI applications into their teaching practices.

Following the completion of the workshop, teachers, now divided into two distinct groups based on their attendance, were monitored for four months. One group consisted of those actively participating in the AI workshop (control group), while the other group comprised teachers who did not participate (experimental group). After 4 months, identical questionnaires were administered to all 330 teachers, capturing their responses on dimensions related to PD, digital burnout, and autonomous behaviors. The collected data were then numerically coded, facilitating a comparative analysis between the two groups using descriptive statistics and independent sample *t*-tests. This robust sampling and procedural approach aimed to explore the impact of AI workshops on teachers' PD, digital burnout, and autonomous behaviors, shedding light on the potential influence of AI applications in the educational context.

Data Collection

Three instruments were employed in this study to assess different aspects of high school teachers' experiences. The first instrument, the Scale for Teacher Autonomous Behavior, was developed by Evers et al., (2017). This scale comprises four key dimensions: primary work processes in the class (6 items), curriculum implementation (6 items), participation in decision-making at school (4 items), and Professional development. It is a comprehensive tool to measure various facets of teacher autonomy within the educational context. The second instrument, i.e., professional development, was adapted from Atmojo (2021). The questionnaire consists of 8 multiple choice items and 2 open-ended questions, but only the multiple-choice items were used. The third instrument used in this study was the Digital Burnout Scale, developed by Erten and Özdemir (2020). This scale is designed to gauge individuals' levels of digital burnout and includes three sub-dimensions: digital aging, digital deprivation, and emotional exhaustion. The researchers of the current study established the validity of this scale in terms of item content and construct validity. Additionally, the Cronbach Alpha coefficient for the scales was above 0.85 for each scale and its components, indicating acceptable levels.

Results

Research Question 1

Question 1 addressed the effects of teachers' use of AI-powered applications on teachers' autonomy. The results of *t*-tests for the two groups are presented in Table 1.

Table 1

Groups' Scores on Autonomous Behaviors (Pretest and Posttest)

Variable	Treatment	Pretest					Posttest			
		M	*SD*	*t*	*df*	*p*	*M*	*t*	*df*	*p*
PWP in the class	E	3.2	0.85	1.23	318	> .05	3.8	14.12	318	< 0.001
	C	3.3	0.79				4.35			
Curriculum implementation	E	3.3	0.83	0.98	318	> .05	3.70	15.12	318	< 0.001
	C	3.26	0.91				4.32			
Participation in decision-making at school	E	2.90	0.79	0.96	318	> .05	3.65	16.11	318	< 0.001
	C	2.83	0.82				4.20			
Autonomy total	E	3.30	0.56	0.87	318	> .05	3.75	17.11	238	< 0.001
	C	3.12	0.56				4.32			

Note. PWP = primary work processes; E = experimental; C = control.

As presented in Table 1, the pretest and posttest scores illuminate different dimensions of teacher autonomy for both the experimental and control groups. In terms of primary work processes (PWP) in the class, the experimental group exhibited a noteworthy improvement, demonstrating a statistically significant increase from 3.2 to 3.8 (t (318) = 1.23, p < 0.001). Conversely, the control group's scores remained relatively stable. Regarding curriculum implementation, the experimental group demonstrated a significant enhancement from 3.3 to 3.70 (t (318) = 0.98, p < 0.001), indicating positive changes in this dimension. The control group also increased, but the change was less pronounced. In the dimension of participation in decision-making at school, the experimental group showed a significant improvement from 2.90 to 3.65 (t (318) = 0.96, p < 0.001). While the control group also experienced an increase, it was less substantial. Examining the total autonomy scores, the experimental group demonstrated a statistically significant improvement, increasing from 3.30 to 3.75 (t (318) = 0.87, p < 0.001). Meanwhile, the control group's scores increased from 3.12 to 4.32.

Interpreting these findings suggests that the intervention positively impacted teacher autonomy across various dimensions, encompassing primary work processes, curriculum implementation, participation in decision-making at school, and the overall autonomy construct. The experimental group exhibited more substantial improvements compared to the control group, emphasizing the potential efficacy of the intervention in fostering positive developments in teacher autonomy.

Research Question 2

The second research question addressed the effects of the treatment on teachers' PD for online teaching. Results are presented in Table 2.

Table 2

Groups' Scores on Professional Development (Pretest and Posttest)

Variable	Treatment	Pretest					Posttest			
		M	*SD*	*t*	*df*	*p*	*M*	*t*	*df*	*p*
PD for online classes	Experimental	3.3	0.91	1.31	318	> .05	3.8	14.12	318	< 0.001
	Control	3.4	0.77				4.46			

Note. PD = professional development.

As seen in Table 2, the experimental group exhibited a statistically significant improvement in the PD for online classes dimension, with scores increasing from 3.4 to 4.46 (t (318) = 1.31, p < 0.001). In contrast, the control group displayed a minor increase, moving from 3.3 to 3.8. Comparing the groups, it is evident that the intervention had a notable impact on the experimental group, resulting in a more substantial improvement in the PD for online classes than the control group. This suggests that the treatment effectively enhanced participants' perceptions and skills related to PD tailored explicitly for online teaching modalities.

Question Three

Question 3 addressed the effects of teachers' use of AI-powered applications on reducing teachers' digital burnout. The results of the *t*-tests for the two groups are presented in Table 3.

Table 3

Groups' Scores on Digital Burnout (Pretest and Posttest)

Variable	Treatment	Pretest					Posttest			
		M	*SD*	*t*	*df*	*p*	*M*	*t*	*df*	*p*
Digital aging	E	3.3	0.85	1.23	318	> .05	1.65	14.30	318	< 0.001
	C	3.2	0.79				3.10			
Digital deprivation	E	3.26	0.83	0.98	318	> .05	1.69	15.25	318	< 0.001
	C	3.3	0.91				3.00			
Emotional exhaustion	E	2.8	0.79	0.96	318	> .05	1.50	16.17	318	< 0.001
	C	2.90	0.82				2.75			
Total score	E	3.12	0.56	0.87	318	> .05	1.70	17.25	318	< 0.001
	C	3.3	0.56				3.05			

Note. E = experimental; C = control.

As seen in Table 3, regarding digital aging, the control group demonstrated a minor decrease from 3.2 to 3.10, while the experimental group showed a more substantial reduction, moving from 3.3 to 1.65. The difference between the groups' scores on the posttest was statistically significant (t (318) = 14.30, p < 0.001). Similarly, for digital deprivation, the control group exhibited a modest reduction from 3.3 to 3.00, while the experimental group displayed a more considerable reduction, moving from 3.26 to 1.69. The difference between the groups' scores on the posttest was statistically significant (t (318) = 14.25, p < 0.001). Analyzing emotional exhaustion, the control group experienced a slight decrease from 2.9 to 2.75 (t (318) = 0.96, p < 0.001), whereas the experimental group showed a more pronounced reduction, moving from 2.83 to 1.50.

The difference between the groups' scores on the posttest was also statistically significant (t (318) = 16.17, p < 0.001). For the total score, which encompasses all dimensions, the control group demonstrated a minor reduction from 3.3 to 3.05, while the experimental group exhibited a more significant improvement, progressing from 3.12 to 1.70. Results of the t-test also verified that the difference between the two groups' digital burnout at the end of the treatment was statistically significant (t (318) = 17.25, p < 0.001). The comparisons between the control and experimental groups indicate that the intervention had a notable impact on digital burnout, with the experimental group consistently showing a more substantial reduction across all dimensions. This suggests the effectiveness of the treatment in reducing participants' digital burnout.

Discussion

The first research question delved into the impact of AI-powered applications on teachers' autonomous behaviors, explicitly focusing on dimensions such as primary work processes (PWP) in the class, curriculum implementation, participation in decision-making at school, and the overall measure of autonomy. To investigate this, independent t-tests were employed to compare the groups' scores on both pretests and posttests. The analysis revealed no statistically significant differences between the groups in the pretest scores. However, in the posttest, significant differences emerged, indicating that teachers who used AI-powered applications experienced greater autonomy than those who did not. These findings align with prior research supporting the positive impact of PD and innovative methods, such as AI integration, on teacher autonomy (Alibakhshi & Dehvari, 2015; Jamal, 2023; Kaplan, 2018). The study contributes to the existing literature by demonstrating the effectiveness of AI applications in fostering teacher autonomy, as reflected in the posttest results.

The study suggests that AI-powered applications play a crucial role in shaping and enhancing teachers' professional behaviors, particularly regarding autonomy. Moreover, the findings resonate with the work of Borg (2018) and Cirocki and Farrell (2019), emphasizing the importance of evaluating the impact of PD. In this context, integrating AI applications is a valuable component of ongoing teacher development efforts.

Considering the extensive literature on burnout and digital well-being (Erten & Özdemir, 2020; Maslach et al., 2001; Sharma et al., 2020), the positive impact of AI-powered applications on autonomy may also contribute to mitigating burnout risks associated with technological advancements.

The findings align with broader discussions on the role of AI in education (Mhlanga, 2023; Qadir, 2022; Tang et al., 2021), emphasizing the need for responsible and ethical use to promote lifelong learning. The references cited provide a comprehensive foundation for understanding the context of AI in education, its potential benefits, and the importance of considering various factors such as burnout and PD. In conclusion, the study demonstrates that AI-powered applications positively influence teachers' autonomous behaviors, as evidenced by significant posttest differences. The discussion integrates insights from diverse references, providing a well-rounded understanding of the implications of AI integration in the context of teacher autonomy and PD.

The second research question aimed to investigate the impact of AI-powered applications on teachers' PD for online classes. The analysis revealed a statistically significant improvement in the experimental group, with scores increasing from 3.4 to 4.46 (t (318) = 1.31, p < 0.001). In contrast, the control group displayed a more minor increase, moving from 3.3 to 3.8. These results indicate that the intervention involving the use of AI-powered applications had a notable impact on the experimental group's perceptions and skills related to PD for online teaching. The statistically significant improvement suggests that the treatment effectively enhanced participants' abilities and understanding in adapting PD practices to the online teaching modality.

Comparing the two groups, it becomes evident that the experimental group experienced a more substantial improvement in PD for online classes compared to the control group. This discrepancy underscores the effectiveness of incorporating AI-powered applications into PD initiatives, particularly those tailored for online teaching. These findings align with previous research emphasizing the importance of targeted PD for online teaching (Compton, 2009; Parsons et al., 2019; Powell & Bodur, 2019). The study contributes to this body of knowledge by highlighting the specific positive impact of AI-powered applications on enhancing teachers' PD in the online context.

The results also resonate with the broader discourse on the role of technology in education, emphasizing the potential of AI to support and enrich PD experiences (Shin & Kang, 2018; Zawacki-Richter et al., 2019). The intervention's success in fostering improvements in PD for online classes speaks to the adaptability and relevance of AI tools in the rapidly evolving landscape of online education. Furthermore, the study's implications extend to the ongoing discussions about the future of education and the transformative potential of AI (Bozkurt et al., 2021; Tang et al., 2021; Terra, et al., 2023; Thurzo et al., 2023). The positive outcomes in the PD for online classes dimension support the idea that AI can be a valuable ally in preparing teachers for the challenges and opportunities presented by digital and online learning environments.

findings are also in line with Cirocki and Farrell (2019), who emphasized the significance of professional development for teachers, particularly in online education. The improvement observed in the experimental group's scores for PD for online classes suggests that tailored professional development, enhanced by AI applications, can positively impact teachers' skills and perceptions in online teaching modalities. Compton's (2009) work on preparing language teachers for online instruction is also relevant to the study's context.

Research question 3 was developed to explore the impact of teachers' use of AI-powered applications on reducing digital burnout, as indicated by dimensions such as digital aging, digital deprivation, emotional exhaustion, and the total score. The study's findings align with previous research highlighting the potential of technology, including AI, in mitigating burnout and stress in various professional contexts (Sharma et al., 2020). Specifically, the study contributes to this literature by demonstrating the positive effects of AI-powered applications on reducing digital burnout among teachers. The findings indicate that AI-powered applications, by aiding teachers in managing digital burnout, contribute to developing necessary skills for effective online teaching, aligning with Compton's focus on skills, roles, and responsibilities in online language teaching.

Additionally, the study resonates with Jamal (2023), who explored the role of AI in teacher education. The positive impact observed in reducing digital burnout aligns with Jamal's discussion on the opportunities and challenges associated with AI in teacher education. Sharma et al. (2020) and Erten and Özdemir (2020) have addressed the issue of digital burnout, providing a theoretical framework for understanding its dimensions. The current study contributes empirically to this understanding, demonstrating that using AI-powered applications can effectively reduce digital burnout among teachers.

Conclusions and Implications

The findings offer valuable insights into the implications of AI-powered applications on teachers' autonomous behaviors, PD for online classes, and the reduction of digital burnout. In addressing the first research question, it was evident that teachers who integrated AI-powered applications into their practices experienced greater autonomy, as reflected in dimensions such as primary work processes, curriculum implementation, participation in decision-making at school, and the overall measure of autonomy. The study highlights the effectiveness of AI applications in fostering teacher autonomy, offering a significant avenue for enhancing professional behaviors.

Furthermore, as revealed in our findings related to the second research question, the positive outcomes in the PD for online classes underscore the potential of AI-powered applications in tailoring practical teacher training for online teaching modalities. The success in enhancing teachers' abilities and understanding in adapting PD practices to the online context signifies the adaptability and relevance of AI tools in the evolving landscape of online education.

In addressing the third research question, the study demonstrates the positive impact of AI-powered applications on reducing digital burnout among teachers. This aligns with broader discussions on the potential of technology, including AI, in mitigating burnout and stress (Sharma et al., 2020). The study contributes empirically to understanding digital burnout dimensions, emphasizing the practical benefits of AI applications in supporting teachers' well-being in a technologically driven educational environment.

These findings hold significant implications for educational stakeholders, policymakers, and practitioners. The study suggests that the strategic integration of AI-powered applications in teacher training and PD initiatives can lead to positive outcomes, enhancing autonomy and the ability to navigate online teaching challenges. Additionally, recognizing the role of AI in mitigating digital burnout highlights the need for a balanced and responsible use of technology in education. The study advocates for ongoing exploration and integration of innovative technologies to support teachers in their professional growth and well-being as the educational landscape evolves.

Limitations and Suggestions for Further Studies

The study, which examines the impact of educational AI on teachers' perceived autonomy, PD for online teaching, and digital burnout, is of great value in the field of modern education. By examining the integration of AI technologies into educational environments, the research provides valuable insights into how these tools impact teachers' autonomy and shape their role in the learning process. Additionally, examining the impact of AI on PD in online teaching is critical to understanding how educators can leverage technology for continuous improvement. Additionally, the digital burnout research addresses a relevant issue in the digital age and highlights potential challenges and stressors that teachers may face when using AI-driven applications. The results of this study have the potential to influence education policy, guide the development of tailored AI tools to support educators, and improve the overall effectiveness of online education platforms.

While this study provides valuable insights into the impact of AI-powered applications on teachers' autonomous behaviors, PD for online classes, and the reduction of digital burnout, several limitations should be considered, paving the way for future research. First, the study's focus on a specific set of AI applications raises the need for broader investigations encompassing a diverse range of AI tools. Understanding how various AI interventions influence teacher practices and well-being could offer a more comprehensive view of the impact on education.

Second, the reliance on self-reported measures introduces potential biases. Future research could adopt a mixed-methods approach, combining qualitative methods such as observations and interviews with quantitative data to better understand teachers' experiences with AI applications. The sample, predominantly consisting of teachers from a specific geographic region and educational context, limits the generalizability of findings. Future studies should aim for a more diverse and representative sample, considering variations in educational levels, cultural backgrounds, and teaching environments. Furthermore, the study's short-term focus may not capture the long-term effects of AI integration. Longitudinal studies could provide insights into the sustained impact of AI-powered applications on teachers' practices, PD, and well-being over an extended period.

In line with the limitations of the study, comparative analyses to evaluate the effectiveness of different AI tools, qualitative investigations to deepen understanding, and contextualized research exploring factors influencing AI's impact are suggested. Additionally, cross-cultural studies, investigations into teacher training programs, and in-depth analyses of digital burnout dimensions could enhance our understanding of the multifaceted role of AI in education. By addressing these limitations and pursuing these research avenues, future studies can contribute to a more nuanced and comprehensive understanding of how AI shapes teacher autonomy, PD, and well-being in the evolving education landscape.

Funding

This work was supported by the Xinjiang Corps 2021 Online-Offline Blended First-Class Course Construction Project "Design and Development of Digital Learning Resources" (HHSYL-2021-21).

This work was supported by the Shihezi University 2022 Foundation of Research on the construction of evidence-based teaching competence model for teachers in the context of digital transformation of education (ZZZC2022107).

Acknowledgments

The authors acknowledge that they used AI-powered applications for editing and proofreading some sections of this paper when they were not sure of accuracy and fluency.

References

Adnan, M., & Kainat, A. (2020). Online learning amid the COVID-19 pandemic: Students' perspectives. *Journal of Pedagogical Sociology and Psychology* 2(1):45–51. https://doi.org/10.33902/JPSP.2020261309

Alibakhshi, G., & Dehvari, N. (2015). EFL teachers' perceptions of continuing professional development: A case of Iranian high school teachers. *PROFILE: Issues in Teachers' Professional Development*, *17*(2), 29–42. https://doi.org/10.15446/profile.v17n2.44374

Allen, J., Mahamed, F., & Williams, K. (2020). Disparities in education: E-Learning and COVID-19, who matters? *Child & Youth Services, 41*(3), 208–210. https://doi.org/10.1080/0145935x.2020.1834125

Atmojo, A. E. P. (2021). EFL teachers' online teacher professional development experiences amidst the COVID-19 pandemic: Practices and perceptions. *Englisia: Journal of Language, Education, and Humanities*, *9*(1), 1-18.

Atmojo, A. E. P., & Nugroho, A. (2020). EFL classes must go online! Teaching activities and challenges during COVID-19 pandemic in Indonesia. *Register Journal, 13*(1), 49–76. DOI:10.18326/rgt.v13i1.49-76

Benavides, L. M. C., Tamayo Arias, J. A., Arango Serna, M. D., Branch Bedoya, J. W., & Burgos, D. (2020). Digital transformation in higher education institutions: A systematic literature review. *Sensors*, *20*(11), Article 3291. https://doi.org/10.3390/s20113291

Borg, S. (2018). Evaluating the impact of professional development. *RELC Journal*, *49*(2), 195–216. https://doi.org/10.1177/0033688218784371

Bozkurt, A., Karadeniz, K., Baneres, D., Guerrero-Roldán, A. E., & Rodríguez, M. E. (2021). Artificial intelligence and reflections from the educational landscape: A review of AI studies in half a century. *Sustainability*, *13*(2), 1–16. https://doi.org/10.3390/su13020800

Bozkurt, A., Xiao, J., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., Farrow, R., Bond, M., Nerantzi, C., Honeychurch, S., Bali, M., Dron, J., Mir, K., Stewart, B., Costello, E., Mason, J., Stracke, C. M., Romero-Hall, E., Koutropoulos, A. … Jandrić, P. (2023). Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education*, *18*(1), 53–130. https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/709

Cao, L. (2023). Trans-AI/DS: Transformative, transdisciplinary, and translational artificial intelligence and data science. *International Journal of Data Science and Analytics*, *15*(1), 119–132. https://doi.org/10.1007/s41060-023-00383-y

Chu, H.-C., Hwang, G.-H., Tu, Y.-F., & Yang, K.-H. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology, 38*(3), 22–42. https://ajet.org.au/index.php/AJET/article/view/7526

Cirocki, A., & Farrell, T. S. C. (2019). Professional development of secondary school EFL teachers: Voices from Indonesia. *System*, *85*, Article 102111. https://doi.org/10.1016/j.system.2019.102111

Compton, L. K. (2009). Preparing language teachers to teach language online: A look at skills, roles, and responsibilities. *Computer Assisted Language Learning*, *22*(1), 73–99.

Deliorman Bakoğlu, R., Boz Taştan, İ., Yiğit, İ., & Yıldız, S. (2009). Tükenmişliği ölçmede alternatif bir araç: Kopenhag tükenmişlik envanterinin Marmara Üniversitesi akademik personeli üzerine uyarlaması [An alternative tool for the measurement of burnout: Adaptation of the Copenhagen Burnout Inventory on Marmara University academics]. *Yönetim Dergisi: İstanbul Üniversitesi İşletme Fakültesi İşletme İktisadı Enstitüsü*, 20(63), 77–98.

Erten, P., & Özdemir, O. (2020). The Digital Burnout Scale development study. *İnönü University Journal of the Faculty of Education*, *21*(2), 668–683. https://www.doi.org/10.17679/inuefd.597890

Esteban-Guitart, M. & Gee, J. P. (2020). "Inside the head and out in the world". An approach to deep teaching and learning. *Multidisciplinary Journal of Educational Research, 10*(1), 1-25. doi:10.4471/remie.2020.4868

Evers, A. T., Verboon, P., & Klaeijsen, A. (2017). The development and validation of a scale measuring teacher autonomous behaviour. *British Educational Research Journal, 43*(4), 805– 821.

Goksel, N., & Bozkurt, A. (2019). Artificial intelligence in education: Current insights and future perspectives. In S. Sisman-Ugur & G. Kurubacak (Eds.), *Handbook of research on learning in the age of transhumanism* (pp. 224–236). IGI Global. https://doi.org/10.4018/978-1-5225-8431-5.ch014

Jamal, A. (2023). The role of artificial intelligence (AI) in teacher education: Opportunities & challenges. *International Journal of Research and Analytical Reviews*, *10*(1), 140–146. https://www.researchgate.net/publication/369384184_The_Role_of_Artificial_Intelligence_AI_in_Teacher_Education_Opportunities_Challenges

Kaplan, H. (2018). Teachers' autonomy support, autonomy suppression and conditional negative regard as predictors of optimal learning experience among high-achieving Bedouin students. *Social Psychology of Education*, *21*, 223–255. https://www.doi.org/10.1007/s11218-017-9405-y

Loughran, J. (2019). Pedagogical reasoning: The foundation of the professional knowledge of teaching. *Teachers and Teaching*, *25*(5), 523–535. https://doi.org/10.1080/13540602.2019.1633294

Marcelo, C., Murillo, P., Marcelo-Martínez, P., Yot-Domínguez, C., & Yanes-Cabrera, C. (2023). Micro-celebrities or teacher leaders? An analysis of Spanish educators' behaviors on Twitter. *Electronic Journal of e-Learning*, *21*(4), 258–273. https://doi.org/10.34190/ejel.21.4.2833

Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Occupational Behavior*, *2*, 99–113.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, *52*, 397–422.

Maslach, C., & Leiter, M. P. (2016). Understanding the Burnout experience: Recent Research and Its Implications for Psychiatry. *World Psychiatry, 15*(2), 103–111. https://doi.org/10.1002/wps.20311

Mhlanga, D. (2023, February 11). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4354422

Monereo, C., & Caride, M. (2022). Becoming a professional: Analysis of the reciprocal influence between I-positions and we-positions in a group of university students. *Journal of Constructivist Psychology*, *35*(4), 1347–1370.

Nataraj, P. (2022, February 22). *How schools in India are integrating AI in their curriculum*. Analytics India Magazine. https://analyticsindiamag.com/integrating-ai-curriculum-cbse-international-

Oplatka, I. (2002). Women principals and the concept of burnout: An alternative voice? *International Journal of Leadership in Education*, *5*(3), 211–226.

Parsons, S. A., Hutchison, A. C., Hall, L. A., Parsons, A. W., Ives, S. T., & Leggett, A. B. (2019). U.S. teachers' perceptions of online professional development. *Teaching and Teacher Education*, *82*, 33–42. https://doi.org/10.1016/j.tate.2019.03.006

Pines, A. & Aronson, E. (1981). Career burnout: causes and cures. New York: Free Press.

Powell, C. G., & Bodur, Y. (2019). Teachers' perceptions of an online professional development experience: Implications for a design and implementation framework. *Teaching and Teacher Education*, *77*, 19–30. https://doi.org/10.1016/j.tate.2018.09.004

Qadir, J. (2022, December 30). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. *TechRxiv*. https://doi.org/10.36227/techrxiv.21789434.v1

Richards, J. C., & Farrell, T. S. (2005). *Professional development for language teachers: Strategies for teacher learning*. Cambridge University Press.

Ruiz-Alfonso, Z., León, J., Santana-Vega, L. E., & González, C. (2021). Teaching quality: An explanatory model of learning in secondary education. *Psicología Educativa*, *27*, 67–76. https://www.doi.org/10.5093/psed2020a18

Ryan, R. M., Deci, E. L., & Vansteenkiste, M. (2016). Autonomy and autonomy disturbances in self-development and psychopathology: Research on motivation, attachment, and clinical process. *Developmental psychopathology*, *1*, 385-438.

Salas-Pilco, S. Z., Xiao, K., & Hu, X. (2022). Artificial intelligence and learning analytics in teacher education: A systematic review. *Education Sciences*, *12*(8), Article 569. https://doi.org/10.3390/educsci12080569

Sandoval-Cruz, R. I., Navarro Rangel, Y., González Calleros, J. M., & Perales-Escudero, M. D. (2022). Pre-service EFL teachers' conceptions of learning about SLA through online discussion forums and WhatsApp in blended learning. *RELC Journal*, *53*(1), 165–179.

Serter, S. (2021). *Denizli'de görev Yapan İlkokul ve Ortaokul Öğretmenlerinin Tükenmişlik Düzeyleri* [Burnout levels of primary and secondary school teachers working in Denizli]; [Unpublished non-thesis master's project]. *Pamukkale University*.

Sharma, M. K., Anand, N., Ahuja, S., Thakur, P. C., Mondal, I., Singh, P., Kohli, T., & Venkateshan, S. (2020). Digital burnout: COVID-19 lockdown mediates excessive technology use stress. *World Social Psychiatry*, *2*(2), 171–172. https://www.doi.org/10.4103/WSP.WSP_21_20

Shin, D.-S., & Kang, H.-S. (2018). Online language teacher education: Practices and possibilities. *RELC Journal*, *49*(3), 369–380. https://doi.org/10.1177/0033688217716535

Şengün, C. (2021, April 16). Dijital Tükenmişlik: Başa Çıkmak İçin 5 Yöntem [Digital burnout: 5 ways to cope]. *The Magger*. https://www.themagger.com/dijital-tukenmislik-basa-cikma-yontemler

Sølvik, R. M., & Glenna, A. E. H. (2021). Teachers' potential to promote students' deeper learning in whole-class teaching: An observation study in Norwegian classrooms. *Journal of Educational Change*, *23*, 343–369. https://www.doi.org/10.1007/s10833-021-09420-8

Tang, K. Y., Chang, C. Y., & Hwang, G. J. (2021). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, *31*(4), 2134–2152. https://doi.org/10.1080/10494820.2021.1875001

Terra, M., Baklola, M., Ali, S., & El-Bastawisy, K. (2023). Opportunities, applications, challenges and ethical implications of artificial intelligence in psychiatry: A narrative review. *The Egyptian Journal of Neurology, Psychiatry and Neurosurgery, 59*(1), Article 80. https://doi.org/10.1186/s41983-023-00681-z

Thurzo, A., Strunga, M., Urban, R., Surovkova, J., & Afrashtehfar, K. (2023). Impact of artificial intelligence on dental education: A review and guide for curriculum update. *Education Sciences*, *13*(2), Article 150. https://doi.org/10.3390/educsci13020150

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*(1), Article 15. https://doi.org/10.1186/s40561-023-00237-x

Wylleman, P., Harwood, C. G., Elbe, A. M., Reints, A., & De Caluwé, D. (2009). A perspective on education and professional development in applied sport psychology. *Psychology of Sport and Exercise*, *10*(4), 435–446. https://www.doi.org/10.1016/j.psychsport.2009.03.008

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zhang, Z., Zhu, H., Zhou, Z., & Zou, K. (2022). How does innovation matter for sustainable performance? Evidence from small and medium-sized enterprises. *Journal of Business Research, 153*, 251–265. https://doi.org/10.1016/j.jbusres.2022.08.034.

Zhao, J., & Qin, Y. (2021). Perceived teacher autonomy support and students' deep learning: the mediating role of self-efficacy and the moderating role of perceived peer support. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.652796

International Review of Research in Open and Distributed Learning
Volume 25, Number 3

August – 2024

AI-Supported Online Language Learning: Learners' Self-Esteem, Cognitive-Emotion Regulation, Academic Enjoyment, and Language Success

Ting Xiao[1,2], Sisi Yi[3,*], and Shamim Akhter[4]
*[1]Department of Genetic Engineering, Faculty of Biology-Oriented Science and Technology, KINDAI University, Wakayama, Japan; [2]Graduate School of Bioscience Major in Bioscience, Nagahama Institute of Bio-Science and Technology, Shiga, Japan; [3]Yangzhou Technician College, Yangzhou, China; [4]Universiti Utara Malaysia, Sintok, Malaysia; *Corresponding author*

Abstract

The consideration of students' emotional and psychological health is crucial to facilitate effective teaching and grading practices. This study set out to shed light on the interplay between self-esteem (S-E), cognitive-emotion regulation (CER), academic enjoyment (AE), and language success (LS) in artificial intelligence (AI)-supported online language learning. To this end, the foreign language learning self-esteem scale, the Cognitive Emotion Control Questionnaire, the foreign language enjoyment scale, and a researcher-made test were distributed to 389 English as a foreign language learners in China. Screening the data with confirmatory factor analysis and structural equation modeling, the effects of S-E, CER, AE, and LS were identified and quantified. These results highlighted the important function that online courses assisted by AI perform in enhancing students' CER and AE. This implied that students who have cultivated a robust sense of self-efficacy are adept at effectively regulating their cognitive and affective processes in AI-supported language learning. Possible improvements in language education are discussed, as are the study's broader implications.

Keywords: self-esteem, cognitive-emotion regulation, academic enjoyment, language success, AI-supported online language learning

Background AI-Supported Online Language Learning: Learners' Self-Esteem, Cognitive-Emotion Regulation, Academic Enjoyment, and Language Success

The emergence of large language models and chatbots, such as ChatGPT, has had a profound impact on scientific discussions, elevating artificial intelligence (AI) to a prominent position. Nevertheless, it is important to acknowledge that research in the field of AI has been ongoing for several decades. As a consequence, there has been a growing fascination with the potential of AI to assist in the field of education, particularly in terms of its ability to personalize learning. However, there are also apprehensions regarding the implications for academic integrity, security, and privacy on a broader scale. The implementations of AI in education have led to considerable improvements in areas including administration, curriculum creation, and assessment (Luan et al., 2020; Qiao & Zhao, 2023).

Regrettably, the issue of AI in language instruction has received little academic attention, and the existing studies have only recently started to pinpoint crucial areas that need further investigation. Therefore, it is important to further explore the possibilities of AI applications to find out how they may be used to teach language skills useful in the actual world. It is also crucial to think about how the AI programs' real-time feedback may help students achieve their language learning objectives, as well as how to improve students' self-study on their computers and mobile devices via the use of different types of feedback. This research analyzed the benefits presented by utilizing AI to build language abilities in language learning settings. Specifically, it delved into the depths to which AI might encourage students to improve their self-esteem (S-E), cognitive-emotion regulation (CER), academic enjoyment (AE), and language success (LS).

Literature Review

The concept of S-E refers to an individual's evaluation of their own worth or value. More specifically, it denotes the extent of individuals' belief in their capabilities and qualities (Bracken, 2009). S-E has had a significant influence on several aspects of student life, including decision-making, relationships, emotional well-being, and general state of health. Furthermore, it impacts motivation, since those with a robust and optimistic self-perception comprehend their capabilities and may be inclined to embrace new endeavors (Rubio-Alcalá, 2017). According to Murk (2006), learners with stronger S-E exhibited higher self-assurance and established more ambitious goals for themselves, even when confronted with challenges and obstacles. Persisting in their endeavors has proven important in enhancing learners' overall effectiveness. Moreover, Hosseinmardi et al. (2021) as well as Prasad et al. (2023) suggested that S-E had the potential to enhance individuals' self-regulatory abilities and emotional well-being. The connection between S-E and reading comprehension has been closely linked to students' autonomy, as highlighted by Zhang (2022). Qiao and Zhao (2023) indicated that the use of AI-based instruction was critical in enhancing oral skills and promoting self-assisted processes in language learners.

This highlights the capacity of AI to augment educational experiences, motivating students to assume agency over their learning and employ meta-cognitive strategies in the oral component of language acquisition. Furthermore, Yang et al. (2022) established a novel task-oriented voice chatbot called Ellie that served as an English conversation companion. The results of their investigation indicated that students exhibited a notable degree of enthusiasm in their engagement with Ellie. Additionally, the chatbot's task design and operational goals were shown to be suitable, as seen by the high rates of task success and S-E improvement. Their study highlighted the advantageous possibilities of chatbots in English as a foreign language (EFL) settings. Martínez-Ramón et al. (2022) supported the efficacy of artificial neural networks in forecasting psychological characteristics such as S-E; enjoyment with education was also supported.

CER in learners refers to deliberate cognitive strategies used to control the processing of emotionally intense information (Aldao & Nolen-Hoeksema, 2010; Garnefski et al., 2004). CER encompasses several internal and external mechanisms that are responsible for monitoring, assessing, and altering emotional responses, particularly in terms of their intensity and duration (Rezaei & Zebardast, 2021). Scholars have recently begun to investigate the cognitive components of emotion regulation, as opposed to other types of methods, such as behavioral strategies (e.g., Griffiths et al., 2021; Rezaei & Zebardast, 2021; Weidi & JeeChing, 2023). Extensive research over time has proven the essential link between mental procedures and emotion control in individuals. This cognitive regulation is critical for individuals to successfully regulate and retain control over their emotions, especially in the face of dangerous or difficult conditions (Weidi & JeeChing, 2023).

The concept of enjoyment serves as a substitute for the emotional response elicited by the successful completion of a task, as discussed in the field of positive psychology (Pekrun, 2006). As Jiang and Dewaele (2019) highlighted, the idea of enjoyment is intricate and diverse, as it is influenced by five separate mechanisms: (a) emotional, (b) mental, (c) inspirational, (d) interpersonal, and (e) physical. Previous research conducted by Elahi Shirvan et al. (2020) confirmed the beneficial impact of pleasure on several aspects of students' educational experiences, including learning attitudes, motivation levels, interactions with peers and teachers, as well as general well-being.

The level of enjoyment experienced in language lessons undergoes a transformation over time, influenced by factors such as the learner's characteristics and the instructional environment (Dewaele et al., 2018). Furthermore, Macintyre et al. (2019) emphasized that the degrees of pleasure and terror experienced by students in a language classroom have a substantial influence on their levels of involvement and performance. The alignment between context-oriented attributes and the cognitive needs of students serves as a driving force for their engagement and enjoyment in the classroom (Chen et al., 2021). Furthermore, the interactions between teachers and students have a pivotal role in fostering a positive and engaging learning environment inside the classroom (Elahi Shirvan et al., 2020).

Additionally, the connections among enjoyment, motivation, positive attitudes, and language success were confirmed by Liu (2022). Higher levels of self-evaluation and thinking skills were associated with greater enjoyment and more effective immunization among EFL college students (Aldosari et al., 2023). In their review of the relevant literature, Heeg and Avraamidou (2023) determined that the use of AI may increase students' enjoyment of in-class participation. Customizing and individualizing lessons and assignments following a learner's skills, aptitudes, and capacities is an essential use of AI that has made education more effective (Neji et al., 2023). AI provides students with a more pleasurable and engaging or immersive learning experience, which in turn improves the student's ability to absorb and remember knowledge, which is the cornerstone of education (Kizilcec, 2023).

The Purpose of this Research

Research has identified a number of factors that determine academic achievement. According to Wei (2020), when students actively participated in S-E, CER, and AE, they had improved learning outcomes in the cognitive, meta-cognitive, and emotional zones. While it is clear that S-E, CER, and AE enhance students' well-being, there is a dearth of research that has investigated the connections between and among these elements in the context of AI-supported online education. This lacuna in the literature prompted the undertaking of this study which aimed to examine the relationships among S-E, CER, AE, and LA in the context of AI-supported online language development. The outcomes were expected to offer evidence to formulate significant conclusions about language training and its assessments in AI-supported online language learning. This inquiry was framed by a single research question. Do higher levels of self-esteem foster the cognitive-emotion regulation, academic enjoyment, and language success in AI-supported online language learning?

Methodology

The current investigation was undertaken with a cohort of 389 students who attended classes in private language institutions at the intermediate level in China. The participants were selected using either opportunity sampling or convenience sampling methods. Duolingo as well as ChatGPT were employed to support their online instruction. There were 201 female participants and 188 men. Their ages ranged from 20 to 27 years old. The participants had sufficient skills to complete the questionnaires in English.

This investigation commenced in February 2023, and persisted until June of the same year. The procedure was executed via a Web-based system linked to the Internet. The questionnaire included the foreign language learning self-esteem scale (FLLSE), the Cognitive Emotion Control Questionnaire (CERQ), the foreign language enjoyment scale (FLES), and a researcher-made test. The receipt of 389 fully completed forms resulted in a return percentage of 76.4%.

FLLSE was used to explore the depth of students' S-E. This tool was designed by Rubio (2007), and uses a five-point Likert scale ranging from one (*strongly disagree*) to five (*strongly agree*). The FLLSE is comprised of a total of 25 questions broken down into four categories. In our investigation, the dependability of this instrument was found to be satisfactory ($\alpha = 0.881$).

Using CERQ, developed by Garnefski et al. (2002), we analyzed the cognitive emotion control techniques participants applied while facing threatening or upsetting life experiences. The CERQ is a 36-item survey that divides a person's reactions to threatening or stressful life events into nine conceptually distinct subscales, each of which consists of four items. Strategies for controlling one's emotions via the use of cognition were evaluated using a five-point Likert scale ranging from five (*almost never*) to one (*almost always*). The internal consistencies of the various subscales were all within the desirable range, from 0.783 to 0.914.

FLES, created and validated by Dewaele and MacIntyre (2016), was used to assess the level of pleasure experienced by students learning foreign languages. The FLES questionnaire comprises 21 questions rated on a five-point Likert scale, from *strongly disagree* to *strongly agree*. The analysis of Cronbach's alpha coefficient came in at 0.789, a satisfactory level of dependability.

To evaluate the participants' LA, the researchers devised a test that was based on the students' classroom materials as well as the participants' existing levels of competence in the target language. Three EFL instructors were asked to evaluate the validity and reliability of this exam; their suggestions were implemented to improve the test. This exam was divided into four parts to assess listening, speaking, reading, and writing skills. The overall score was 20 points.

The Kolmogorov-Smirnov (K-S) test was conducted to ascertain if the data adhered to a normal distribution. The results of further examination revealed that the data followed a normal pattern. Consequently, the statistical techniques of confirmatory factor analysis (CFA) and structural equation modeling (SEM) were applied to screen the data. CFA was used to validate the observed variables' component structure, and to determine whether there was a connection among the variables and their latent constructs. Linear structural relations (LISREL) 8.80 software, a proprietary statistical software package used in SEM for manifest and latent variables (Jöreskog, 1990), was used throughout these analyses.

Results

In this section we provide information on the reports that were generated from the data analysis, along with in-depth explanations for each component. First, we examine descriptive data relevant to the various parts of each instrument (Table 1).

Table 1

Descriptive Statistics

Variable	*N*	Minimum	Maximum	Mean	Std. deviation
Language capability	389	6	30	19.165	6.445
Real in-class language use	389	6	30	20.044	5.585
In-class correlations	389	6	30	20.195	5.404
Attitude toward behavior in the class	389	7	35	23.638	6.064
Self-esteem	389	29	125	83.041	20.934
Self-blame	389	4	20	13.614	3.805
Other-blame	389	4	20	13.272	4.023
Rumination	389	4	20	13.630	3.948
Catastrophizing	389	4	20	13.563	3.500
Putting into perspective	389	7	20	15.185	2.430
Positive refocusing	389	4	20	14.648	3.411
Positive reappraisal	389	8	20	14.409	2.506
Acceptance	389	7	20	15.604	3.417

Planning	389	5	20	13.887	3.956
Cognitive-emotion regulation	389	80	167	127.812	17.527
Academic enjoyment	389	26	105	71.689	15.606
Academic success	389	9	20	17.129	2.836

The most frequent answer, after accounting for SE, was attitude toward behavior in the class (*M* = 23.638, *SD* = 6.064). When the main variables of the CER scale were dissected into their constituent pieces, acceptance had the highest mean value (*M* = 15.604, *SD* = 3.417) of the variables in the scale. On the third instrument, the results for AE were *M* = 71.689 and *SD* = 15.606. The mean score for AS, was 17.129, while the associated standard deviation was 2.836.

Next, the data was subjected to the K-S test to search for unusual patterns. The results are shown in Table 2.

Table 2

Results of the K-S Test

Variable	K-S Z Score	Sig. (2-tailed)
Language capability	1.196	0.114
Real in-class language use	1.024	0.245
In-class correlations	1.162	0.134
Attitude toward behavior in the class	0.896	0.398
Self-esteem	0.777	0.582
Self-blame	1.104	0.175
Other-blame	1.316	0.063
Rumination	1.168	0.131
Catastrophizing	1.045	0.225
Putting into perspective	1.236	0.082
Positive refocusing	1.099	0.201
Positive reappraisal	1.304	0.072
Acceptance	1.238	0.092
Planning	1.031	0.238
Cognitive-emotion regulation	0.712	0.691
Academic enjoyment	0.566	0.905
Academic success	0.822	0.508

All instruments and their components exhibited statistically non-significant values, as shown in Table 2. This observation may suggest that parametric approaches were well-suited for the analysis of the data. Pearson product-moment correlation was then used to examine the connections among S-E, CER, AE, and AS.

Table 3

Correlation Coefficients: S-E, CER, AE, and AS

Variable	1	2	3	4	5	6	7
1. Language capability	—						
2. Real in-class language use	0.543**	—					
3. In-class correlations	0.613**	0.625**	—				
4. Attitude toward behavior in the class	0.578**	0.607**	0.615**	—			
5. Cognitive-emotion regulation	0.856**	0.825**	0.813**	0.894**	—		
6. Academic enjoyment	0.753**	0.788**	0.675**	0.704**	0.601**	—	
7. Academic success	0.635**	0.665**	0.573**	0.604**	0.635**	0.557**	—

**Correlation is significant at the 0.01 level (2-tailed)

The findings summarized in Table 3 revealed notable associations among the different components of the subscales of S-E, CER, AE, and AS.

Subsequently, a causal analytic framework and SEM were used to investigate the interplay among S-E, CER, AE, and AS. The statistical analysis was conducted using LISREL 8.80. Assessment of the concordance between the model and the data included the use of many measures, such as the size of the chi-squared statistic, the root mean squared error of approximation (RMSEA), the goodness of fit index (GFI), the normed fit index (NFI), and the comparative fit index (CFI).

Table 4

Model-Fitness Measures (Model 1)

Fitting index	$\chi2$	df	$\chi2$/df	RMSEA	GFI	NFI	CFI
Cut value			< 3	< 0.1	> 0.9	> 0.9	> 0.9
Model 1	4020.72	1374	2.926	0.070	0.953	0.962	0.948
Model 2	7496	2609	2.873	0.069	0.942	0.961	0.973

Table 4 displays the outcomes, showing that all fitness levels for Model 1 were suitable. These values consisted of the chi-square/*df* ratio (2.926), RMSEA (0.070), GFI (0.953), NFI (0.962), and CFI (0.948). Furthermore, as shown in Table 4, the Model 2 parameters have been fulfilled, suggesting a satisfactory correspondence. The chi-square/*df* ratio (2.873), RMSEA (0.069), GFI (0.942), NFI (0.961), and CFI (0.973) were the parameters evaluated.

Figure 1

Schematic Visualization of Path Coefficient Values (Model 1)

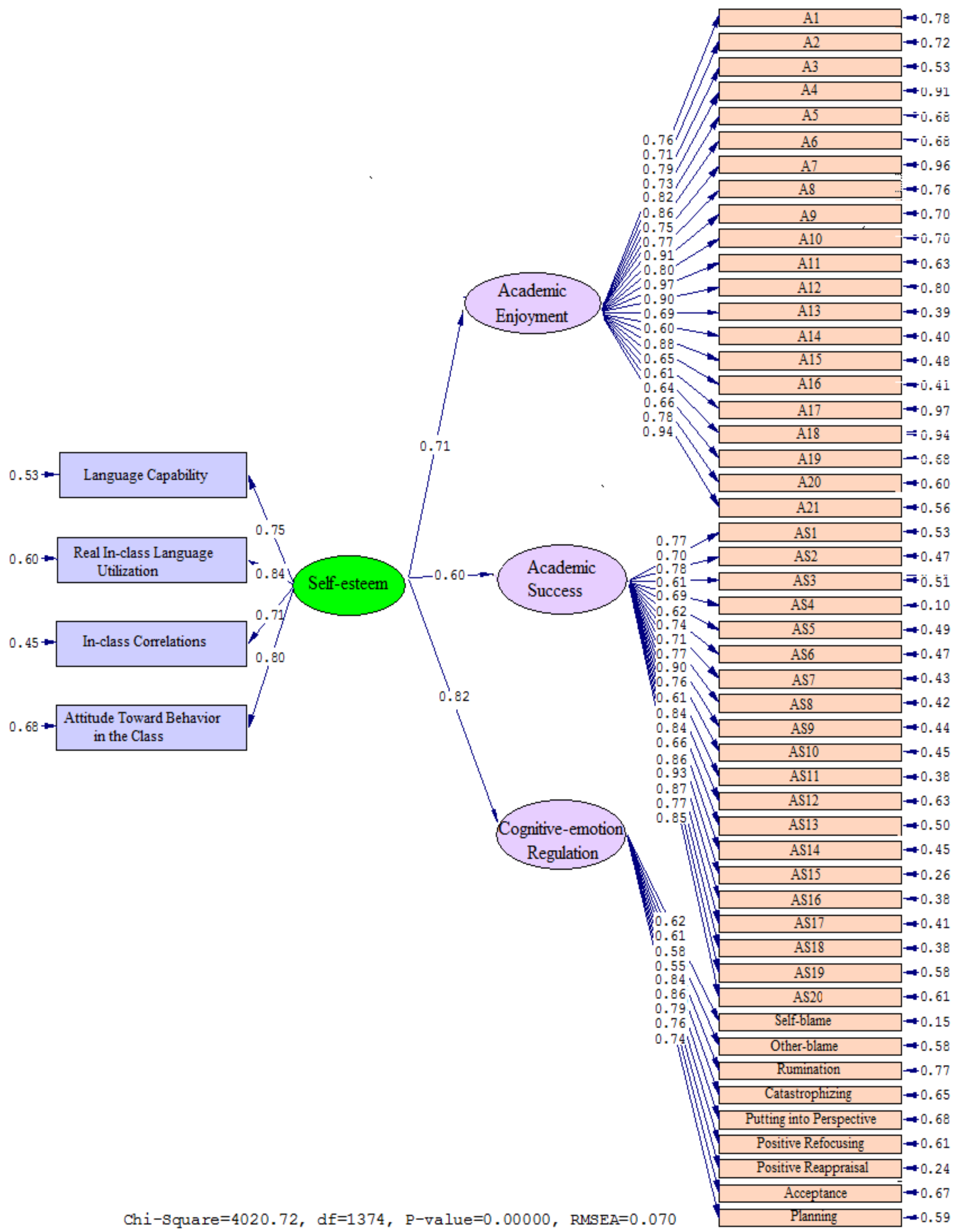


Figure 2

Path Coefficient Importance t-Values (Model 1)

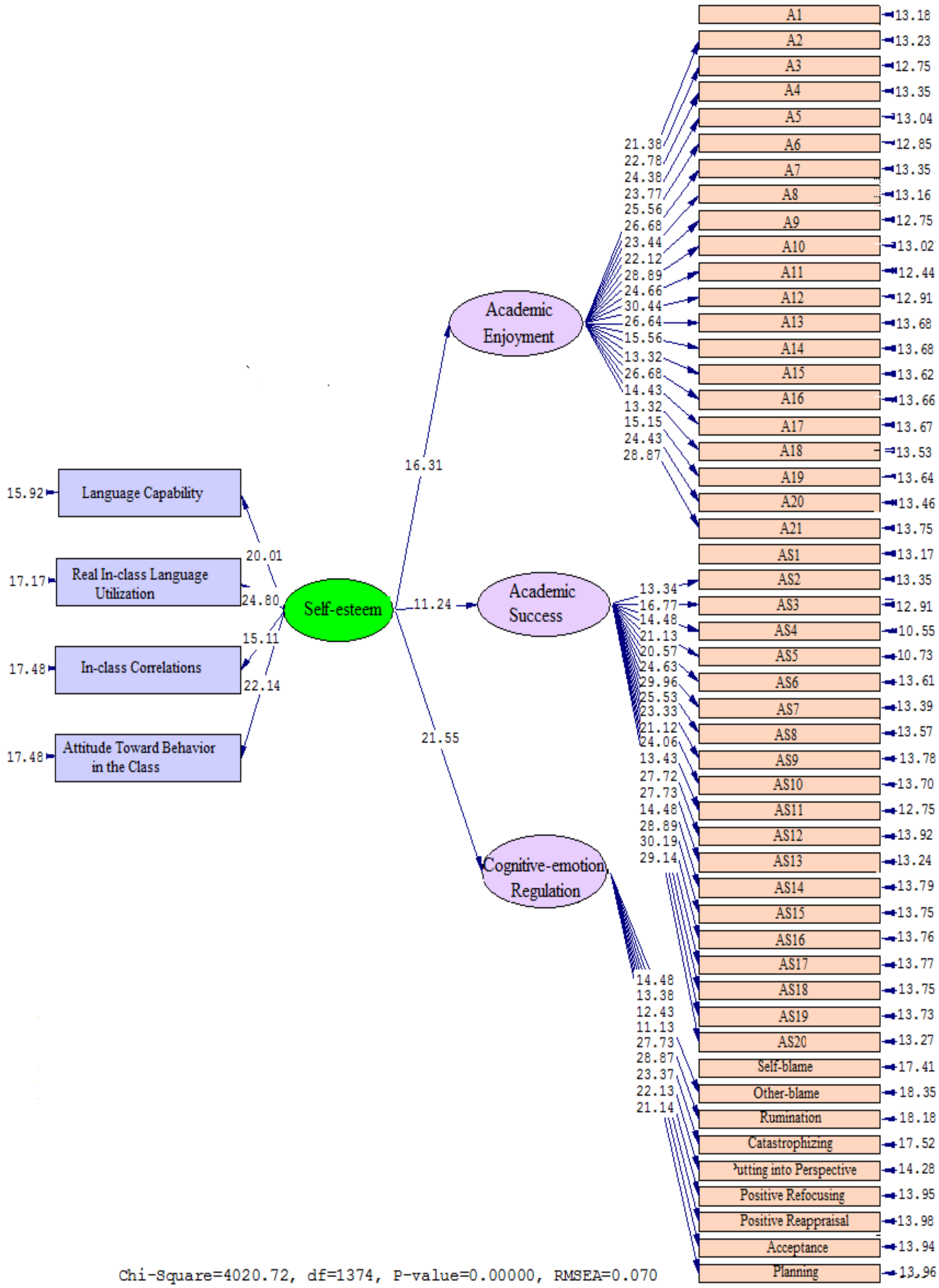


Table 5

Synopsis of the Results Obtained From Model 1

Path			Path coefficient	*t*-Statistic	Test result
Self-esteem	→	Cognitive-emotion regulation	0.82	21.55	Supported
Self-esteem	→	Academic enjoyment	0.71	16.31	Supported
Self-esteem	→	Academic success	0.60	11.24	Supported

Figures 1 and 2 provide a graphical illustration of the connections among the elements, and Table 5 provides additional information. The standard estimates and *t*-values mirror a notable connection between S-E and CER ($\beta = 0.82$, $t = 21.55$), AE ($\beta = 0.71$, $t = 16.31$), and AS ($\beta = 0.60$, $t = 11.24$).

Figure 3

Schematic Visualization of the Path Coefficients' Values (Model 2)

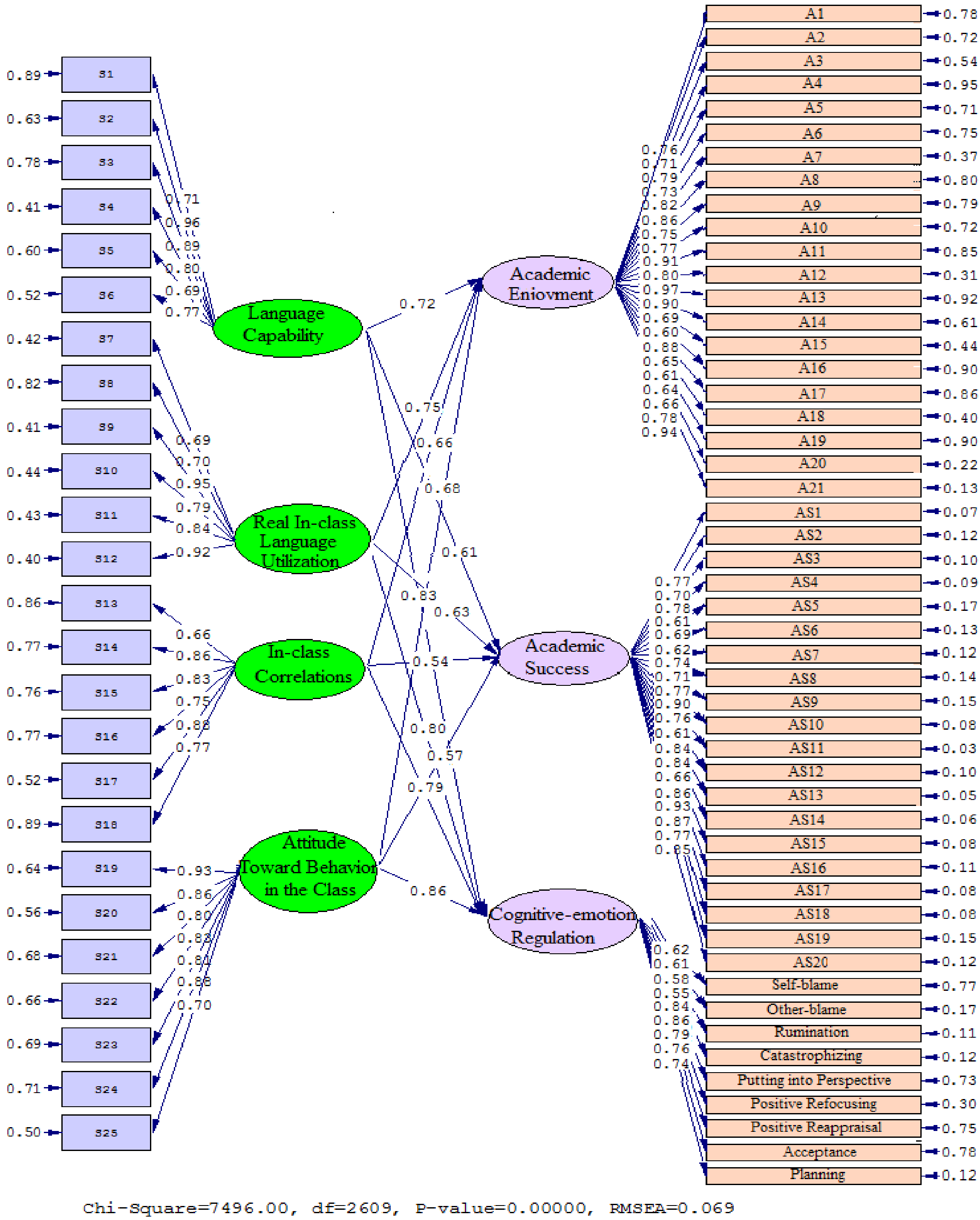


Figure 4

Path Coefficient Importance t-Values (Model 2)

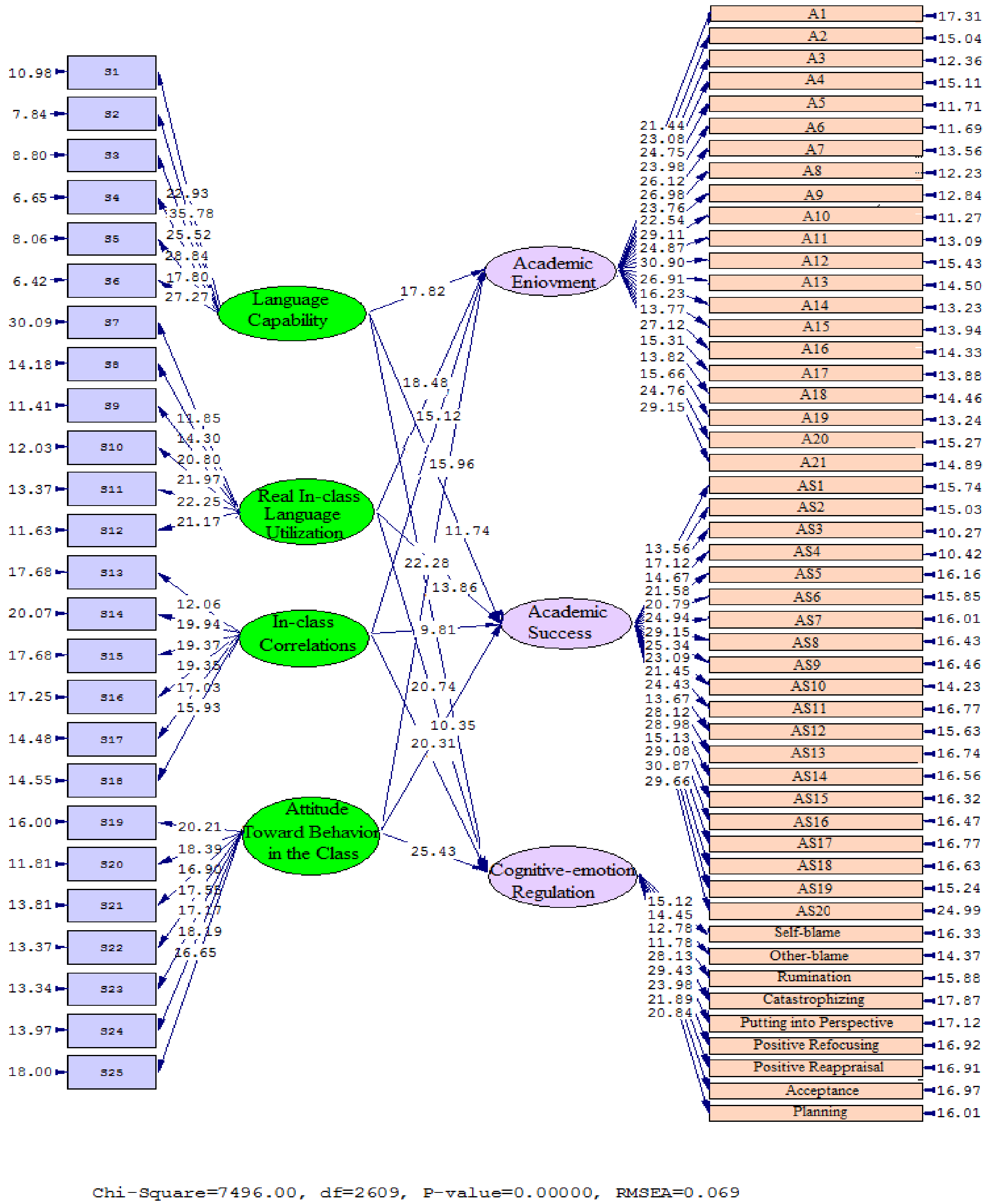


Table 6

Synopsis of the Results Obtained From Model 2

Path			Path coefficient	*t*-Statistic	Test result
Language capability	→	Cognitive-emotion regulation	0.83	22.28	Supported
Real in-class language use	→	Cognitive-emotion regulation	0.80	20.74	Supported
In-class correlations	→	Cognitive-emotion regulation	0.79	20.31	Supported
Attitude toward behavior in the class	→	Cognitive-emotion regulation	0.86	25.43	Supported
Language capability	→	Academic enjoyment	0.72	17.82	Supported
Real in-class language use	→	Academic enjoyment	0. 75	18.48	Supported
In-class correlations	→	Academic enjoyment	0.66	15.12	Supported
Attitude toward behavior in the class	→	Academic enjoyment	0.68	15.96	Supported
Language capability	→	Academic success	0.61	11.74	Supported
Real in-class language use	→	Academic success	0.63	13.86	Supported
In-class correlations	→	Academic success	0.54	9.81	Supported
Attitude toward behavior in the class	→	Academic success	0.57	10.35	Supported

The results indicated a significant and favorable correlation between CER and the subsequent sub-factors: language Capability (β = 0.83, t = 22.28), real in-class language Use (β = 0.80, t = 20.74), in-class correlations (β = 0.79, t = 20.31), and attitude toward behavior in the class (β = 0.86, t = 25.43). Likewise, a statistically significant relationship was observed between AE and various sub-scales, namely language capability (β = 0.72, t = 17.82), real in-class language use (β = 0. 75, t = 18.48), in-class correlations (β = 0.66, t = 15.12), and attitude toward behavior in the class (β = 0.68, t = 15.96). In line with the findings, there were positive and statistically noteworthy connections between AS and the following sub-components: language capability (β = 0.61, t = 11.74), real in-class language use (β = 0.63, t = 13.86), in-class correlations (β = 0.54, t = 9.81), and attitude toward behavior in the class (β = 0.57, t = 10.35).

Discussion

The purpose of this research was to examine how S-E, CER, AE, and LS are connected. More specifically, the influence of S-E on CER, AE, and LS in AI-supported online language learning was explored. To achieve this objective, the study employed empirical research methods among EFL students who were enrolled in English-language institutions. These students used online classes to enhance their English language proficiency. The findings showed that students who developed and practiced S-E improved in CER and AE; they also achieved better scores on their language tests. These findings brought to light the significant role that AI-supported online classes can play in improving students' psychological well-being as well as their engagement in their academic work. The findings of the study are discussed in greater detail below.

This study had one key research question—do higher levels of self-esteem foster the cognitive-emotion regulation, academic enjoyment, and language success in AI-supported online language learning? Our findings suggested that S-E guided CER of EFL learners. That means the strategies encompassed within S-E contributed to a state of balance in students' educational experiences, enabling them to engage in critical evaluation of their learning process. This suggested that students were capable of enhancing their learning outcomes by engaging in critical evaluation of their learning processes. The outcomes of this study suggested that there was a beneficial connection between students' psychological well-being and their self-concept, in addition to their beliefs regarding their monitoring and meta-cognitive skills.

The cognitive and meta-cognitive learning experiences of students, and language learners in particular, can be enhanced when they are provided with emotional support and when they employ appropriate strategies for learning (i.e., AI) in the face of chaos and complexity. It was also agreed that S-E established the standard for language learning and evaluation (Ritonga et al., 2023). Accordingly, ensuring successful language learning and increasing learners' engagement requires an investment in providing related knowledge via efficient materials and applications.

The results also suggested that enhanced S-E increased EFL learners' engagement. In particular, model 2 of CER subfactors reflected S-E in the following areas: (a) course worth, (b) involvement with instructors, (c) involvement with classmates, (d) interactions with others, and (e) involvement with online evaluations. These are some of the attributes of students feeling they belong and are valued. This result aligned with Riswanto et al. (2022), who revealed a substantial influence of S-E and critical thinking on EFL students' motivation. It can be argued that facilitating students' autonomy through the use of online language learning courses contributes to the development of their linguistic abilities. Ismail and Heydarnejad (2023), as well as Soodmand Afshar and Jamshidi (2022), also established direct connections among S-E, learning autonomy, and personal best goals. The promotion of independence and autonomy among students can be facilitated by providing them with the necessary resources to achieve success in their future academic pursuits.

In general, the incorporation of AI applications into language curricula yielded a positive effect on students' academic performance, as evidenced by the successful attainment of the curriculum's intended learning outcomes. AI applications have the potential to alleviate the onerous workload of language educators by eliminating laborious duties like evaluation and feedback. Furthermore, AI applications have the potential to pique the generally apathetic interest of students in language through the provision of interactive and amusing intelligent tutors or unconventional individualized educational settings. AI applications can also assist in optimizing instructional operations to boost generally poor educational results in language subjects.

Moreover, in accordance with Bandura's (1989) social cognitive theory, the process of learning takes place through the acts of observing, imitating, and modeling the behaviors exhibited by others. Within the realm of AI-supported relationships, students are afforded the chance to look into and actively participate in AI systems that exhibit self-regulatory behaviors. These behaviors include providing responsive input and assisting students in determining and planning their goals. Through the process of observing these behaviors, learners have the opportunity to internalize and subsequently replicate self-regulatory strategies. Consequently, this process enables individuals to foster the growth of their self-regulation abilities. Furthermore, AI technologies provide notable advantages, such as dynamic methods of learning and actual time analysis of data. These features allow learners to receive prompt feedback and track their progress. The provision of prompt feedback enables learners to assess their achievement, discover aspects requiring enhancement, and adapt their methods of learning correspondingly. Through the process of self-reflection and incorporating feedback obtained from AI systems, individuals can cultivate cognitive consciousness and self-regulatory practices (Zimmerman, 2002).

Conclusion and Implications

Implementing an AI environment is advisable due to its ability to create an engaging and intriguing technical setting. This will enhance learners' ability to effectively interact with AI and their peers, eventually resulting in increased competence in autonomously evaluating and improving their oral communication skills. The findings of this research indicated that the interactive speaking activities, aided by AI, led to improvements in EFL learners' (a) speaking and listening skills, (b) critical thinking abilities, (c) affective engagement, and (d) language success. This implies that AI technology may help learners oversee and regulate their educational processes. These technologies may assist in establishing objectives, monitoring accomplishments, and implementing any necessary adjustments. AI-driven training empowers learners to take control of their educational process and enhance their oral communication skills by offering personalized coaching and adaptive tasks that promote the development of meta-cognitive strategies.

This research has certain drawbacks that should be noted, in addition to the implications. Nevertheless, it offers intriguing novel insights into the matter at hand. Participants in this research were Chinese EFL students at the intermediate level—future investigations could repeat this study in other classrooms and compare the results. It is recommended that additional areas of inquiry investigate the possible impacts of S-E on CER, AE, and LS. Students' self-reporting survey answers were the only source of data used in the research, raising concerns about generalizability. Integrating qualitative and quantitative approaches can be useful for future studies. Furthermore, it is suggested that the study's findings be pooled and that a future survey investigate how students' socioeconomic backgrounds might influence the correlations among AI-supported language acquisition, CER, AR, and S-E.

Acknowledgments

The authors declare that they used AI applications (i.e., ChatGPT and QuillBot) to edit and proofread some sections to address language accuracy and fluency.

References

Aldao, A., & Nolen-Hoeksema, S. (2010). Specificity of cognitive emotion regulation strategies: A transdiagnostic examination. *Behavior Research and Therapy, 48*(10), 974–983. https://doi.org/10.1016/j.brat.2010.06.002

Aldosari, M., Heydarnejad, T., Hashemifardnia, A., & Abdalgane, M. (2023). The interplay among self-assessment, using refection for assessment, classroom enjoyment, and immunity: Into prospects of effective language learning. *Language Testing in Asia, 13*, 1. https://doi.org/10.1186/s40468-022-00213-1

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, *44*, 1175–1184. https://doi.org/10.1037/0003-066X.44.9.1175

Bracken, B. A. (2009). Positive self-concepts. In R. Gilman, E. S. Huebner, & M. J. Furlong (Eds.), *Handbook of positive psychology in schools*. Routledge. https://doi.org/10.4324/9780203884089

Chen, P., Bao, C., & Gao, Q. (2021). Proactive personality and academic engagement: The mediating effects of teacher-student relationships and academic self-efficacy. *Frontiers in Psychology*, *12*, 1824. https://doi.org/10.3389/fpsyg.2021.652994

Dewaele, J. -M., & MacIntyre, P. D. (2016). Foreign language enjoyment and foreign language classroom anxiety: The right and left feet of the language learner. In P. D. MacIntyre, T. Gregersen, & S. Mercer (Eds.), *Positive psychology in SLA* (pp. 215–236). Multilingual Matters. https://doi.org/10.21832/9781783095360-010

Dewaele, J. -M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Language Teaching Research*, *22*, 676–697. https://doi.org/10.1177/1362168817692161

Elahi Shirvan, M., Taherian, T., & Yazdanmehr, E. (2020). The dynamics of foreign language enjoyment: An ecological momentary assessment. *Frontiers in Psychology, 11*, 1391-1405.

Garnefski, N., Kraaij, V., & Spinhoven, P. (2002). *Manual for the use of the cognitive emotion regulation questionnaire*. DATEC.

Garnefski, N., Teerds, J., Kraaij, V., Legerstee, J., & van den Kommer, T. (2004). Cognitive emotion regulation strategies and depressive symptoms: Differences between males and females. *Personality and Individual Differences, 36*, 267–276. https://doi.org/10.1016/S0191-8869(03)00083-7

Griffiths, S., Suksasilp, C., Lucas, L., Sebastian, C. L., & Norbury, C. (2021). Relationship between early language competence and cognitive emotion regulation in adolescence. *Royal Society Open Science, 8*(10), 210742. https://doi.org/10.1098/rsos.210742

Heeg, D. M., & Avraamidou, L. (2023). The use of artificial intelligence in school science: A systematic literature review. *Educational Media International, 60*(2), 125–150. https://doi.org/10.1080/09523987.2023.2264990

Hosseinmardi, A. A., Ghorban Shiroodi, S., Zarbakhshbahri, M. R., & Tizdast, T. (2021). The relationship of academic engagement, school engagement and school belonging with academic achievement by mediated the academic achievement motivation in male students. *Iranian Journal of Educational Society, 14*(2), 178–189. https://doi.org/10.22034/ijes.2021.534830.1123

Ismail, S. M., & Heydarnejad, T. (2023). Probing into the influence of EFL learners' self-assessment and evaluation apprehension in predicting their personal best goals and self-efficacy skills: A structural equation modeling. *Language Testing in Asia*, *13*, 8. https://doi.org/10.1186/s40468-023-00219-3

Jiang, Y., & Dewaele, J. -M. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners? *System*, *82*, 13–25. https://doi.org/10.1016/j.system.2019.02.017.

Jöreskog, K. G. (1990). New developments in LISREL: Analysis of ordinal variables using polychoric correlations and weighted least squares. *Quality and Quantity 24*, 387–404. https://doi.org/10.1007/BF00152012

Kizilcec, R. (2023). To advance AI use in education, focus on understanding educators. *International Journal of Artificial Intelligence in Education, 34,* 12–19. https://doi.org/10.1007/s40593-023-00351-4

Liu, S. (2022). Toward the role of L2 enjoyment in EFL students' academic motivation and engagement. *Frontiers in Psychology, 12*, 822588. https://doi.org/10.3389/fpsyg.2021.822588

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J. H., Ogata, H., Baltes, J., Guerra, R., Li, P., & Tsai, C. -C. (2020). Challenges and future directions of big data and artificial intelligence in education. *Frontiers in Psychology, 11*, 580820. https://doi.org/10.3389/fpsyg.2020.580820

Macintyre, P. D., Gregersen, T., & Mercer, S. (2019). Setting an agenda for positive psychology in SLA: Theory, practice, and research. *The Modern Language Journal*, *103,* 262–274. https://doi.org/10.1111/modl.12544

Martínez-Ramón, J. P., Morales-Rodríguez, F. M., Ruiz-Esteban, C., & Méndez, I. (2022). Self-esteem at university: Proposal of an artificial neural network based on resilience, stress, and sociodemographic variables. *Frontiers in Psychology, 13,* 815853. https://doi.org/10.3389/fpsyg.2022.815853

Murk, J. C. (2006). *Self-esteem research, theory, and practice: Toward a positive psychology of self-esteem* (3rd ed.). Springer.

Neji, W., Boughattas, N., & Ziadi, F. (2023). Exploring new AI-based technologies to enhance students' motivation. *Issues in Informing Science and Information Technology*, *20*, 95–110. https://doi.org/10.28945/5149

Prasad, K., Aladini, A., Normurodova, N. Z., & Belton, B. (2023). Take language assessment easy: The mediator impacts of self-assessment, test-taking skills in predicting student evaluation apprehension, foreign language learning self-esteem, and language achievement in online classes. *Language Testing in Asia*, *13*, 34. https://doi.org/10.1186/s40468-023-00246-0

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: Illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology, 14*, 1255594. https://doi.org/10.3389/fpsyg.2023.1255594

Rezaei, S., & Zebardast, A. (2021). The mediating role of cognitive emotion regulation strategies on mindfulness, anxiety, and academic procrastination in high schoolers. *Practice in Clinical Psychology, 9*(2), 133–142. https://doi.org/10.32598/jpcp.9.2.731.1

Riswanto, Heydarnejad, T., Saberi Dehkordi, E., & Parmadi, B. (2022). Learning-oriented assessment in the classroom: The contribution of self-assessment and critical thinking to EFL learners' academic engagement and self-esteem. *Language Testing in Asia, 12*(1). https://doi.org/10.1186/s40468-022-00210-4

Ritonga, M., Shaban, A. A., Al-Rashidi, A. H., & Chilani, N. (2023). Engagement in on-line language assessment: Are test-taking skills, self-assessment, resilience, and autonomy critical? *Language Testing in Asia, 13,* 25. https://doi.org/10.1186/ s40468-023-00236-2

Rubio-Alcalá, F. D. (2017). Self-esteem and self-concept in foreign language learning. In C. Gkonou, M. Daubney, & J. -M. Dewaele (Eds.), *New insights into language anxiety: Theory, research, and educational implications* (pp. 198–206). Multilingual Matters.

Rubio, F. (2007). Self-esteem and foreign language learning: An introduction. In F. Rubio (Ed.), *Self-esteem and foreign language learning* (pp. 2–12). Cambridge Scholars Publishing.

Soodmand Afshar, H., & Jamshidi, B. (2022). EFL learners' language learning strategy use, instrumental motivation, self-efficacy, self-regulation, autonomy, and L2 achievement: A SEM analysis. *Applied Research on English Language, 11*(4), 133–160. https://doi.org/10.22108/are.2022.133123.1880

Wei, B. (2020). COVID-19 and online teaching in higher education: A case study of Peking University. *Human Behavior and Emerging Technologies*, *2*(2), 113–115. https://doi.org/10.1002/hbe2.191

Weidi, Z., & JeeChing, P. (2023). Cognitive emotion regulation and life satisfaction among students from Mainland China in Malaysian universities. *Frontiers in Education, 8,* 1028548. https://doi.org/10.3389/feduc.2023.1028548

Yang, H., Kim, H., Lee, J. H., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL, 34*, 327–343. https://doi.org/10.1017/S0958344022000039

Zhang, F. (2022). A theoretical review on the impact of EFL/ESL students' self-sabotaging behaviors on their self-esteem and academic engagement. *Frontiers in Psychology, 13,* 873734. https://doi.org/10.3389/fpsyg.2022.873734.

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

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

The Effects of Duolingo, an AI-Integrated Technology, on EFL Learners' Willingness to Communicate and Engagement in Online Classes

Zhiqun Ouyang[1,3], Yujun Jiang[2,3], and Huying Liu[2,3,*]
*[1]School of Foreign Languages, Hunan International Economics University, Changsha, China; [2]SWAN College, Central South University of Forestry & Technology, Changsha, China; [3]Graduate School, Stamford International University, Bangkok, Thailand; *Corresponding author*

Abstract

This study, which is quasi-experimental in nature, looks into how language learners' willingness to communicate and engagement in English as a foreign language (EFL) classrooms are affected by Duolingo. The control and experimental groups comprised two complete classes with forty EFL students. To compare learner engagement and communication willingness scores before and after treatment, the study used independent samples *t*-tests. The results demonstrated the groups' initial homogeneity by showing no discernible differences prior to the intervention. The results confirmed the effects on learner engagement, which showed significant gains in affective, cognitive, and behavioral domains, indicating Duolingo's beneficial impact on engagement in general. Furthermore, the significant effect sizes observed confirmed Duolingo's contribution to improved language attitudes, engagement, and communicative confidence. Compared to the control group, the experimental group's willingness to speak, read, write, comprehend, and communicate generally improved in a manner that was statistically significant. The significant effect sizes demonstrate how well Duolingo works to improve different aspects of willingness to share. The study emphasizes the pedagogical tool's adaptability and encourages teachers to integrate Duolingo for a comprehensive and technologically enhanced language learning experience. Practical implications arise for EFL teachers who use online learning resources.

Keywords: AI-integrated technology, AI, Duolingo, willingness to communicate, engagement, online class, EFL

Introduction

The rapid advancement of technology has drastically changed the way that languages are taught and acquired in today's ever-changing educational environment. Educators have access to cutting-edge methods and resources that stimulate students' curiosity and push their intellectual limits (Blake, 2013, 2016; Stanley, 2013). Acquiring new languages is made easier for learners by the abundance of authentic and creative resources (Walker & White, 2013). According to Larsen-Freeman and Anderson (2011), the use of technology in language education has improved learning experiences by giving students more access to the target language and letting them advance at their own speed. Technical resources such as podcasts, vodcasts, and online dictionaries are also being made available to teachers.

The Internet has become a change agent and a better means of education in the modern era of multimedia and the Internet (Adesote & Fatoki, 2013). Teachers need to embrace this change and be creative and inventive in order to keep up with the modern demands of education. Information and communication technology (ICT) has advanced as a result of the shift in educational tools and platforms from traditional approaches to technology-driven programs such as Memrise, Babbel, HelloTalk, and Duolingo (Yunus et al., 2009). In order to enable students to study anytime, anywhere, and at their convenience, mobile devices—such as laptops, tablets, smartphones, and other gadgets—play a vital role in education (Ogata & Yano, 2004; Yang, 2006). More than ten years ago, Keegan (2003) predicted that mobile learning (m-learning), a subset of e-learning, would influence education. Wagner (2005) noted that there is no denying the influence of mobile devices on different groups of people, indicating a mobile revolution in the field of education.

New smartphone apps are constantly improving language learning, making it more engaging and customized for users as technology advances. This essay examines the possibilities offered by the smartphone language learning app Duolingo (https://www.duolingo.com/) for acquiring a second language. Users who speak English have a choice of 16 languages, from French to Esperanto. Other language speakers, however, have fewer choices. While Spanish speakers have six language options and French speakers have four, English speakers can choose from 16 languages. In contrast to comparable apps like Babbel and Busuu, Duolingo offers a wider selection of languages despite this variability. This wide variety is a result of the diverse learning community on Duolingo as well as the platform's encouragement of user contributions, which promotes continuous expansion of the language courses that are offered.

The world's most popular language learning app, Duolingo, with over 300 million users, is the subject of this paper's exploration of mobile language learning. Specifically focused on teaching English as a second or foreign language, Duolingo is a game-based application that strives to make education free, enjoyable, and accessible. Duolingo uses technology to create a fully immersive and technologically advanced learning environment, including compatibility with computers and mobile phones. But even with Duolingo's extensive use, there is a noticeable lack of research on the precise effects of the app on the willingness to communicate (WTC) and general engagement of English as a foreign language (EFL) learners. Although Duolingo was shown to be beneficial in teaching vocabulary, grammar, and language proficiency in earlier research, a thorough investigation of the complex dynamics underlying learners' communication willingness and sustained engagement is still lacking. To close this gap, this paper examines how learners' attitudes toward communication are shaped by Duolingo and how much attention it can generate. Two research questions were raised:

1. Does EFL learners' use of Duolingo affect their willingness to communicate in online classes?
2. Does EFL learners' use of Duolingo affect their engagement in online EFL classes?

Review of Literature

The study is mainly based on two leading theories. The first theory is a sociocultural and cultural theory, which posits a significant connection between an individual's psychology and the cultural and institutional context in which they are situated (Scott & Palincsar, 2013). As Ahmed (2017) articulated, culture encompasses inherited beliefs and practices that exert a substantial influence on the course of our lives. Central to this theory is the role of social interactions and cultural engagements in shaping psychological development. It underscores that development is not solely an internal process but is profoundly impacted by external social interactions. The surroundings in which individuals find themselves play a pivotal role in shaping behavior and learning. In this view, language mirrors and communicates the very fabric of culture.

Numerous contemporary domains of investigation and expression are increasingly reliant on and bolstered by computational resources. An emergent common ground has already materialized, encompassing artificial intelligence (AI), learning analytics, educational data mining, machine learning, and complexity theory (Dawson et al., 2018; Tsai et al., 2019).

Feng and Law's (2021) analysis of AI in education research from 2010 to 2019 highlighted a diverse range of research topics, primarily focused on intelligent tutorial systems and massive open online courses. Keywords such as neural networks, personalized learning, eye tracking, and deep learning were also prominent during that period. As AI capabilities continue to advance, AI systems are poised to become commonplace tools in activities such as article or essay writing, paper outlining, artistic creation, and collaboration on academic research projects. With the advent of state-of-the-art machine learning, particularly large language models (Huang et al., 2022; Zhou et al., 2022), AI agents may assume a significant role in these activities. These advancements prompt reconsideration of longstanding assumptions about learning, posing questions about granting degrees to individuals using AI systems, hiring graduates based on their knowledge and skills, and potential concerns about fairness and cheating.

In addition to potential harms, educators contemplating the incorporation of AI in education must navigate the overhyped potentials and pitfalls—an aspect referred to as "AI theatre" by Selwyn (2022), perhaps fueled in part by "enchanted determinism" (Campolo & Crawford, 2020). This notion implies the belief that technology possesses magical, superhuman powers to address educational shortcomings or rescue humanity from its own challenges. Our narrative seeks to avoid extremes of wishful thinking or alarmist rhetoric by recognizing that AI and its foundation in global big data enhance and amplify human thinking and performance, magnifying human potential.

As a magnifying tool, AI can either exacerbate negative aspects or enhance positive ones, necessitating careful and deliberate usage. All tools capable of enhancing human potential mediate the intentions of those employing them. For instance, steam engines and later, oil and gas burning engines, reduced labor costs for massive earth-moving projects, facilitating the construction of vast dams and cities. Nevertheless, they also brought about displacements, injuries, and unemployment. AI in education constitutes a double-edged sword with the potential for unintended consequences, prompting a reassessment of assumptions regarding learning, knowledge, skill, performance, creativity, and innovation. Essential to its positive use is the intention of those wielding power, exercised with caution and vigilance regarding consequential impacts (Gibson et al., 2023).

EFL learners enhance their language proficiency through interactions with native speakers and under mentorship (Adilbayeva et al., 2022). Digital communication platforms immerse learners in sociocultural environments, with the language used on these platforms significantly affecting language improvement. The depth of word meanings is gleaned through communication (Adilbayeva et al., 2022; Ahmed, 2017; Alibakhshi & Mohammadi, 2016). Peer interactions are equally crucial, guided by instruction as emphasized by the theory (Scott & Palincsar, 2013).

Willingness to Communicate

The concept of willingness to communicate (WTC) refers to individuals' willingness to engage in verbal interactions with specific individuals or groups using a second language (L2; MacIntyre et al., 1998). It can also be interpreted as a consistent inclination for discourse when given the freedom of choice (MacIntyre et al., 1998). Kruk (2019) extended this idea by suggesting that WTC reflects a learner's cognitive consideration in employing the target language for communicative purposes. In this context, MacIntyre and Vincze (2017) argued that WTC is the primary objective in foreign language acquisition, given its potential to encourage authentic communicative behavior and enhance proficiency in the L2. The comprehensive framework of WTC, as outlined by Öz, Demirezen, and Pourfeiz (2015), encompasses affective, sociopsychological, linguistic, and communicative aspects. This framework elucidates and predicts language learners' communicative tendencies within the L2 domain. MacIntyre et al.'s (1998) theoretical framework presented a threefold analysis of WTC, examining it through trait-oriented, dynamic, and contextual lenses.

The psychological aspect of WTC is closely linked with foreign language anxiety, self-confidence, and motivation (Dewaele & Dewaele, 2018). Conversely, the dynamic and contextual dimensions of WTC are intertwined with the socioenvironmental and situational elements of the learning process, including factors such as conversational partners (Lee & Hsieh, 2019) and collaborative peers (Zarei et al., 2019). Recent literature has emphasized that WTC is best understood as a dual-faceted construct, combining the learner's enduring traits and situational dispositions (Khajavy et al., 2019). This dual perspective underscores WTC's origin from stable learner traits, such as age, gender, and personality (MacIntyre & Charos, 1996), while acknowledging its susceptibility to fluctuation based on situational cues, including interlocutors, pedagogical methods, and thematic contexts (Zhang et al., 2018). Due to its close connection with learners' inclination to actively seek communicative opportunities and participate in interactive exchanges, WTC plays a crucial role in language acquisition.

A prevailing proposition in the L2 domain asserts WTC as a decisive determinant of L2 communicative behavior, thereby contributing to L2 proficiency (MacIntyre et al., 1998). Several studies have investigated WTC's ability to predict L2 communicative patterns, revealing a positive correlation between increased WTC and enhanced L2 engagement. Additionally, inquiries have examined the relationship between WTC and L2 competence, uncovering a constructive association between the two. More recently, scholarly investigations have highlighted that L2 performance depends on learners' WTC, transcending mere communicative behaviors.

Student Engagement

Emotional engagement refers to students' affirmative and adverse reactions towards peers, educators, educational institutions, and learning outcomes. Conversely, cognitive engagement is characterized by students' intellectual investment in and comprehension of subject matter, encompassing meticulous contemplation and a willingness to invest substantial effort in comprehending intricate concepts and mastering arduous skills (Fredricks & McColskey, 2012). The ramifications of academic engagement are manifold and enduring, encompassing endeavors such as pursuing advanced education, sustaining consistent learning habits, enhancing vocational opportunities, nurturing constructive self-conception and well-being, and mitigating symptoms of depression (Eccles & Wang, 2012). Consequently, dynamic involvement in academic pursuits engenders positive outcomes that transcend the confines of educational contexts. Furthermore, intellectual engagement evinces a robust nexus with academic motivation and performance, as students who actively participate in scholarly endeavors are inclined to accord higher evaluations to their studies, attain elevated scores, and evince diminished levels of academic disengagement and evasion (Li & Lerner, 2011).

Recently, engagement has garnered substantive consideration as a pivotal determinant of academic triumph (King, 2015). It is posited that positive emotions indirectly influence educational outcomes through motivational mechanisms, prominently exemplified by engagement (Gobert et al., 2015). In this paradigm, engagement is a pivotal driver of academic aspirations. Students who manifest keen interest are apt to channel augmented exertions toward academic tasks, culminating in successful task completion and increased academic performance (Ketonen et al., 2019). In professional milieus, engagement is understood as a mental state characterized by heightened vigor, unwavering dedication, and complete engrossment (Schaufeli et al., 2002). Vigor underscores heightened cognitive strength during work; dedication encapsulates a sense of self-value, enthusiasm, inspiration, pride, and challenge, while engrossment entails complete absorption and gratification in one's undertakings, leading to a swift passage of time. This conceptual framework has been transposed into the academic realm, focusing on students' academic tasks and activities (Appleton et al., 2006). Engaged students experience heightened vitality, a fervent attachment to their academic pursuits, and an active integration into their scholarly journey (Avcı & Ergün, 2022). Empirical substantiation buttresses the proposition that engaged university students exhibit enhanced academic performance (Zhou et al., 2010), with practical designs unveiling a positive correlation between engagement and educational attainment (Avcı & Ergün, 2022). Engagement correlates with elevated academic grades, scholastic accomplishment, and self-reported learning achievements (Zhou et al., 2010). Succinctly, engagement emerges as a pivotal catalyst for academic success, wherein affirmative emotional states catalyze augmented engagement, ultimately leading to enhanced academic performance. Engaged students are predisposed to channel escalated effort into their educational undertakings, thus fostering triumphant task execution and elevated scholastic accomplishment. Therefore, educators are urged to cultivate academic engagement by developing a favorable pedagogical milieu, nurturing positive affective states, and fostering active participation in academic pursuits.

Studies on the Use of Duolingo

Studies have examined how Duolingo affects students' academic performance (Anugerahwati, 2016; Crompton, 2023; Poureau & Wright, 2013; Ratzlaff, 2015). These studies investigated multiple facets, including postsecondary students' accomplishments, attitudes, attributions, and language learning abilities (Bain et al., 2010); distinct case studies on the use of Duolingo as a tool for second language acquisition (Anugerahwati, 2023); and the advantages and difficulties of using Duolingo and other mobile learning resources (Crompton, 2013). Furthermore, studies on Duolingo's effectiveness compared to conventional language courses (Ratzlaff, 2015) and its integration into language education have been done (Munday, 2016). Additional relevant studies have looked at how gifted students perceive learning English as a foreign language (Okan & Işpınar, 2009), how Malaysian talented students struggle with language learning (Yunus, Sulaiman, Kamarulzaman, et al., 2013), and Malaysian gifted students' application of English language learning techniques (Yunus, Sulaiman, & Embi, 2013) The studies above enhance comprehension of how Duolingo, as a language learning aid, influences the accomplishments and encounters of students in various educational settings and learner types.

The previous related studies show that the use of Duolingo in EFL learning has greatly led to students' higher achievement in learning English (Alfuhaid, 2021; Arumsari & Octaviani, 2022; Habibie, 2020; Hakimantieq et al., 2022; Hernadijaya, 2020; Redjeki & Muhajir, 2021; Ünal & Güngör, 2021; Zheng & Fisher, 2023). According to the findings of earlier research, Duolingo is acknowledged for its numerous advantages; however, it is imperative to recognize the challenges it presents. Notably, Nushi and Eqbali (2017) asserted that one such challenge lies in the absence of direct human interaction within the Duolingo platform, emphasizing its primary focus on individual user development. Consequently, the enhancement of communication skills may not be optimally facilitated.

Building on this perspective, Perez (2020) underscored that the use of Duolingo necessitates a reliable Internet connection and compatible devices. However, a significant limitation arises from the unequal access to such technological facilities among students. This potential discrepancy has the capacity to give rise to challenges in the educational context (Irzawati, 2023).

The extant research collectively underscores both the merits and demerits associated with incorporating Duolingo into EFL instruction. To harness the benefits and mitigate potential drawbacks, users are advised to engage in proactive planning, address potential issues, and capitalize on available opportunities. A crucial element for users to wield Duolingo effectively in contributing to the success of language learning and teaching endeavors involves a comprehensive understanding of the platform's advantages and limitations (Irzawati, 2023).

Methodology

Sample and Procedure

A quasi-experimental research approach was used to investigate the effects of AI-infused technology, Duolingo, on the communication readiness and involvement of EFL learners. Specifically, a pretest/posttest control/experimental group research design was used. The quantitative data obtained from this design were analyzed using statistical methods.

Eighty first-year language learners from the Foreign Language Department of Hunan International Economics University in China participated in this study. The corresponding author was a researcher as well as a teacher there. Each participant was a native Chinese speaker pursuing a foreign language education in English. Chinese language proficiency and freshman enrollment at Hunan International Economics University were prerequisites for selection. Based on their capacity to give thorough and in-depth accounts of their experiences with the intervention, as well as their willingness to participate in semi-structured interviews, participants for the qualitative phase were selected. The sample sizes for the quantitative and qualitative phases of the study were 80 and 20, respectively, based on power analysis and study design. Using a pretest/posttest control/experimental group research design, the quantitative phase employed a sample size calculation to ensure adequacy.

Instrumentation

The present investigation employed two instruments to gather data: a learner engagement scale and the WTC scale. The interview checklist functioned as a qualitative measure, and the WTC and Foreign Language Classroom Anxiety (FLCA) scales as quantitative measures. We used an assessment tool adapted from MacIntyre et al. (1998) to measure the WTC of English language learners. Participants indicated their degree of agreement or willingness on a 5-point scale for each of the 27 items in this tool. Responses ranged from 1 (*almost never willing*) to 5 (*almost always willing*). The Cronbach's alpha coefficient was employed to evaluate items' internal consistency. With Cronbach's alpha for each dimension exceeding 0.78, the findings demonstrated a high degree of internal consistency and the scale's consistency in measuring the intended construct.

The Student Engagement Scale, a self-report instrument used to assess students' participation in class activities, was the second tool used. This scale was developed and validated by Fredricks and McColskey (2012). Three aspects of engagement were assessed using this scale: behavioral, cognitive, and affective enjoyment. Higher scores on this scale indicated higher levels of engagement. The range of scores was 12 to 60. The internal consistency of the scale was estimated using Cronbach's alpha, and the estimated alpha for each component exceeded 0.80, indicating that the instrument enjoyed an acceptable level of internal consistency.

Procedure

We employed a quasi-experimental research method to investigate the effectiveness of an intervention (use of Duolingo) designed to increase language learners' willingness to communicate in English and engagement in learning processes. First, the study was conducted through a structured series of steps, ensuring a systematic approach. Second, learners were randomly assigned to either the control or experimental group after screening and obtaining consent. Third, before any intervention, participants' baseline WTC and engagement levels were measured through validated pretest instruments. These instruments assessed participants' comfort and readiness to communicate in English and their motivation and involvement in learning activities. Fourth, following placement tests, the participants in the experimental group were instructed to engage with Duolingo for a minimum of 30 minutes every day over two months. Additionally, they were tasked with providing daily voice messages detailing their perceptions of progress after each session. The participants conscientiously shared with us screenshots showing their daily advancements. Notably, there needed to be a stipulation for the participant to complete a specific set of lessons or skills; instead, they had the autonomy to decide when they had sufficiently grasped the necessary knowledge. The approach encouraged participants to study at their current proficiency level, progress at their pace, and establish their learning strategies, whether these involved moving forward, revisiting topics, or maintaining a consistent practice. Finally, after the intervention, participants' WTC and engagement levels were once again measured, this time through validated posttest instruments.

Data Analysis

The collected information, which included the experimental and control groups' pretest and posttest results, was subjected to a number of statistical analyses. To provide a summary of participants' initial levels of engagement and communication as well as any changes after the intervention, descriptive statistics such as means, standard deviations, and frequency distributions were first computed. The significance of the observed differences was then evaluated using inferential statistical techniques. To find out if the changes seen in the experimental group were significantly different from those in the control group, several independent samples *t*-tests were run.

Results

Research Question 1

The first research question was designed to investigate the effects of Duolingo on language learners' engagement in classroom activities. The groups' scores on learner engagement and its components before and after the treatment were submitted to independent samples *t*-tests. The results of *t*-tests on learner engagement before the treatment are presented in Table 1.

Table 1

Results of t-*Tests Examining Learner Engagement: Pretest*

Type of engagement	Control		Experimental		Independent samples ***t***-tests		
	M	*SD*	*M*	*SD*	*t*	*df*	*p*
Affective	3.30	1.1	3.43	0.88	0.99	78	.28
Cognitive	3.20	0.97	3.27	1.07	-1.4	78	.37
Behavioral	3.35	0.85	3.20	0.88	1.45	78	.15
Total engagement	3.33	2.60	3.35	2.95	1.92	78	.29

The results indicate that with regard to affective engagement the experimental group scored slightly higher (than the control group. The groups did not differ significantly at the beginning of the study (t (78) = 0.99, p = .28). When it came to cognitive engagement, the t-test produced a non-significant result (t (78) = -1.4, p = .37), indicating that there was no discernible difference between the experimental (M = 3.27, SD = 1.07) and control group (M = 3.20, SD = 0.97) at the beginning of the study. The t-test result for behavioral engagement was not statistically significant (t (78) = 1.45, p = .15), meaning that there was no discernible difference in behavioral engagement between the groups at the start of the study. There was no statistically significant difference in the groups' scores for the variable engagement (total), indicating that the groups were similar in terms of learner engagement and its constituent parts (t (78) = 1.92, p = .29). In Table 2, the posttest results are displayed.

Table 2

Results of t-*Tests Examining Learner Engagement: Posttest*

Type of engagement	Control		Experimental		Independent samples t-tests			
	M	*SD*	*M*	*SD*	*t*	*df*	*p*	Cohen's *d*
Affective	3.43	0.90	4.20	1.00	12.3	78	.001	0.66
Cognitive	3.40	0.96	4.30	1.1	13.2	78	.001	0.95
Behavioral	3.45	0.85	4.35	1.30	14.3	78	.001	0.88
Total engagement	10.28	2.91	12.85	3.35	13.2	78	.001	2.03

As shown in Table 2, the experimental group (M = 4.20, SD = 1.00) exhibited notably higher levels of affective engagement than the control group (M = 3.43, SD = 0.90). This difference proved highly significant (t = 12.3, df = 78, p < .001), signifying a substantial increase in affective engagement attributable to the intervention. Additionally, a statistically significant disparity was observed in cognitive engagement between the experimental group (M = 4.30, SD = 1.1) and the control group (M = 3.40, SD = 0.96), as indicated by the t-test (t = 13.2, df = 78, p = .001). Thirdly, a remarkable augmentation in behavioral engagement was evident in the experimental group (M = 4.35, SD = 1.30) compared to the control group (M = 3.45, SD = 0.85). The t-test result was highly significant (t = 14.3, df = 78, p < .001), underscoring the intervention's substantial impact on behavioral engagement. Finally, the total learner engagement score for the experimental group (M = 12.85, SD = 3.35) surpassed that of the control group (M = 10.28, SD = 2.91). The t-test statistic was highly significant (t = 13.2, df = 78, p < .001), confirming a substantial positive effect of the intervention on overall learner engagement.

We compared effect sizes for each aspect to investigate the consistency of outcomes regarding learner engagement and its components through implementing digital communication activities. The results revealed that digital communication activities moderately affected EFL learners' affective engagement (Cohen's d = 0.66); however, there was a significant effect on other aspects of engagement (behavioral and cognitive) as well as the overall engagement (Cohen's d > 0.80).

Research Question Two

The impact of Duolingo on language learners' readiness to interact in the classroom was measured in response to the second research question. Independent samples t-tests were performed using the groups' WTC scores, both before and after the treatment, and its components. Table 3 displays the findings of t-tests for the groups' WTC scores prior to the intervention.

Table 3

Results of t*-Tests Examining Willingness to Communicate: Pretest*

WTC variable	Control		Experimental		Independent samples t-tests		
	M	*SD*	*M*	*SD*	t	*df*	*p*
Speaking in class	25.1	3.10	24.00	2.40	0.82	78	.15
Reading in class	17.3	4.20	18.00	3.50	0.93	78	.27
Writing in class	22.3	2.00	23.00	2.00	0.80	78	.21
Comprehension	13.2	4.10	13.00	2.10	1.40	78	.19
Total WTC	77.9	13.30	78.00	13.20	2.20	78	.41

Note. WTC = willingness to communicate.

When it came to reading in class, the experimental group scored an average of 18 points (SD = 3.5), while the control group scored an average of 17.3 points (SD = 4.20). The t-test produced a non-significant result (t (78) = 0.93, p = .27), suggesting that there was no statistically significant difference in the control and experimental groups' willingness to read at the start of the study. In terms of writing assignments for class, the experimental group showed a somewhat higher mean score of 23 (SD = 2.00) than the control group, which showed a mean score of 22.3 (SD = 2.00). The results of the t-test showed a non-significant difference (t (78) = 0.80, p = .21), indicating that at the beginning of the study, there was no significant difference between the two groups' willingness to write in class. The experimental group produced a mean score of 13 (SD = 2.10) for comprehension, while the control group's mean score was 13.2 (SD = 4.10). At the beginning of the study, there was no discernible difference in comprehension between the control and experimental groups, according to the non-significant result of the t-test (t (78) = 1.40, p = .17). The experimental group recorded a slightly lower mean score of 84.9 (SD = 13.20) than the control group, which recorded a mean score of 88.00 (SD = 13.30) regarding the WTC (total). The lack of a significant difference at the start of the study in the t-test result (t (78) = 2.1, p = .31) indicates that the result did not reach statistical significance.

Results of the independent samples *t*-tests to compare posttest scores between the groups for various aspects of their WTC is shown in Table 4.

Table 4

Results of t-*Tests Examining Willingness to Communicate: Posttest*

WTC Variable	Control		Experimental		Independent samples *t*-tests			
	M	*SD*	*M*	*SD*	*t*	*df*	*p*	Cohen's *d*
Speaking in class	23.60	3.33	32.3	3.66	12.39	78	.001	1.9
Reading in class	18.60	4.13	24.39	4.16	13.29	78	.001	1.4
Writing in class	24.50	2.3	27.00	2.4	14.3	78	.001	1.13
Comprehension	12.70	4.20	18.53	2.35	10.6	78	.001	1.69
Total WTC	79.40	13.9	107.8	12.4	14.61	78	.001	2.20

Note. WTC = willingness to communicate.

The comparison of the experimental group (*M* = 32.3, *SD* = 3.66) and control group (*M* = 23.6, *SD* = 3.33) regarding willingness to speak showed a statistically significant difference, as shown in Table 4. When comparing the experimental group to the control group, the *t*-test result (*t* (78) = 12.39, *p* = .001) showed a significant increase in the effect size (Cohen's *d* = 1.9). Moreover, a significant distinction was observed in the experimental group (*M* = 24.39, *SD* = 4.16) and control group (*M* = 18.60, *SD* = 4.13) with regard to willingness to read. There was a significant difference in the effect size (Cohen's *d* = 1.40) between the experimental and control groups, as indicated by the *t*-test (*t* (78) = 13.29, *p* = .001). The experimental group (*M* = 27, *SD* = 2.4) showed statistically significant improvement in willingness to write compared to the control group (*M* = 24.5, *SD* = 2.3). The experimental group appears to have had significantly higher writing engagement, as suggested by the substantial effect size (Cohen's *d* = 1.13) indicated by the *t*-test (*t* (78) = 14.3, *p* = .001). Also, a noteworthy distinction can be seen in the experimental group's (*M* = 18.53, *SD* = 2.35) and control group's (*M* = 12.7, *SD* = 4.20) performance on the WTC comprehension scale. There was a significant difference in the effect size (Cohen's *d* = 1.69) between the experimental and control groups, as demonstrated by the *t*-test (*t* (78) = 10.6, *p* = .001).

Discussion

Investigating how Duolingo affected language learners' participation in classroom activities—with a particular emphasis on affective, cognitive, and behavioral engagement—provided insightful information. There were no statistically significant differences found between the control and experimental groups in Table 1's pretest score analysis across different engagement dimensions, suggesting homogeneity at the start of the study. It appears that any subsequent differences can be attributed to the effects of Duolingo, as there were no significant differences in affective, cognitive, behavioral, or total engagement prior to the intervention. When looking at the posttest scores in Table 2, it is evident that the experimental group had significantly higher engagement levels overall than the control group. As assessed by students' emotional involvement, affective engagement showed a significant improvement (Cohen's *d* = 0.66), suggesting that Duolingo had a beneficial effect on students' emotional attachment to the language learning process. This result is consistent with studies (Blake, 2013; Crompton, 2013) that highlight the contribution of technology, like apps such as Duolingo, to the development of favorable affective outcomes in language learning.

Similarly, cognitive engagement, reflecting learners' mental investment in language learning activities, substantially improved in the experimental group (Cohen's d = 0.95). This result is consistent with studies emphasizing the cognitive benefits of technology in language education (Carneiro & Simao, 2011; Klopfer et al., 2002). The positive impact on cognitive engagement supports the notion that Duolingo's interactive features and gamified elements contribute to heightened cognitive involvement (Blake, 2016). Behavioral engagement, representing learners' active participation in language learning tasks, showed a remarkable increase in the experimental group (Cohen's d = 0.88). The statistically significant improvement suggests that Duolingo effectively promotes learner involvement and participation in language-related activities (Thornton & Houser, 2005; Vesselinov & Grego, 2012).

The total learner engagement score also revealed a substantial positive effect of Duolingo (Cohen's d = 2.03). This comprehensive measure suggests that integrating Duolingo into language instruction significantly enhances engagement. The effect size's magnitude underscores the intervention's practical significance (Fredricks & McColskey, 2012). In summary, the results indicate that Duolingo positively influences language learners' engagement in classroom activities, encompassing affective, cognitive, behavioral, and total engagement. These findings align with previous research highlighting the potential of technology, specifically Duolingo, in enhancing various dimensions of engagement in language learning contexts.

Next, we look at how Duolingo affects language learners' WTC in the classroom. The findings (tables 3 and 4) offer important new information about how the intervention affects various facets of communication. There were no discernible variations in WTC between the control and experimental groups in speaking, reading, writing, comprehension, or overall willingness to communicate, according to the pretest analysis shown in Table 3. This initial homogeneity implies that the intervention can be held responsible for any changes in learners' willingness to communicate in future. When comparing the experimental group to the control group, there was a noticeable increase in communication willingness, as indicated by the analysis of the posttest scores shown in Table 4. The experimental group was substantially more willing to speak in class (Cohen's d = 1.9), highlighting Duolingo's beneficial effects on oral communication. This result is in line with studies showing how well technology can improve speaking abilities (Blake, 2013; Klopfer et al., 2002). The experimental group showed a significantly higher willingness to read aloud in class (Cohen's d = 1.4). This finding is consistent with research showing how technology, particularly language learning applications, can increase reading engagement (Lee & Hsieh, 2019; Stanley, 2013). The experimental group showed noticeably higher willingness to write in class.

The experimental group's total WTC score was substantially higher than the control group's (Cohen's d = 2.2). According to this thorough assessment, Duolingo significantly improves students' willingness to interact with one another in the classroom. Research by Zhang, Beckmann, and Beckmann (2018) adds credence to these conclusions by indicating that digital communication activities significantly increase learners' willingness to communicate. The study highlights how technology can help language learners create a supportive environment that increases their confidence and drive to share. Furthermore, Mystkowska-Wiertelak and Pawlak (2017) emphasized the connection between affective, cognitive, and contextual factors that affect communication willingness. The results of this study support this viewpoint since Duolingo's influence on learners' affective and cognitive engagement in the context of language learning increases their willingness to communicate. The study's findings prove that Duolingo positively impacts language learners' participation in class activities and openness to communication. The results are consistent with other studies showing how technology—especially Duolingo—can improve communication and engagement in language learning environments. The significant effect sizes are seen in several dimensions.

Conclusion and Implications

The study's findings shed light on how Duolingo has a revolutionary effect on language learners' willingness to communicate and participate in class activities. Significant gains in affective, cognitive, and behavioral engagement, as well as an increased willingness to speak, read, write, comprehend, and communicate in general, are consistently shown in the results. The results highlight Duolingo's efficacy as a flexible instrument for language learning and are consistent with recent studies that support the use of technology in language learning. The significant effect sizes observed in various engagement and communication dimensions underscore Duolingo's capacity to incite favorable transformations in language learners' perspectives, involvement, and communicative assurance. Duolingo has surfaced as a helpful ally in creating a rich and dynamic language learning environment as educators look for new ways to involve students and support efficient language acquisition. The study's implications go beyond the immediate context to encompass more general language education issues. Above all, the favorable results highlight how important it is for teachers to use technology wisely. Duolingo is a prime example of how digital platforms can improve student engagement and communication skills. Including Duolingo in language courses could benefit teachers who want to create a vibrant and welcoming learning environment.

Acknowledgements

This work was sponsored in part by Key Scientific Research Project of Hunan Provincial Department of Education (23A0655) and Teaching Reform Research Project of HEIs in Year 2023 in Hunan Province (HNJG-20231611).

This research was funded by the Foreign Language Research Joint Project of Hunan Social Science Foundation (No. 23WLH32).

References

Adilbayeva, U., Mussanova, G. A., Mombekova, N. B., & Suttibayev, N. A. (2022). Digital communication technology for teaching a foreign language and culture through reading. *International Journal of Society, Culture & Language, 10*(3), 21–30. https://www.ijscl.com/article_252481_8ed1f0934dafd81701135be0fca47728.pdf

Ahmed, S. (2017). *Impact and implementation of sociocultural theory in ESL classroom* [Unpublished doctoral dissertation]. BRAC University.

Alfuhaid, S. R. (2021). The utilization of Duolingo to enhance the speaking proficiency of EFL secondary school students in Saudi Arabia. *English Language Teaching, 14*(11), 9–15. https://doi.org/10.5539/elt.v14n11p9

Alibakhshi, G., & Mohammadi, M. J. (2016). Synchronous and asynchronous multimedia and Iranian EFL learners' learning of collocations. *Applied Research on English Language*, *5*(2), 237–254. https://www.doi.org/10.22108/ARE.2016.20428

Anugerahwati, M. (2023). Can Duolingo help people master a foreign language? *KnE Social Sciences*, *8*(7), 133–139. https://www.doi.org/10.18502/kss.v8i7.13244

Arumsari, & Octaviani, S. K. (2022). Duolingo mobile application for English listening skill improvement of vocational school students. *Journal of English Language and Education, 7*(2), 84–95. https://doi.org/10.31004/jele.v7i2.306

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology*, *44*(5), 427-445. https://doi.org/10.1016/j.jsp.2006.04.002

Avcı, Ü., & Ergün, E. (2022). Online students' LMS activities and their effect on engagement, information literacy, and academic performance. *Interactive Learning Environments, 30*(1), 71–84. https://www.doi.org/10.1080/10494820.2019.1636088

Bain, S. K., McCallum, R. S., Bell, S. M., Cochran, J. L., & Sawyer, S. C. (2010). Foreign language learning aptitudes, attitudes, attributions, and achievement of postsecondary students identified as gifted. *Journal of Advanced Academics, 22*(1), 130–156. https://doi.org/10.1177/1932202X1002200106

Blake, R. (2013). *Brave new digital classroom: Technology and foreign language learning* (2nd ed.). Georgetown University Press.

Blake, R. (2016). Technology and the four skills. *Language Learning & Technology, 20*(2), 129–142. http://hdl.handle.net/10125/44465

Campolo, A., & Crawford, K. (2020). Enchanted determinism: Power without responsibility in artificial intelligence. *Engaging Science, Technology, and Society, 6*(2020), 1–19. https://doi.org/10.17351/ests2020.277

Carneiro, R., & Simao, A. M. V. (2011). Technology-enhanced environments for self-regulated learning in teaching practices. In R. Carneiro, P. Lefrere, K. Steffens, & J. Underwood (Eds.), *Self-regulated learning in technology enhanced learning environments: A European perspective* (pp. 75–101). Sense Publishers.

Crompton, H. (2013). The benefits and challenges of mobile learning. *Learning and Leading with Technology, 41*, 38–39.

Dawson, S., Poquet, O., Colvin, C., Rogers, T., Pardo, A., & Gasevic, D. (2018). Rethinking learning analytics adoption through complexity leadership theory. In A. Pardo, K. Bartimote-Aufflick, & G. Lynch (Chairs), *LAK '18: Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 236–244). ACM. https://doi.org/10.1145/3170358.3170375

Dewaele, J. M., & Dewaele, L. (2018). Learner-internal and learner-external predictors of willingness to communicate in the F.L. classroom. *Journal of the European Second Language Association, 2*(1), 24–37. https://www.doi.org/10.22599/jesla.37

Eccles, J. & Wang, M. T. (2012) So, what is student engagement anyway? In S.L. Christenson, A.L. Reschly, & C. Wylie (Eds.). *Handbook of research on student engagement* (pp.133-145). Springer. https://doi.org/10.1007/978-1-4614-2018-7_6

Feng, S., & Law, N. (2021). Mapping artificial intelligence in education research: A network-based keyword analysis. *International Journal of Artificial Intelligence in Education, 31*(2), 277–303. https://doi.org/10.1007/s40593-021-00244-4

Gibson, D., Kovanovic, V., Ifenthaler, D., Dexter, S., & Feng, S. (2023). Learning theories for artificial intelligence promoting learning processes. *British Journal of Educational Technology, 54*(5), 1125–1146. https://doi.org/10.1111/bjet.13341

Gobert, J. D., Baker, R. S., & Wixon, M. B. (2015). Operationalizing and detecting disengagement within online science microworlds. *Educational Psychologist, 50*(1), 43–57. https://doi.org/10.1080/00461520.2014.999919

Habibie, A. (2020). Duolingo as an educational language tool to enhance EFL students' motivation in learning. *British (Jurnal Bahasa dan Sastra Inggris), 9*(1), 13–26. https://dx.doi.org/10.31314/british.9.1.13-26.2020

Hakimantieq, H., Suherdi, D., & Gunawan, W. (2022). Duolingo as a mobile-assisted language learning: A new supplementary of learning basic English reading for EFL students. *Edukatif: Jurnal Ilmu Pendidikan, 4*(6), 7548–7558. https://doi.org/10.31004/edukatif.v4i6.4104

Hernadijaya, N. S. (2020). The use of Duolingo application to enhance junior high school students' English vocabulary. *RETAIN, 8*(2), 17–24. https://ejournal.unesa.ac.id/index.php/retain/article/view/33055

Huang, J., Gu, S. S., Hou, L., Wu, Y., Wang, X., Yu, H., & Han, J. (2022). *Large language models can self-improve*. arXiv. https://doi.org/10.48550/arXiv.2210.11610

Irzawati, I. (2023). The integration of Duolingo into EFL learning. *Esteem Journal of English Education Study Programme*, *6*(2), 328-337. DOI:10.31851/esteem.v6i2.12317

Keegan, D. (2002). *The future of learning: From e-learning to m-learning*. Zentrales Institut für Fernstudienforschung, Fern Universität. https://files.eric.ed.gov/fulltext/ED472435.pdf

Ketonen, E. E., Malmberg, L.-E., Salmela-Aro, K., Muukkonen, H., Tuominen, H., & Lonka, K. (2019). The role of study engagement in university students' daily experiences: A multilevel test of moderation. *Learning and Individual Differences, 69,* 196–205. https://doi.org/10.1016/j.lindif.2018.11.001

Khajavy, G. H., Makiabadi, H., & Navokhi, S. A. (2019). The role of psychological capital in language learners' willingness to communicate, motivation, and achievement. *Eurasian Journal of Applied Linguisitcs, 5*(3), 495–513. https://doi.org/10.32601/ejal.651346

King, R. B. (2015). Sense of relatedness boosts engagement, achievement, and well-being: A latent growth model study. *Contemporary Educational Psychology, 42,* 26–38. https://doi.org/10.1016/j.cedpsych.2015.04.002

Klopfer, E., Squire, K., & Jenkins, H. (2002). Environmental detectives: PDAs as a window into a virtual simulated world. In B. Werner (Ed.), *Proceedings: IEEE International Workshop on Wireless and Mobile Technologies in Education* (pp. 95–98). IEEE Computer Society. https://www.doi.org/10.1109/WMTE.2002.1039227

Kruk, M. (2022). Dynamicity of perceived willingness to communicate, motivation, boredom and anxiety in second life: The case of two advanced learners of English. *Computer Assisted Language Learning, 35*(1–2), 190–216. https://www.doi.org/10.1080/09588221.2019.1677722

Larsen-Freeman, D., & Anderson, M. (2011). *Techniques and principles in language teaching* (3rd ed.). Oxford University Press.

Lee, J. S., & Hsieh, J. C. (2019). Affective variables and willingness to communicate of EFL learners in in-class, out-of-class, and digital contexts. *System, 82*, 63–73. https://www.doi.org/10.1016/j.system.2019.03.002

Li, Y., & Lerner, R. M. (2011). Trajectories of school engagement during adolescence: Implications for grades, depression, delinquency, and substance use. *Developmental Psychology, 47*(1), 233–247. https://www.doi.org/10.1037/a0021307

MacIntyre, P. D., & Charos, C. (1996). Personality, attitudes, and affect as predictors of second language communication. *Journal of Language and Social Psychology, 15*(1), 3–26. http://dx.doi.org/10.1177/0261927X960151001

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in an L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545–562. https://www.doi.org/10.1111/j.1540-4781.1998.tb05543.x

MacIntyre, P. D., & Vincze, L. (2017). Positive and negative emotions underlie motivation for L2 learning. *Studies in Second Language Learning and Teaching, 7*(1), 61–88. https://www.doi.org/10.14746/ssllt.2017.7.1.4

Mystkowska-Wiertelak, A., & Pawlak, M. (2017). *Willingness to communicate in instructed second language acquisition: Combining a macro- and micro-perspective*. Multilingual Matters. https://doi.org/10.21832/9781783097173

Nushi, M., & Eqbali, M. H. (2017). Duolingo: A mobile application to assist second language learning. *The Journal of Teaching English with Technology, 17*(1), 89–98. https://www.ceeol.com/search/article-detail?id=478437

Ogata, H., & Yano, Y. (2004). Context-aware support for computer-supported ubiquitous learning. In B. Werner (Ed.), *Proceedings: The 2nd IEEE International Workshop on Wireless and Mobile Technologies in Education* (pp. 27–34). IEEE Computer Society. https://www.doi.org/10.1109/WMTE.2004.1281330

Öz, H., Demirezen, M., & Pourfeiz, J. (2015). Willingness to communicate with EFL learners in the Turkish context. *Learning and Individual Differences, 37*, 269–275. https://www.doi.org/10.1016/j.lindif.2014.12.009

Perez, L. C. D. (2020). English language acquisition via Duolingo application: Effectivity and challenges. *Lingua Pedagogia: Journal of English Teaching Studies, 2*(2), 76–87. C:\Users\shenderson\Copy\articles\4 https://journal.uny.ac.id/index.php/lingua-pedagogia/article/view/37984/0

Ratzlaff, N. J. (2015). A *cross-comparison and efficacy study of Duolingo and an entry-level German 1A course* [Unpublished doctoral dissertation]. California State University.

Redjeki, I. S., & Muhajir, R. (2021). Gamification in EFL classroom to support teaching and learning in 21st century. *JEES (Journal of English Educators Society)*, *6*(1), 68–78. https://doi.org/10.21070/jees.v6i1.882

Poureau, L., & Wright, J. (2013). Owning It: An evaluation of language applications and software for second language acquisition mastery. *Online Submission*. https://eric.ed.gov/?id=ED542662

Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic approach. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being, 3*(1), 71–92. https://doi.org/10.1023/A:1015630930326

Scott, S., & Palincsar, A. (2013). *Sociocultural theory*. IXL Learning. http://dr-hatfield.com/theorists/resources/sociocultural_theory.pdf

Selwyn, N. (2022). Digital degrowth: Toward radically sustainable education technology. *Learning, Media and Technology*, 1-14. https://doi.org/10.1080/17439884.2022.2159978

Stanley, G. (2013). *Language learning with technology: Ideas for integrating technology in the classroom*. Cambridge University Press.

Thornton, P., & Houser, C. (2005). Using mobile phones in English education in Japan. *Journal of Computer Assisted Learning, 21*(3), 217–228. http://dx.doi.org/10.1111/j.1365-2729.2005.00129.x

Tsai, Y.-S., Poquet, O., Gašević, D., Dawson, S., & Pardo, A. (2019). Complexity leadership in learning analytics: Drivers, challenges, and opportunities. *British Journal of Educational Technology, 50*(6), 2839–2854. https://doi.org/10.1111/bjet.12846

Ünal, E., & Güngör, F. (2021). The continuance intention of users toward mobile assisted language learning: The case of DuoLingo. *Asian Journal of Distance Education, 16*(2), 197–218. https://doi.org/10.5281/zenodo.5811777

Vesselinov, R., & Grego, J. (2012). *Duolingo effectiveness study: Final report*. City University of New York. https://www.languagezen.com/pt/about/english/Duolingo_Efficacy_Study.pdf

Wagner, E. (2005). Enabling mobile learning. *EDUCAUSE Review, 40*(3), 41–52. https://er.educause.edu/articles/2005/5/enabling-mobile-learning

Walker, A., & White, G. (2013) Technology enhanced language learning. Oxford University Press. https://elt.oup.com/catalogue/items/global/teacher_development/oxford_handbooks_for_language

Yang, S. J. H. (2006). Context-aware ubiquitous learning environments for peer-to-peer collaborative learning. *Educational Technology and Society, 9*(1), 188–201. https://www.researchgate.net/publication/220374333

Yunus, M. M., Lubis, M. A., Lin, C. P., & Wekke, I. S. (2009). Language learning via ICT: Students' experience. In Proceedings of the 5*th WSEAS/IASME International Conference on Educational Technologies (EDUTE'09*; pp. 136-142). https://www.academia.edu/1562306/LANGUAGE_LEARNING_VIA_ICT_STUDENTS_EXPERIENCE

Yunus, M. M., Sulaiman, N. A., & Embi, M. A. (2013). Malaysian gifted students' use of English language learning strategies. *English Language Teaching, 6*(4), 97–109. https://www.doi.org/10.5539/elt.v6n4p97

Yunus, M. M., Sulaiman, A., Kamarulzaman, M. H., & Ishak, N. M. (2013). Language learning difficulties among Malaysian gifted students. *Asian Social Science, 9*(15), 130–137. https://www.doi.org/10.5539/ass.v9n15p130

Zarei, N., Saeidi, M., & Ahangari, S. (2019). Exploring EFL teachers' socio-affective and pedagogic strategies and students' willingness to communicate with a focus on Iranian culture. *Hindawi Education Research International*, *2019*, Article 3464163. https://www.doi.org/10.1155/2019/3464163

Zhang, J., Beckmann, N., & Beckmann, J. F. (2018). To talk or not to talk: A review of situational antecedents of willingness to communicate in the second language classroom. *System, 72*, 226–239. https://www.doi.org/10.1016/j.system.2018.01.003

Zheng, C., & Fisher, L. (2023). Opening the "black box": How out-of-class use of Duolingo impacts Chinese junior high school students' intrinsic motivation for English. *ECNU Review of Education*. Advance online publication. https://www.doi.org/10.1177/20965311231171606

Zhou, Q., Main, A., & Wang, Y. (2010). The relations of temperamental effortful control and anger/frustration to Chinese children's academic achievement and social adjustment: A longitudinal study. *Journal of Educational Psychology, 102*(1),180–196 https://www.doi.org/10.1037/a0015908

Zhou, Y., Muresanu, A. I., Han, Z., Paster, K., Pitis, S., Chan, H., & Ba, J. (2022). *Large language models are human-level prompt engineers*. ArXiv. http://arxiv.org/abs/2211.01910

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

The Auxiliary Role of Artificial Intelligence Applications in Mitigating the Linguistic, Psychological, and Educational Challenges of Teaching and Learning Chinese Language by non-Chinese Students

Jingfang Xia[1], Yao Ge[2], Zijun Shen[3], and Dr Mudasir Rahman Najar[4]
[1]College of Education Guangxi Normal University, Guilin, 541004, China; [2]IOE, UCL's Faculty of Education and Society University College London; [3]Department of Foreign Languages, Sichuan University of Media and Communications, Chengdu, China; [4]Dept of English, SRM University, Delhi NCR, Haryana, India

Abstract

Learners might have several challenges while attempting to learn a second/foreign language. Learners of Chinese face linguistic, psychological, and educational challenges. The integration of technology, especially artificial intelligence (AI), into teaching foreign languages is a blessing for teachers and learners. This study delved into the auxiliary role of AI-powered applications in mitigating the linguistic, psychological, and educational challenges which non-Chinese learners face while learning Chinese/Mandarin language. Qualitative research was employed, and 20 teachers of Chinese language were selected through theoretical sampling. In-depth interviews were used for collecting data, and MAXQDA was used for thematic analysis. Findings revealed that AI-powered educational applications are useful for helping language learners overcome the commonly reported linguistic, psychological, and educational challenges which non-Chinese learners and teachers of Mandarin might encounter. Findings verify the effectiveness of AI-powered applications, such as ChatGPT, Poe, Brainly, and so forth, in helping teachers and learners of Chinese language learn grammar, structure, idioms, and cultural issues of Chinese language. Findings have implications for foreign language (Chinese) learners and teachers, educational technologists, as well as syllabus designers.

Keywords: AI-powered application, auxiliary role of AI, AI, Chinese language, challenges of learning, Chinese language learner, Chinese language teacher

Introduction

Over numerous years, the People's Republic of China (PRC) has exerted significant influence on the global political stage. This can be attributed primarily to its extensive military power, expansive territorial control, and its privileged position as a permanent member of the United Nations Security Council (Wang & Lemmer, 2015; Wu, 2017; Wu et al.,2022). Following significant internal political changes, China has taken on a role of great importance as a trade partner and a rapidly growing market for regions including the Middle East, the Western world, and Africa (Yalun, 2019). Chinese, being one of the official languages of the United Nations, holds a central position in international discussions.

AI is an expanding area of computer science that engenders intelligent machines capable of executing human activities. The implementation of AI technology is progressively more prevalent in the domains of healthcare, banking, and transportation. The potential of AI to revolutionize diverse industries lies in its efficacy, accuracy, and decision-making capabilities. AI acquires knowledge and enhances its performance through the use of machine-learning algorithms. These algorithms enable robots to assess extensive datasets, identify patterns, and uncover insights that are beyond the reach of human cognition. Consequently, this has propelled the advancement of natural language processing, computer vision, and speech recognition. Nevertheless, the advent of AI may engender job displacement and predispose biases in decision-making processes. Therefore, the ethical dimensions of AI research and development must be taken into account. While AI has the potential to enhance society and improve lives, its ethical use is imperative.

A few related studies (Wang, 2004, Ye, 2013; Yue, 2017) have argued that there is a clear tendency among youth in several Western countries to study Mandarin as a foreign language (MFL) or Chinese as a foreign language (CFL). The term Chinese Mandarin as a foreign language (CMFL) has also gained popularity in academic discourse (Wang & Lemmer, 2015). Increasing enrollment numbers demonstrate the growing interest in CFL, the need for more Chinese language teachers, and rising support for standardized tests of Chinese language proficiency in nations with MFL/CFL teaching institutions (Lu, et al., 2016; Orton, 2011, 2016; Pérez-Milans, 2015). The need for more instructors to be trained and recruited domestically in China and internationally to better support Chinese language learners has become urgent due to the rising interest in Chinese language acquisition and the increasing number of international students studying this language. China's efforts to enhance its global influence have led to increased investments in promoting Mandarin, often referred to interchangeably as Chinese. These initiatives also aim to elevate Mandarin's official language status in regions such as Taiwan and Singapore (Hartig, 2015). It's important to highlight that Chinese is one of the six official languages at the United Nations (Yalun, 2019). In English-speaking nations, diverse national initiatives have played a role in the expansion of CFL/MFL programs in educational institutions, both within China and globally, including countries such as Iran, Japan, the United Kingdom, the United States, South Africa, and Australia. Notably, in 2015, the US government reported that one million American students were engaged in studies in China, with ambitious plans to increase this number by 2020 (Everson, 2008; Wu & Itayi, 2022).

Teaching and learning Chinese, like other second/foreign languages such as English, French, and German, is associated with a set of linguistic, psychological, and educational problems, which can be solved through educational applications powered by AI. Therefore, educators in Chinese language teaching must critically reevaluate traditional communication media and teaching methodologies through the lens of practical teaching experiences. This entails a paradigm shift in ideas and roles, the organization of disparate oral Chinese teaching resources, and the seamless integration of contemporary information technologies with traditional language teaching methods (Guo et al., 2019). Using network resources adeptly, educators should design the learning process and leverage the network to access individualized learning information for each student, thereby facilitating regulation of the learning process. Developing intelligent computer-assisted instruction systems involves the convergence of artistic intelligence, computer science, pedagogy, psychology, and behavioral science (Graham et al., 2019). This research aims to confer intelligence upon computer systems, allowing them to assume specific educational and instructional responsibilities and, consequently, partially substitute for teachers to achieve optimal teaching outcomes (Johnson et al., 2018). More specifically, it explores how AI-powered applications such as ChatGPT, Poe, and Brainly can be used to help Chinese language learners and teachers overcome linguistic, psychological, and educational problems. More specifically, the following questions were posed:

1. How can AI-powered educational applications help Chinese language learners overcome the linguistic problems of learning the Chinese language?

2. How can AI-powered educational applications help Chinese language learners overcome the psychological problems of learning Chinese?

3. How can AI-powered educational applications help Chinese language teachers overcome the educational problems of teaching Chinese?

Literature Review

One promising approach to resolving the complex issues faced by non-Chinese students learning Chinese is the integration of AI applications into language education (Guo et al., 2019). The theoretical underpinnings of this investigation are found at the nexus of linguistics, psychology, and education. The subtleties of learning Chinese are closely related to linguistics, which is a critical component. AI applications can help overcome language barriers by providing learners with a thorough and immersive linguistic experience by using advanced language processing capabilities. AI applications in the psychological field can be made to adjust to different learning styles and offer customized feedback, resulting in a positive and personalized learning environment. Finally, AI technologies have the potential to provide interactive and captivating learning environments, leading to better teaching and increased student involvement (Kang & Kang, 2023). The integration of these theoretical viewpoints offers a strong foundation for comprehending AI's auxiliary function in easing the difficulties non-Chinese learners encounter while learning the language.

The contemporary openness of network education undeniably broadens the learning space available for Chinese language instruction, with increasing interest in network learning and the continual emergence of network teaching platforms. However, the complexity and temporal-spatial separation within the network education environment poses challenges for managers and educators in acquiring dynamic learning information from online learners (Zhu et al., 2019). This often results in a simplistic reproduction of textbook resource content and a unilateral pursuit of quantity and scale in the push mode of teaching resources (Yalun, 2019). Addressing this issue becomes crucial, given the diverse group of online learners and the need to determine effective ways of collecting reliable learning status information and providing personalized learning services (Derun et al., 2019).

The rapid development of multimedia and network technologies necessitates urgently establishing a novel Chinese language model that transcends traditional constraints of region and time (Chen et al., 2014). To enhance teaching efficiency and cultivate talents effectively, continuous exploration and experimentation with new technologies and methods for improving teaching and learning approaches are essential (Li, 2018). Simultaneously, there is a desire to tailor education to students' aptitudes and implement differentiated education based on individual learning foundations, abilities, and other characteristics (Miao et al., 2018). The proposal of an intelligent teaching system (ITS) makes it feasible to achieve these goals. By establishing an open teaching environment, Internet-based modern education surpasses traditional education's temporal and spatial limitations (Agarwal et al., 2021). Effectively leveraging the resource advantages of diverse existing education systems, rational resource allocation and educational development become possible, providing a practical solution to this challenge. Similarly, Kang and Wang (2022) proposed a model for an intelligent Chinese language network teaching system to address shortcomings in existing network teaching systems.

Intelligent teaching represents a pivotal direction in computer-aided teaching, characterized by an open interactive teaching approach that uses computers to simulate the teaching thought processes of experts (Jin, 2011). This approach centers on students as the focal point, with computers serving as the medium. In modern educational theory, intelligent teaching incorporates advancements from AI, psychology, and cognitive science into computer-aided teaching (Kang & Kang, 2023). It seeks to understand the mode of learning cognition by examining the characteristics and processes of human learning thinking, enabling students to acquire knowledge through personalized adaptive learning and achieving the goal of genuine individualized teaching (Peng, et al., 2019). Research on intelligent computer-assisted instruction systems in China commenced relatively late, but it can potentially play a positive role in advancing the country's education reform (Yi et al., 2020). By analyzing the substantial drawbacks related to adaptability and personalization in current online learning systems and drawing on critical technologies such as fuzzy evaluation algorithms and neural networks (NN) within the ITS, the paper discusses the extraction of parameters such as behavioral data and performance information from online learners (Xiao & Liu, 2017).

Educators in a networked environment, particularly those involved in language teaching, are urged to reevaluate conventional communication channels and instructional methodologies through the lens of practical teaching experience (Guo et al., 2019). This necessitates a shift in perspectives and roles and the systematic organization of disparate Chinese teaching resources. Graham et al., 2019 contributed a framework and design pattern capable of efficiently constructing a college English teaching system with a well-defined structure and dependable performance. This is achieved through research on the modeling of college English teaching systems using universal modeling language (UML).

According to research, the current network teaching platforms (Zhu et al., 2019) provide a wide range of educational resources; however, they frequently focus on enhancing particular student competencies rather than encouraging all-encompassing speaking, listening, reading, and writing skills. The literature (Li, 2018) supports the seamless integration of contemporary IT with conventional language instruction techniques. Research has highlighted how crucial it is for teachers to skillfully use network resources to plan the learning process and gather personalized learning data for every student in order to successfully manage the learning journey (Makhambetova, et al., 2021)

The student model, which methodically depicts students' knowledge, cognitive abilities, learning motivations, learning styles, and prior learning patterns is crucial to ITS, as stated in Hirschmann et al., 2019. Yi et al., 2020 concluded that in order to solve urgent problems with the current Chinese language network teaching platform, it is imperative to develop a comprehensive, interactive, personalized, and feedback-oriented Chinese language teaching platform. According to Miao et al., 2018, the main goal of a network teaching platform is to showcase students' characteristics and attitudes, which serves as the basis for putting creative teaching objectives, resources, and methods into practice. Educational objectives ought to encompass motor skills, cognitive abilities, and emotions, as indicated by existing literature (Xiao & Liu, 2019).

Using UML, Kang and Kang (2023) explored the creation of network-based college. The framework and modeling diagram for an online college English teaching system are designed after extensive research. As per Johnson et al. (2018), the student model should be influenced by the responses and interactions that students have had with the system. This enables the system to apply personalized teaching, allowing for dynamic adjustments based on the learning circumstances of students.

To get around the drawbacks of current teaching systems, Xiao and Liu (2019) offered a model for a networked ITS. Their established a deep learning (DL)-based Chinese language teaching system model, building on earlier research and taking into account the state of Chinese language network teaching systems (Jiang et al., 2021). The model categorizes the learning characteristics of students and adjusts various teaching methods and materials based on these attributes. In order to achieve authentic individualized teaching, this model can dynamically create a personalized learning environment based on unique student traits.

Methodology

Sample and Procedure

For this study, we sought out a few teachers to serve as informants. The study's starting point did not specify a specific sample size. Given this, we first selected 20 teachers from various countries using a theoretical sampling technique. We conducted interviews until data saturation was achieved. After interviewing the fifteenth teacher, data saturation was reached. Thus, the final sample consisted of fifteen teachers. When choosing these instructors, special consideration was given to universities that offer Mandarin as a foreign language in the UK, Russia, and Australia. All of the instructors were non-native Mandarin speakers who started studying the language in their home countries or China after reaching puberty. After being fully informed about the purpose of the study, all participants signed the required paperwork to give their informed consent. The demographic information of the participating teachers is shown in Table 1.

Table 1

Demographic Information of Teacher Informants

Characteristic	*n*
Location of university	
UK	4
Russia	5
Australia	6
Gender	
Male	8
Female	7
Teaching experience	
1–5 years	5
> 5 years	10

Note. N = 15.

Research Method

We used a phenomenological research approach, which thoroughly explores the lived experiences of people affected by a particular phenomenon. Phenomenology is frequently used in studies with little prior knowledge (Cohen et al., 2018). Every participant received a thorough explanation of the study's procedures, and participation was voluntary (Bogdan & Biklen, 2007). We chose interviews, which comprised individual meetings, phone calls, and online exchanges, from among the different techniques for gathering data for qualitative research (Creswell, 2014). We conducted electronic interviews in two formats: online and offline (via email) because of the pandemic's effects and the participants' geographic distances from us. The option to respond in Mandarin or English was offered. Language, psychological, and educational barriers were among the main issues that participants were asked to address when discussing their experiences teaching or learning Mandarin. We listened to and carefully examined each online interview before transcribing it. After that, the recordings were played back to capture the exact words of the participants. The goal was to accurately capture colloquial expressions and phrases, considering the informal nature of the interviews. We transcribed the participants' comments daily following each interview session. Each interview was roughly 33 minutes, although discussions varied between 20 and 50 minutes. Generally, we carried out one or two interviews daily while allocating the rest of the time to transcription.

Data Analysis

Using MAXQDA software (Version 2022), data analysis was completed in accordance with Creswell's (2014) recommendations. The primary analytical unit was sentences, and the analysis's primary focus was manifest content as opposed to latent content. Every step of the qualitative data collection, analysis, and reporting process was done in English. Since this study was not based on theories or frameworks that already existed, an inductive method of content analysis was employed. To analyze qualitative data, we followed a five-step process based on the framework proposed by Gao and Zhang (2020). First, the data underwent a rigorous cleaning process to eliminate linguistic errors, ambiguities, and repetitions. Second, after reading the data multiple times, we generated open codes. Third, these open codes were used to develop axial codes and subthemes. Fourth, by organizing axial codes and subthemes, higher-order general themes and selective codes were produced. To document each stage of the data analysis and interpretation procedure, a comprehensive report was written.

Results

Research Question 1

The first research question addressed the Chinese language teachers' perceptions of how AI-powered applications (ChatGPT and Poe) help Chinese language learners overcome linguistic problems. The interviews with experts were content analyzed, and the themes which emerged are presented in Table 2.

Table 2

AI-Powered Applications Usage in Addressing Learners' Linguistic Challenges

Theme	Frequency, ***n*** (%)
AI applications can solve pronunciation challenges.	20 (100)
AI applications can raise learners' awareness of tonal discrimination.	18(90)
AI applications can raise learners' awareness of tonal variation in dialects.	18 (90)
AI applications can help learners learn characters and writing.	14 (70)
AI applications can help learners learn grammar and sentence structure.	13 (65)
AI applications can help learners learn idiomatic expressions.	12 (60)
AI applications can help learners learn idiomatic usage.	12 (60)
AI applications can help learners learn cultural nuances.	12 (60)
AI applications can help learners learn proverbs and symbolism.	16 (80)
AI applications are useful for vocabulary acquisition and expansion.	17 (85)

All participants stated that speech recognition algorithms powered by AI can analyze learners' pronunciation, providing real-time feedback. AI can identify specific areas of difficulty, suggest corrections, and offer targeted exercises to improve pronunciation. About 90% of teachers argued that AI-driven language learning platforms can incorporate interactive exercises focusing on tonal discrimination. Speech synthesis technology can generate various tones, allowing learners to practice distinguishing between them in a controlled and supportive environment.

Moreover, 90% of participants stated that AI can expose diverse dialects through audio samples and interactive exercises. Speech recognition can assist learners in adapting to tonal variations, helping them understand and communicate effectively across different regional variations. Seventy per cent of the informants argued that AI-powered language learning apps often include handwriting recognition and character analysis tools. These tools can evaluate learners' writing skills, offer corrections, and provide character stroke order guidance to enhance writing proficiency.

Grammar and sentence structure were identified by 65% of participants as another thematic area in which AI-powered applications could help learners. Natural language processing (NLP) algorithms can analyze written text to identify grammatical errors and suggest corrections. AI-driven chatbots or language tutors can engage learners in conversations, correct sentence structure, and reinforce proper grammar usage.

Among participants, 60% stated that language learning applications can integrate databases of idiomatic expressions. AI algorithms can provide explanations, examples, and context for using idioms. Interactive exercises and quizzes can reinforce understanding and application. The same number of informants stated that AI chatbots or virtual language partners can simulate real-life conversations, incorporating idiomatic expressions in context. Learners can engage in dialogue to practice the correct usage of idioms, receiving feedback from the AI system.

Participants (60%) also stated that AI can provide cultural context through multimedia content, including videos, articles, and interactive scenarios. Natural language understanding algorithms enable AI to explain cultural nuances, ensuring learners comprehend the broader context of language use. Furthermore, 80% of participants said that the contribution of language learning platforms can include modules focused on proverbs and symbolism. AI algorithms can break down the meaning of proverbs, explain cultural symbolism, and offer exercises to reinforce understanding and usage. Finally, 85% of informants stated that AI-driven apps employ spaced repetition algorithms to optimize vocabulary learning. These algorithms adapt to learners' proficiency levels, ensuring they review and practice words optimally for effective retention and expansion.

Research Question 2

The second research question addressed the Chinese language teachers' perceptions of how AI-powered applications (ChatGPT and Poe) help Chinese language learners overcome psychological problems. The interviews with experts were content analyzed, and the themes which emerged are presented in Table 3.

Table 3

AI-Powered Applications Usage in Addressing Learners' Linguistic Challenges

Theme	Frequency, *n* (%)
AI applications can solve interference from the learners' L1.	20 (100)
AI applications can reduce age-related challenges.	18(90)
AI applications can positively affect learners' motivation and attitude.	17 (85)
AI applications can affect students' preferences for other languages.	14 (70)
AI applications can affect parents' and learners' resilience to learn Mandarin.	13 (65)
AI applications can decrease learners' fear of making mistakes.	11 (55)
AI applications can reduce learners' anxiety in speaking and conversing.	11 (55)
AI applications can remedy the lack of personalized support.	11 (55)

Note. L1 = first language.

As seen in Table 3, the majority of participants stated that AI-driven language learning platforms can incorporate personalized modules that specifically target common challenges arising from learners' native language interference. Adaptive exercises, pronunciation analysis, and targeted lessons can address this interference. Similarly, participants argued that AI can tailor language learning content and methodologies based on age. Interactive games, storytelling, and visually engaging content can be employed for younger learners. AI can customize lessons for adults to align with their cognitive abilities and learning preferences.

Furthermore, participants believe that AI-integrated applications can positively affect the learners' motivation and attitude. AI can employ motivational strategies, such as gamification, rewards, and personalized learning paths. Virtual language tutors powered by AI can adapt teaching styles based on individual preferences, fostering a positive attitude towards learning Mandarin. Participants also stated that AI can affect Chinese learners' preferences for other languages. They believed that AI could engage language learning by incorporating content related to students' interests. For instance, if a learner prefers a particular language or cultural context, AI can integrate relevant materials, making Mandarin learning more appealing and relevant.

The next theme was labeled parents' and learners' resilience to learn Mandarin, which is affected by AI-generated applications. Participants stated that AI-powered language learning platforms can provide continuous support and encouragement. Virtual tutors can offer positive reinforcement, progress tracking, and personalized feedback. Additionally, AI can facilitate community building, connecting learners with similar goals and fostering a supportive learning environment.

Participants also argued that AI-powered educational applications can reduce learners' anxiety and fear of making mistakes when speaking. The teachers believed that language learners often fear making mistakes, hindering their willingness to practice and engage actively in language learning activities. AI can create a non-judgmental learning environment. Virtual tutors can provide constructive feedback, emphasizing learning from mistakes. AI-powered language learning apps can offer safe spaces for trial and error, boosting learners' confidence.

Moreover, participants stated that learners may experience anxiety, particularly when it comes to speaking and holding conversations in the target language. However, AI can facilitate speaking practice through virtual interactions. AI-driven chatbots or virtual conversation partners can engage learners in realistic dialogues, providing a low-pressure environment to practice conversational skills.

The last psychologically related theme was the lack of personalized support for which AI-powered applications can be practical. For instance, some participants stated that in traditional classroom settings, teachers may need help to provide personalized attention to each learner's unique needs and challenges. AI can offer individualized learning experiences. Adaptive learning algorithms can identify areas of difficulty for each learner and tailor lessons accordingly. AI-powered virtual tutors can provide personalized guidance and support, addressing specific learning gaps.

Research Question 3

The third research question addressed Chinese language teachers' perceptions of how AI-powered applications (ChatGPT and Poe) help Chinese language learners overcome the educational problems of teaching the Chinese language to non-Chinese language learners. The interviews were content analyzed, and the themes which emerged are presented in Table 4.

Table 4

AI-Powered Applications Used in Solving Educational Problems

Theme	Frequency, *n* (%)
AI applications can help teachers develop course and curriculum.	18 (90)
AI applications can remedy the lack of resources and materials.	14 (70)
AI applications can help teachers know about pedagogical approaches and strategies.	13 (65)
AI applications can improve teachers' pedagogical knowledge.	12 (60)
AI applications can solve teacher training problems.	12 (60)
AI applications can be used for assessment and feedback.	12 (60)
AI applications can be used for individualized learning paths.	16 (80)
AI applications are useful for cultural understanding.	17 (85)

AI can play a crucial role in addressing various educational problems faced by learners of the Chinese language. The first type of educational problem AI can solve is thematically coded as course design and curriculum. Participants mentioned that this type of problem can be solved in three different ways. First, as suggested by most teachers, AI can personalize learning experiences by assessing individual student performance and tailoring course content accordingly. This adaptive learning approach ensures learners progress at their own pace, reinforcing concepts they find challenging while advancing quickly through familiar material. Second, AI-driven recommendation systems can suggest supplementary materials, practice exercises, and multimedia resources based on learners' proficiency levels, interests, and learning styles. This enhances the overall learning experience and provides a more comprehensive language understanding. Third, AI can analyze language trends, cultural changes, and real-time linguistic data to update and modify curriculum content. This ensures that learners are exposed to current and relevant language usage, keeping the curriculum dynamic and engaging.

The second theme that emerged from interviews is participants' view that AI can remedy the need for more resources and materials for teaching the Mandarin language. Participants stated that AI can assist in generating language learning materials, including interactive exercises, quizzes, and culturally relevant content. This addresses resource shortages by continuously supplying new and diverse learning materials. Participants also stated that AI-powered translation tools aid learners in understanding and translating complex Chinese texts. These tools can also offer multilingual support, helping learners of different native languages.

The teachers we interviewed also stated that AI-powered technology can raise Chinese language teachers' awareness of pedagogical approaches and strategies. Participants mentioned that intelligent teaching systems (ITS) can simulate one-on-one interactions with a teacher, providing instant feedback, guidance, and targeted support. This personalized learning experience helps students overcome specific language challenges and reinforces positive learning behaviors. Participants also mentioned that AI can facilitate gamified learning experiences, making language acquisition enjoyable and engaging. Interactive scenarios, role-playing, and language games enhance the learning process, making it more effective and enjoyable.

A fourth theme that emerged concerns how teachers can use AI applications to improve their pedagogical knowledge. AI can contribute to ongoing teacher training programs by offering individualized professional development modules. These modules can focus on emerging pedagogical techniques, incorporating the latest language education and technology advancements. Teachers also believe that AI-driven assessment tools can assist teachers in evaluating student performance efficiently. This allows educators to focus on providing targeted feedback and addressing specific learning needs.

The fifth theme we noted concerns participants' view that AI-powered applications can be used to address teacher training problems through virtual simulations and data-driven insights. Teachers stated that AI can create virtual classroom simulations where teachers can practice various instructional strategies, manage diverse student needs, and receive feedback. This virtual training environment helps teachers develop effective classroom management and teaching skills. Participants also stated that AI analytics can provide valuable insights into teachers' performance and areas that require improvement. This data-driven approach helps in tailoring professional development programs to address specific needs.

The sixth theme mentioned in the interviews concerns how AI-powered technology could be used for assessment and feedback purposes through automated grading systems. Teachers believe AI-powered grading systems can efficiently evaluate written assignments, essays, and language assessments. These systems save educators time and provide instant feedback to learners, helping them understand their mistakes and areas for improvement. Teachers also believe that speech recognition for pronunciation can be used for assessment purposes. Providing real-time feedback on intonation, accent, and fluency helps learners refine their spoken Chinese language skills.

The seventh extracted contribution of AI-powered applications to educational problems was thematically labeled as individualized learning paths. Informants stated that AI can analyze learners' interactions with educational content to identify individual strengths and weaknesses. This data-driven approach enables the creation of personalized learning paths, recommending specific exercises or activities tailored to each learner's needs. They also stated that AI tutors can adapt to learners' cognitive styles and preferences, providing customized learning experiences. By understanding individual learning patterns, AI can offer targeted support, ensuring that each student masters Chinese language concepts at their own pace.

Finally, participants believe that AI-powered applications can be used for cultural understanding. For instance, AI can enhance language learning by incorporating cultural context into lessons. Virtual cultural experiences, language immersion scenarios, and AI-generated content related to Chinese culture help learners better understand language nuances, idioms, and cultural references. Participants believe AI algorithms can monitor and analyze current events and trends in Chinese-speaking regions, providing learners with real-time cultural insights. This dynamic approach ensures that learners are linguistically proficient and culturally aware.

Discussion

Teaching and learning methodologies have been profoundly impacted by the incorporation of AI into a variety of educational fields. With regard to the first research question, it was found that AI applications offer comprehensive support for language learners, addressing pronunciation challenges and enhancing awareness of tonal discrimination and variation in dialects. These technologies aid in mastering characters, writing, grammar, sentence structure, and idiomatic expressions, including their usage. Furthermore, AI helps learners understand cultural nuances, proverbs, and symbolism, thereby providing a holistic approach to language acquisition. This finding is echoed in the study by Ajabshir (2013).

The study's conclusions demonstrate that AI can assist students in resolving linguistic-related issues when learning Chinese by drawing on insights from a variety of fields, including medical education, language learning, and the difficulties faced by Chinese language teachers.

As shown by Cheng et al. (2020), AI is being used in medical education. Its use goes beyond diagnosing and includes helping medical students understand complicated cases like hip fractures. These studies highlight how by offering sophisticated tools for interpretation and analysis, AI can improve medical education.

Findings also revealed that AI applications effectively address various challenges in language learning, such as interference from the learners' first language (L1) and age-related difficulties. They positively impact learners' motivation and attitudes, influence preferences for other languages, and enhance resilience in learning Mandarin. Additionally, AI reduces the fear of making mistakes and speaking anxiety, while also providing personalized support that is often lacking in traditional educational settings. Therefore, it can be strongly argued that AI-applications can be used to enhance personalized learning which was supported by a number of researchers (Alibakhshi, 2013). In constructing and validating self-assessment inventories and teaching motivation scales, Alibakhshi and Nezakatgoo (2019) emphasized the importance of personalized approaches in language education. These findings align with the broader theme of AI's role in tailoring language learning experiences based on individual needs.

The One Belt and One Road (OBOR) initiative's advantages and disadvantages are examined in relation to Chinese language instruction. The OBOR, also known as the Belt and Road Initiative (BRI), is a global development strategy adopted by the Chinese government in 2013. It aims to enhance regional connectivity and embrace a brighter economic future through building infrastructure and broadening trade links between Asia, Africa, and Europe. Furthermore, the research conducted by Kang and Kang (2023) highlights the importance of developing a deep learning based Chinese language teaching system model within the context of AI. These observations emphasize how the field of language education is changing and how AI has the potential to significantly influence teaching approaches and curriculum development. The results corroborate those of Gao and Zhang (2020), who investigated the beliefs of foreign language instructors regarding online instruction in difficult times, demonstrating the flexibility of AI-enabled teaching strategies. The enhancement of motivation and disposition, the mitigation of fear and anxiety by means of nonjudgmental surroundings, and the provision of ongoing assistance are consistent with the wider psychological motifs deliberated by educators. Cognitive difficulties in combining AI and deep learning for breast cancer screening (Derun et al., 2019) and the more general difficulties and possibilities in AI applications (Wang, 2021) draw attention to the necessity of ongoing study and modification. The results highlight the significance of tackling obstacles and optimizing the advantages offered by AI, underscoring a well-rounded and knowledgeable strategy for execution. Generally speaking, this study illuminates the revolutionary potential of AI in resolving numerous educational obstacles faced by Chinese language learners. Through the integration of participant perspectives and pertinent scholarly works, this conversation clarifies the consistency and resilience of the recognized themes. In line with Gligorea et al. (2023), AI clearly emerges as a key component for customizing language instruction to each student's needs. The participants' focus on an approach to adaptive learning is in perfect alignment with the findings of Gligorea et al.'s (2023) demonstration of AI's revolutionary potential in enabling customized learning experiences that meet the needs of each individual student. The insights from the participants regarding AI producing a variety of educational resources and offering multilingual assistance are consistent with Kang and Kang's (2023) acknowledgment of AI's function in mitigating resource limitations in Mandarin language instruction.

Regarding innovative pedagogy and teacher assistance, the research finds resonance with Wang's (2021) and Cheng et al.'s (2020) work. Wang (2021) emphasized how AI can provide teachers with tailored professional development, and Cheng et al. (2020) emphasized the significance of personalized feedback powered by AI in medical education. Their results are consistent with the opinions of the participants regarding how AI can revolutionize education. As we move on to teacher training, Wu and Itayi's (2022) investigation of AI's function in offering virtual training environments for teachers is in line with integrating AI in virtual simulations and data-driven insights. This alignment highlights the potential of AI analytics to provide insightful information about teachers' performance and areas in need of development, thereby reinforcing participants' views on AI in teacher training. Regarding evaluation, the research is consistent with Agarwal et al. (2021), who used AI to classify diseases in medical imaging. The participants agreed with Gligorea et al. (2021) regarding the effectiveness of AI in assessing written assignments and language assessments. Regarding personalized learning paths, this research concurs with Kang and Kang (2023) who claimed that AI can analyze learner interactions and generate tailored learning paths. The focus on AI's capacity to assess how students engage with instructional materials aligns with participants' perceptions that AI can customize learning opportunities according to each student's unique strengths and shortcomings. Lastly, the participants' perceptions of AI's value in fostering cultural understanding are consistent with more general discourse about AI's potential to improve language acquisition. Yi et al.'s (2020) investigation into the application of deep learning and hyperspectral imaging technology in traditional Chinese medicine supports the participants' belief that AI can enhance language learning by incorporating cultural context.

Conclusions and Implications

In conclusion, this study highlights the transformative impact of AI on teaching and learning practices in second language acquisition. Drawing insights from the challenges faced by Chinese language teachers, the findings underscore AI's potential in addressing linguistic, psychological and education related issues for Chinese language learners. The study aligns with existing literature, showcasing AI's role in tailoring language learning experiences, mitigating resource shortages, reshaping pedagogy, and providing valuable insights into teaching and assessment.

The study focuses on AI's ability to customize language instruction to meet the needs of each learner, which is consistent with the literature's emphasis on personalized learning experiences. AI becomes a key factor in curriculum design and teaching approaches in the context of teaching Chinese, where opportunities and challenges are abundant because of programs such as One Belt and One Road. The study is consistent with earlier research, bolstering the notion that AI can produce a variety of learning resources and efficiently address resource constraints in language instruction. The study further supports the multifaceted impact of AI on the changing distance education landscape through pedagogical innovation, teacher support, training, assessment, and customized learning paths. The participants' perceptions of AI's role in enhancing cultural awareness are consistent with contemporary discourse about how AI can improve language learning experiences by incorporating cultural context, which represents a paradigm shift in how distance learning is perceived.

This study has several implications for educational practice and policy. The positive impact of AI on motivation, the reduction of anxiety, and adaptability in challenging teaching environments emphasize the need for continued exploration and integration of AI in education. Policymakers should consider fostering a supportive environment for AI integration in educational institutions, promoting research, and addressing challenges while maximizing opportunities. Continuous collaboration between educators, researchers, and AI developers is crucial to ensure a balanced and informed approach to AI implementation in education. The transformative potential of AI in addressing linguistic and cultural challenges positions it as a valuable ally in shaping the future of language education.

Acknowledgments

We acknowledge that this paper is novel, and all references were mentioned in the text. We also acknowledge that AI paraphrasing and editing tools such as Grammarly and ChatGPT were used for editing some parts of the manuscript.

This study is supported via funding from Prince Sattam bin Abdulaziz University Project Number (PSAU/2024/R/1445).

References

Agarwal, M., Saba, L., Gupta, S. K., Carriero, A., Falaschi, Z., Paschè, A., Danna, P., El-Baz, A., Naidu, S., & Suri, J. S. (2021). A novel block imaging technique using nine artificial intelligence models for COVID-19 disease classification, characterization, and severity measurement in lung computed tomography scans on an Italian cohort. *Journal of Medical Systems*, *45*, Article 28. https://www.doi.org/10.1007/s10916-021-01707-w

Ajabshir, Z. F. (2023). A review of the affordances and challenges of artificial intelligence technologies in second language learning. *Technology Assisted Language Education*, *1*(4), 96-115. DOI: 10.22126/TALE.2024.10104.1028

Alibakhshi, G. (2013). Construction and validation of self-assessment inventory for English for academic purposes: A case of Iranian tertiary students. *Journal of Research in Applied Linguistics, 4*(2), 93–109.

Alibakhshi, G., & Nezakatgoo, B. (2019). Construction and validation of Iranian EFL teachers' teaching motivation scale. *Cogent Education, 6*(1), Article 1585311. https://www.doi.org/10.1080/2331186X.2019.1585311

Bogdan, R. C., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theory and methods* (5th ed.). Allyn & Bacon.

Cheng, C. T., Chen, C. C., Fu, C. Y., Chaou, C. H., Wu, Y. T., Hsu, C. P., Chang, C. C., Chung, I. F., Hsieh, C. H., Hsieh, M. J., & Liao, C. H. (2020, November 23). Artificial intelligence-based education assists medical students' interpretation of hip fractures. *Insights Imaging, 11*(1), Article 119. https://doi.org/10.1186/s13244-020-00932-0

Chen, M., Mao, S., Zhang, Y., & Leung, V. C. (2014). *Big data: Related technologies, challenges and future prospects* (Vol. 100). Heidelberg: Springer.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education*. Routledge.

Cox, S. E. (2020). *Perceptions and influences behind teaching practices: Do teachers teach as they were taught?* [Master's thesis, Brigham Young University]. BYU Scholars Archive. https://scholarsarchive.byu.edu/etd/5301

Creswell, J. W. (2014). *Research design: Qualitative, quantitative and mixed methods approaches* (4th ed.). Sage.

Derun, P., Genggeng, Q., & Weiguo, C. (2019). Application progress of artificial intelligence technology based on deep learning in breast cancer screening and imaging diagnosis. *International Journal of Medical Radiology*, *42*(01), 15–18. https://inis.iaea.org/search/search.aspx?orig_q=RN:54023411

Everson, M. (2008). Issues in Chinese literacy learning and implications for teacher development. In P. Duff & P. Lester (Eds.), *Issues in Chinese language education and teacher development* (pp. 70–78). University of British Columbia, Centre for Research in Chinese Language and Literacy Education. https://www.academia.edu/9682120/Issues_in_Chinese_Language_Education_and_Teacher_Development_2008

Gao, L. X., & Zhang, L. J. (2020). Teacher learning in difficult times: Examining foreign language teachers' cognitions about online teaching to tide over COVID-19. *Frontiers in Psychology, 11,* Article 549653. https://doi.org/10.3389/fpsyg.2020.549653

Gligorea, I., Cioca, M., Oancea, R., Gorski, A. T., Gorski, H., & Tudorache, P. (2023). Adaptive learning using artificial intelligence in e-learning: A literature review. *Education Sciences*, *13*(12), 1216.

Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. – C., & Jeste, D. V. (2019). Artificial intelligence for mental health and mental illnesses: An overview. *Current Psychiatry Reports, 21*, Article 116. https://www.doi.org/10.1007/s11920-019-1094-0

Guo, Y., Ren, X., Chen, Y.-X., & Wang, T.-J. (2019). Artificial intelligence meets Chinese medicine. *Chinese Journal of Integrative Medicine, 25*(9), 648–653. https://www.doi.org/10.1007/s11655-019-3169-5

Hartig, F. (2015). *Chinese public diplomacy: The rise of the Confucius Institute*. Routledge.

Hirschmann, A., Cyriac, J., Stieltjes, B., Kober, T., Richiardi, J., & Omoumi, P. (2019, June). Artificial intelligence in musculoskeletal imaging: Review of current literature, challenges, and trends. *Seminars in Musculoskeletal Radiology*, *23*, (3), 304–311. https://www.doi.org/10.1055/s-0039-1684024

Jiang, Y., Chen, Y., Lu, J., & Wang, Y. (2021). The effect of the online and offline blended teaching mode on English as a foreign language learners' listening performance in a Chinese context. *Frontiers in Psychology*, *12*, 742742. DOI: 10.3389/fpsyg.2021.742742

Jin, Q. (Ed.). (2011). *Intelligent learning systems and advancements in computer-aided instruction: Emerging studies*. IGI Global.

Johnson, K. W., Torres Soto, J., Glicksberg, B. S., Shameer, K., Miotto, R., Ali, M., Ashley, E., & Dudley, J. T. (2018). Artificial intelligence in cardiology. *Journal of the American College of Cardiology*, *71*(23), 2668–2679. https://www.doi.org/10.1016/j.jacc.2018.03.521

Kang, B., & Kang, S. (2023). Construction of a Chinese language teaching system model based on deep learning under the background of artificial intelligence. *Scientific Programming*, *2022*, Article 3960023. https://doi.org/10.1155/2022/3960023

Li, H. (2018). Deep learning for natural language processing: Advantages and challenges. *National Science Review*, *5*(1), 24–26. https://doi.org/10.1093/nsr/nwx110

Lu, X., Ai, W., Liu, X., Li, Q., Wang, N., Huang, G., & Mei, Q. (2016, September). Learning from the ubiquitous language: an empirical analysis of emoji usage of smartphone users. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing* (pp. 770-780). https://doi.org/10.1145/2971648.2971724

Makhambetova, A., Zhiyenbayeva, N., & Ergesheva, E. (2021). Personalized learning strategy as a tool to improve academic performance and motivation of students. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, *16*(6), 1-17. DOI:10.4018/IJWLTT.286743

Miao, S., Xu, T., Wu, Y., Xie, H., Wang, J., Jing, S., Zhang, Y., Zhang, X., Yang, Y., Zhang, X., Shan, T., Wang, L., Xu, H., Wang, S., & Liu, Y. (2018). Extraction of BI-RADS findings from breast ultrasound reports in Chinese using deep learning approaches. *International Journal of Medical Informatics*, *119*, 17–21. https://www.doi.org/10.1016/j.ijmedinf.2018.08.009

Orton, J. (2011). Educating Chinese language teachers—Some fundamentals. In L. Tsung & K. Cruickshank (Eds.), *Teaching and learning Chinese in global contexts: Multimodality and literacy in the new media age* (pp. 151–164). Continuum.

Peng, H., Ma, S., & Spector, J. M. (2019). Personalized adaptive learning: an emerging pedagogical approach enabled by a smart learning environment. *Smart Learning Environments*, *6*(1), 1-14. https://slejournal.springeropen.com/articles/10.1186/s40561-019-0089-y

Pérez-Milans, M. (2015). Mandarin Chinese in London education: Language aspirations in a working-class secondary school. *Language Policy*, *14*, 153–181. https://doi.org/10.1007/s10993-014-9345-8

Wang, R. (2021). Research on the application of deep learning in artificial intelligence courses. *Journal of Electronic Research and Application*, *5*(6), 14–18. https://www.doi.org/10.26689/jera.v5i6.2760

Wang, X. (2004). Encouraging self-monitoring in writing by Chinese students. *ELT Journal, 58*(3), 238–246. https://doi.org/10.1093/elt/58.3.238

Wang, Y., & Lemmer, V. (2015). Teaching Chinese as a foreign language in higher education in China and South Africa: Lecturers' views. *Per Linguam, 31*(2), 35–52. https://doi.org/10.5785/31-2-608

Wu, W. J., & Mareya, I. A. (2022). Benefits and challenges faced by foreign students of Chinese language after graduation. *Journal of Applied Linguistics and Language Learning, 5*(1), 12–18. https://doi.org/10.5923/j.jalll.20220501.02

Xiao, Y., & Liu, S. (2019). Collaborations of industry, academia, research and application improve the healthy development of medical imaging artificial intelligence industry in China. *Chinese Medical Sciences Journal*, *34*(2), 84–88. https://doi.org/10.24920/003619

Yalun, A. (2019). International promotion of Chinese language in the new era. *International Education Studies, 12*(7), 67–79. https://www.doi.org/10.5539/ies.v12n7p67

Ye, L. (2013). Shall we delay teaching characters in teaching Chinese as a foreign language? *Foreign Language Annals, 46*(4), 610–627. https://doi.org/10.1111/flan.12049

Yi, T., Lin, C., En-Ci, J., & Ji-Zhong, Y. (2020). Application and prospects of hyperspectral imaging and deep learning in traditional Chinese medicine in context of AI and industry 4.0. *Zhongguo Zhong yao za zhi= Zhongguo Zhongyao Zazhi= China Journal of Chinese Materia Medica*, *45*(22), 5438-5442.https://www.doi.org/10.19540/j.cnki.cjcmm.20200630.603

Yue, Y. (2017). Teaching Chinese in K-12 schools in the United States: What are the challenges? *Foreign Language Annals, 50*(3), 601–620. https://doi.org/10.1111/flan.12277

Zhu, Z., Shan, J., & Yan, H. (2019). International investigation and development strategy for artificial intelligence maker education. *Open Education Research, 25*(1), 47–54. https://doi.org/10.1155/2022/3960023

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

Artificial Intelligence in Higher Education: A Cross-Cultural Examination of Students' Behavioral Intentions and Attitudes

Dongmin Ma[1], Huma Akram[2,*], and I-Hua Chen[3]
*[1]School of Foreign Studies, North China University of Water Resources and Electric Power, Zhengzhou, Henan, China; [2]School of International Education, North China University of Water Resources and Electric Power, Zhengzhou, Henan, China; [3]Chinese Academy of Education Big Data, Qufu Normal University, Shandong, China; *Corresponding author*

Abstract

Artificial intelligence (AI) has undergone considerable advancement in the contemporary period and represents an emerging technology in higher education. Cultural contexts significantly shape individuals' perceptions, attitudes, and behaviors, particularly in the realm of technology acceptance. By adopting a cross-cultural lens, this research explores the potential variations across Chinese and international students from diverse countries in terms of attitudes and their behavioral intentions toward AI use. With a technology acceptance model (TAM) framework, the research used a survey approach, employing questionnaires as the primary means of data collection. The data were then analyzed through structural equation modeling and descriptive statistics. A substantial discrepancy was found in the prevalence, attitudes, and behavioral intentions toward AI use between Chinese and international students. Findings further revealed a stronger effect of perceived ease of use on both attitudes and behavioral intentions among international students compared with their Chinese counterparts. Findings suggest that cultural backgrounds and prior technological exposure play intricate roles in shaping perceptions of AI technology. The study emphasizes the need for tailored educational strategies to regulate diverse cultural perspectives, provide language-specific support, and ensure user-friendly interfaces. These insights contribute to the evolving discourse on technology acceptance in higher education and offer practical implications for educators and institutions toward optimizing AI integration in pedagogical practices.

Keywords: artificial intelligence, higher education, technology acceptance model, attitudes, behavioral intentions

Artificial Intelligence in Higher Education: A Cross-Cultural Examination of Students' Behavioral Intentions and Attitudes

Over the last few years, artificial intelligence (AI) has emerged as a transformative force, reshaping various aspects of our lives. Its influence has extended into the realm of education, promising to revolutionize traditional teaching and learning methods. The integration of AI in higher education stands out as a beacon of innovation, holding the potential to produce sustainable growth in students' learning experiences (Ouyang & Jiao, 2021), foster retention (L. Chen et al., 2020), strengthen academic motivation (Yilmaz & Yilmaz, 2023), increase academic performance (Zhou, 2023), promote self-directed learning skills (Lasfeto & Ulfa, 2023), support language learning (Crompton & Burke, 2023), and develop problem-solving skills (Zhang & Zhu, 2022). AI refers to the development of computers that can perform tasks that are normally associated with human intelligence (Ertel, 2018) and encompasses a wide array of functions such as problem-solving, language understanding, visual perception, and decision-making. In the context of education, AI presents an exceptional opportunity to tailor learning experiences to individual needs, providing personalized pathways for students to explore and master their academic pursuits (Alam, 2021).

The integration of AI in education is a multifaceted process influenced by various factors, such as technological infrastructure (Matsika & Zhou, 2021), institutional policies (Cheng & Wang, 2022), self-efficacy (Yilmaz & Yilmaz, 2023), and pedagogical approaches (Chan, 2023). In addition to other factors, students' personal perceptions and attitudes play essential roles in the successful implementation of AI in higher education. Studies have constantly shown that students' learning experiences are greatly affected by their beliefs about using or engaging with technologies in their educational practices (Akinoso, 2023; Zulaiha & Triana, 2023). As active participants in the learning process, students show preferences for technological applications that correspond with their attitudes and strategies for learning (Abdelrady & Akram, 2022; Guo et al., 2023). This alignment stems from a desire to improve their educational experience via the effective incorporation of technology (Akram et al., 2022; Linardatos & Apostolou, 2023). On the other hand, individuals with a negative attitude toward technology are less likely to be interested in technology-driven learning and hold negative views of tech-based resources, hindering their abilities to effectively incorporate technology into their learning (Akram, Yingxiu, et al., 2021a; Y. Wang, 2023). Therefore, individuals' willingness and ability to use AI in education are key factors that determine how well the technology is integrated. Furthermore, the debut of new technology requires a thorough examination of individuals' attitudes toward it and the elements that shape those perceptions (Makumane, 2023). In terms of AI in higher education, examining students' perspectives is a critical step in understanding the dynamics of their relationship with this novel technology.

Cultural differences can also greatly influence attitudes toward technology in the educational process. Diverse cultures exhibit different policies regarding and accessibility toward digital technologies (Akram & Yang, 2021), leading to gaps in digital competences such as computational thinking and AI literacy between people of different backgrounds (Kayalar, 2016; Savicki, 2023). These differences can result in distinct ways of interacting with technology, which can influence people's attitudes, perceptions, and beliefs (Vargo et al., 2021). Over time, these disparities may lead to significant differences in professional paths, economic standing, and various other aspects of life.

Over the last few decades, the landscape of higher education in China has undergone a revolutionary transition due to significant growth in the number of international students arriving to study (Akram et al., 2020; Jiani, 2017). This influx has brought cultural diversity to Chinese institutions, offering a robust platform for intellectual enrichment and signifying China's dedication to globalization and cross-cultural interactions (Dai & Hardy, 2023). The interaction between cultures brings different elements or characteristics to the academic landscape, making it an interesting area for research, especially concerning how new technologies like AI can be introduced and used in education. As the landscape of AI in education continues to evolve, it is crucial to understand how students from diverse cultural backgrounds, sharing the common space of Chinese educational institutions, view and engage with it. Therefore, the study aims to address the following research objectives (ROs):

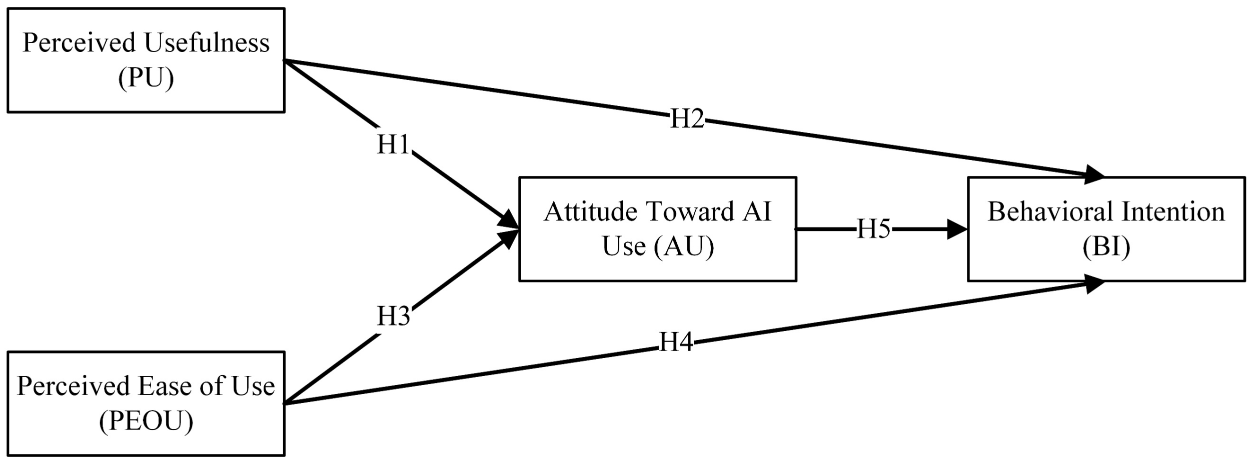
> RO1: To explore and compare the attitudes and behavioral intentions of both Chinese and international students regarding the integration of AI in higher education.
>
> RO2: To identify and analyze the key determinants influencing the attitudes and behavioral intentions of students toward using AI in higher education.

Theoretical Framework and Hypotheses

Examining individuals' attitudes is crucial to determining the success rate of newly introduced technology. Several theories have been developed to investigate and explain the reasons that lead users to embrace, reject, or continue to use a new technology (Venkatesh & Davis, 2000; Venkatesh et al., 2003). The technology acceptance model (TAM) is a widely recognized theoretical framework that helps explain users' intentions and attitudes to use technology (see Figure 1). In the context of students' intentions and attitudes to use AI in higher education, the TAM is particularly relevant as it offers insights into the factors that influence users' decisions to integrate AI into their learning practices (Davis, 1989). This model consists of two primary constructs, which are believed to be the most important factors in determining the level of user approval.

Figure 1

Conceptual Framework



Perceived Usefulness

Perceived usefulness (PU) refers to the extent to which individuals believe that using a specific technology will enhance their ability to accomplish their work (Davis, 1989). Students who perceive technological applications as useful in their studies are more likely to integrate them into their educational practices (Akram, Aslam, et al., 2021). In the current study, PU refers to how AI can improve students' learning and conserve their time and energy. Conversely, individuals who do not perceive technology as useful may find it challenging to use (Davis, 1989). X. Chen et al. (2020) found that PU significantly influenced students' behavioral intentions to use ChatBot in Chinese learning. Similarly, Alfadda and Mahdi (2021) demonstrated a positive and significant correlation between PU and students' attitudes about using Zoom in language learning. Anthony et al. (2022) also established a strong correlation between the perceived advantages of technical courses and their effective adoption. This suggests that students are more likely to incorporate technology into their learning process if they view it as useful and applicable. Therefore, the following hypotheses were put forward:

> H1: PU would significantly and positively influence users' attitudes toward the use of AI in both Chinese and international students.

> H2: PU would significantly and positively influence users' behavioral intentions toward the use of AI in both Chinese and international students.

Perceived Ease of Use

Perceived ease of use (PEOU) refers to a user's comfort and ease with using technology (Davis, 1989). It influences students' behavioral predisposition of ultimate use (Sathye et al., 2018). Several studies provide evidence that PEOU plays a key role in shaping users' attitudes and behavioral intention toward use of technology (e.g., Al-Hattami, 2023; Elfeky & Elbyaly, 2023). Elfeky and Elbyaly (2023) conducted a quasi-experimental study and found PEOU as a significant determinant of students' attitude toward and their behavioral intention to use a learning management system. In agreement with this finding, Barrett et al. (2023) emphasized that perceived technological barriers might prevent users from successfully embracing and accepting technology. Similarly, a study conducted in the Chengdu region of China brought forth a noteworthy revelation regarding the influence of PEOU on the behavioral intention of students (Min et al., 2023), underscoring the significance of promoting the benefits and user-friendliness of online learning systems to encourage students' adoption and boost their satisfaction when using digital learning resources. Therefore, the following were hypothesized:

> H3: PEOU significantly and positively influences users' attitudes toward the use of AI in both Chinese and international students.

> H4: PEOU significantly and positively influences users' behavioral intentions toward the use of AI in both Chinese and international students.

Attitudes Toward Use

A growing body of research supports the idea that users' attitudes toward system usage (AU) strongly influence their behavioral intentions, which ultimately affect their behavior (Al-Mamary, 2022; Yang et al., 2022). This synchronization underscores the critical significance of attitudes in influencing the intentions of users and, eventually, their subsequent behaviors with diverse systems (Akram, Yingxiu, et al., 2021b). Considering this, Almaiah et al. (2022) observed that students often find themselves with insufficient knowledge and incapable of making effective use of newly introduced technologies. Al-Mamary (2022) believes that assessing students' intentions to adopt a new technology is important for analyzing their actual use. In another study, Yang et al. (2022) revealed that across various regions in China, college students' attitudes toward using metaverse technology contributed 33% of the influence on their behavioral intention to use the technology. In other words, students' attitudes about using technology affect how they respond to the technology, leading to the final hypothesis:

> H5: AU significantly and positively influences users' behavioral intentions toward the use of AI in both Chinese and international students.

Methodology

To examine the research objectives and suggested research hypotheses, this study employed a quantitative research method. An online survey consisting of two sections and 16 items was used to collect empirical data. The first section collected the demographic characteristics of the participants, including gender, age, educational level, major, nationality, and the AI platform they used, using a nominal scale. The second section included items derived from the four-construct TAM model and included five items to measure PU and PEOU, adapted from Lewis (2019). It also included three items each measuring AU and BI, adopted from Venkatesh et al. (2003) and Teo (2009). All questions were rated on a 5-point Likert scale, where 1 indicated strong disagreement and 5 indicated strong agreement.

The aim of this cross-sectional study was to investigate the attitudes and behavioral intentions of both Chinese and international students toward the use of AI. To ensure translation quality and consistency, a professional translator first wrote the questionnaire in English, which was then translated into Chinese. The Chinese version was retranslated into English by another translator (Epstein et al., 2015).

This study employed convenience sampling, a non-probability sampling method that allows a researcher to collect data from accessible respondents (Emerson, 2021). An online survey was distributed through the WeChat platform in a Chinese university, targeting both Chinese and international students. Furthermore, to ensure content validity of the self-administered questionnaire, the researchers consulted one English and one Chinese linguistic expert. The questionnaire's clarity and readability were checked through a pilot study involving 50 prospective participants, including both Chinese and international students (Taherdoost, 2016). Based on the pilot study results, two questions were removed. All participants voluntarily participated in the study and were informed of their right to withdraw at any time. The online survey included a cover letter explaining the study's purpose and the participants' rights. No monetary incentives or rewards were provided for participation.

Participants

After screening the data, the study included 689 valid cases from diverse schools of a Chinese university to examine students' attitudes and behavioral intentions toward AI usage. The participants consisted of 372 Chinese and 317 international students.

Table 1 shows the demographic characteristics of participants. Among Chinese participants, 54.8% were female and 45.2% were male, with the majority between the ages of 17 and 24. All Chinese students were pursuing undergraduate education, with over 50% from applied sciences fields and over 40% from social sciences. Of international students, 63.1% were male and 36.9% were female, with the majority (58.1%) between the ages of 21 and 24. Most international participants were pursuing undergraduate education, and 12.6% were postgraduate students. Over 40% of international participants were from natural sciences fields, about 36.5% from applied sciences, and over 20% from social sciences. Most international students were from Asia (66.9%) and Africa (32.5%), with only a few from Australia (0.3%) and Europe (0.3%).

We found a substantial discrepancy in the adoption rate of AI between Chinese and international students. International students showed a higher adoption rate (78%) compared with Chinese students (53%). Diverse preferences for AI platforms were also revealed among both groups. Among Chinese students, ChatGPT was the most widely used platform, followed by Baidu and Bing AI. Other platforms including ChatBot, Google, Youdao, Zhidao, and Bard were used by smaller percentages, and 47.1% of Chinese students did not use any specific AI platform. Similarly, ChatGPT emerged as the top choice for 47.3% of international students, followed by Bing AI and Google. Other platforms including ChatBot, Baidu, Freenome, Midjourney, Nova, Perplexity AI, and QuillBot were used by smaller percentages, and 22.4% of international students expressed reluctance to use any particular AI platform.

Table 1

Demographic Features of Both Samples

Demographic category	Chinese		International	
	Frequency	%	Frequency	%
Gender				
Male	168	45.2	200	63.1
Female	204	54.8	117	36.9
Age				
17–20	281	75.5	72	22.7
21–24	89	23.9	184	58.1
25–28	1	0.2	41	12.9
29+	1	0.2	20	6.3
Educational level				
Undergraduate	372	100.0	277	87.4
Postgraduate	0	0.0	40	12.6
Major				
Social sciences	174	46.7	71	22.4
Applied sciences	198	53.3	116	36.5
Natural sciences	0	0.0	130	41.1
Nationality				
African	0	0.0	103	32.5
Asian	0	0.0	212	66.9
Australian	0	0.0	1	0.3
European	0	0.0	1	0.3
AI platform				
ChatBot	4	1.07	6	1.89
Baidu	52	13.97	9	2.8
Bing AI	59	15.8	35	11.0
ChatGPT	72	19.3	150	47.3
Google	3	0.8	29	9.14
Freenome	0	0.0	1	0.31
Midjourney	0	0.0	4	1.3
Nova	0	0.0	3	0.9
Perplexity AI	0	0.0	2	0.6

QuillBot	0	0.0	2	0.6
None	175	47.1	71	22.4
Youdao	2	0.6	0	0.0
Zhidao	2	0.6	0	0.0
Bard	3	0.8	0	0.0

Data Analysis

Structural equation modeling (SEM) is a statistical technique that allows researchers to estimate connections within a conceptual model and to find correlations between independent and dependent variables (Y. A. Wang & Rhemtulla, 2021). This multivariate approach is an integration of factor analysis and multiple regression that estimates several associations in a single step (Breitsohl, 2019). Given its capacity to manage complex mathematical models and its confirmatory modeling approach (Hair & Alamer, 2022), SEM is ideal for investigating attitudes toward the use of technology and behavioral intentions, especially when dealing with changes between independent and dependent variables (J. Wang & Wang, 2019). Therefore, this study used SEM to analyze the proposed model's data, employing a maximum-likelihood covariance-based SEM approach (SPSS Amos v.26.0), which aligns with the study's emphasis on theory testing and confirmation. Using all constructs treated as latent variables, this study modeled PU and PEOU as predictors, AU as mediator, and BI as outcome. Each of these constructs was measured by uniting distinct indicators—namely, questionnaire items assessing participants' insights.

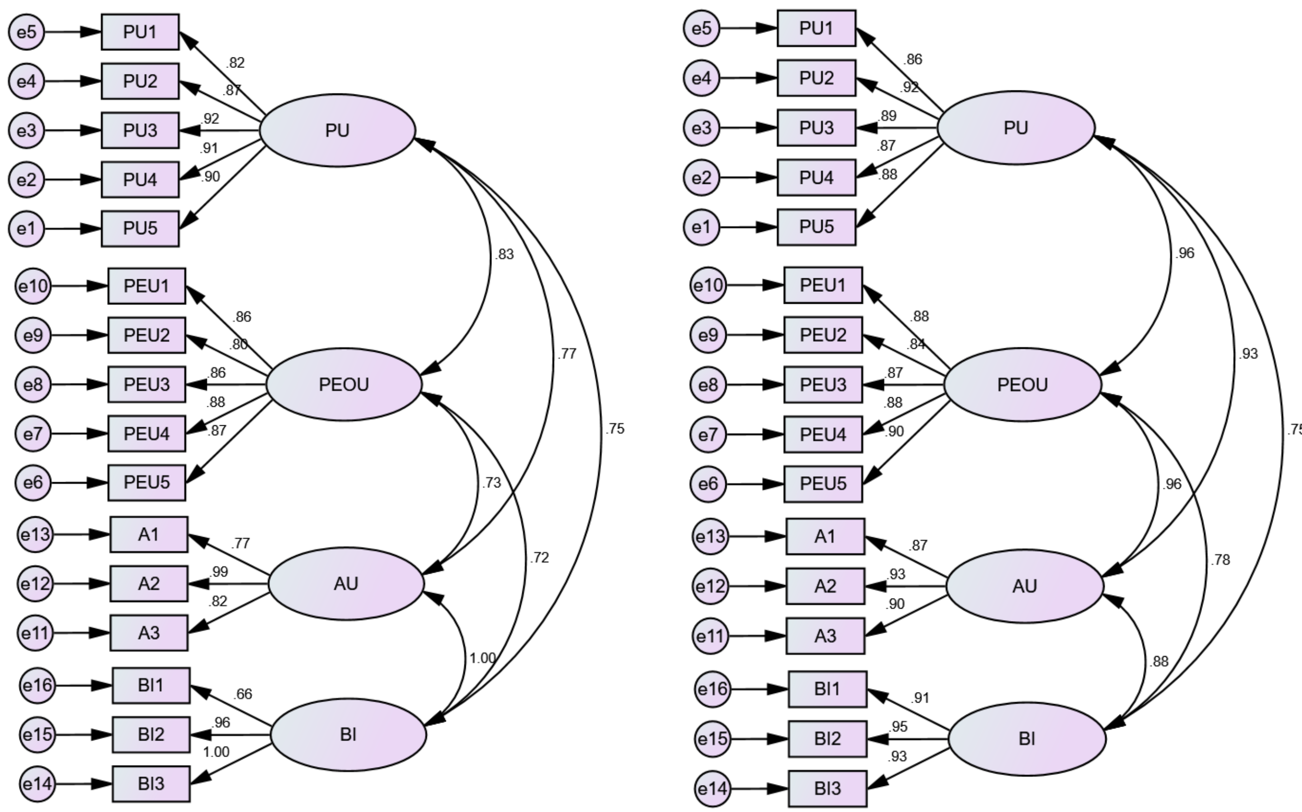
Results

Assessment of the Model

The study conducted a comprehensive examination of all constructs across both samples to evaluate the consistency of the research model, as outlined by Hair and Alamer (2022). Critical measures were used to assess the model: goodness of fit and reliability and validity. The initial model demonstrated a remarkable degree of consistency with the collected data, and all items showed strong factor loadings across both Chinese and international students' samples, as depicted in Figure 2. Notably, each obtained value exceeded the recommended threshold of 0.6, in line with the guidelines provided by Alavi et al. (2020).

Figure 2

Confirmatory Factor Analysis Model of Both Samples



(a) Chinese students' sample (b) International students' sample

Note. PU = perceived usefulness; PEOU = perceived ease of use; AU = attitude toward use; BI = behavioral intention to use.

Table 2 presents a comprehensive assessment of the model fit indices, which includes chi-square (χ^2/df), Root Mean Square Residual (RMSR), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), and Tucker-Lewis Index (TLI). All seven indicators were statistically significant across both samples and fell within the recommended range, consistent with the criteria outlined by Sarstedt et al. (2021). The consistency observed across these various indices strongly supports the reliability and validity of the model for both samples, affirming its appropriateness for the study's intended research objectives.

Table 2

Fit Indices Summary

Fitting indices	Recommended values	Model values for Chinese sample	Model values for international sample
χ^2/df	< 3	2.51	2.62
RMSR	< 0.08	0.071	0.75
RMSEA	< 0.08	0.067	0.072
CFI	> 0.90	0.92	0.94
GFI	> 0.90	0.93	0.94
AGFI	> 0.8	0.83	0.86
TLI	> 0.90	0.91	0.93

Note. RMSR = Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; GFI = Goodness of Fit Index; AGFI = Adjusted Goodness of Fit Index.

To assess the consistency of the research measurement model across both Chinese and international students' samples, reliability and validity assessments were conducted. Cronbach's alpha was used to ensure the reliability of the variables, and all constructs achieved a satisfactory level exceeding 70% across both samples, in line with the guidelines by Taber (2018) (Table 3).

Validity was evaluated using both convergent and discriminant validation tests. A convergent validity test was conducted to determine the degree of alignment among all instrument constructs from two distinct perspectives: composite reliability (CR) and average variance extracted (AVE). All variables across both samples were validated, as the obtained values exceeded the predefined thresholds, specifically, CR > .70 and AVE > .50 (Lai, 2021). In parallel, a discriminant validity test was conducted to explore the distinctions among the overlapping constructs. The acquired values for each construct across both samples were found to be consistent with the established threshold values (> .7) (Hair and Alamer, 2022). All variables exhibited significant correlations with each other. In the Chinese sample, correlation coefficients ranged from .76 to .82, and in the international students' sample, they ranged from .79 to .83, demonstrating high correlation.

Table 3

Reliability and Validity Matrix of Both Samples

S.No	Variables	Chinese sample							International sample						
		α	CR	AVE	1	2	3	4	α	CR	AVE	1	2	3	4
1	PU	.79	.90	.78	**.88**				.77	.90	.76	**.87**			
2	PEOU	.78	.90	.79	.78*	**.88**			.78	.91	.77	.83*	**.87**		
3	AU	.80	.91	.79	.80*	.76*	**.88**		.79	.92	.79	.85*	.87*	**.88**	
4	BI	.81	.90	.80	.78*	.77*	.82*	**.89**	.80	.90	.75	.70*	.73*	.80*	**.86**

Note. Bold values reflect discriminant validity. CR = composite reliability; AVE = average variance extracted; PU = perceived usefulness; PEOU = perceived ease of use; AU = attitude toward use; BI = behavioral intention to use.

**p* < .01.

Descriptive Analysis

Table 4 and Figure 3 illustrate the descriptive results for both samples. Chinese students' reported mean values were 3.46 for PU, 3.23 for PEOU, 3.38 for AU, and 3.41 for BI. Meanwhile, international students indicated slightly higher mean values across all constructs: 3.72 for PU, 3.58 for PEOU, 3.64 for AU, and 3.69 for BI. To assess the normality of the data distribution, both skewness and kurtosis were examined using descriptive statistics. All the obtained values were within the determined range, with skewness between −3 and +3 and kurtosis between −10 and +10 (Matore & Khairani, 2020). Following established criteria, especially when using SEM, ensures the appropriateness of the data distribution (Demir, 2022).

Table 4

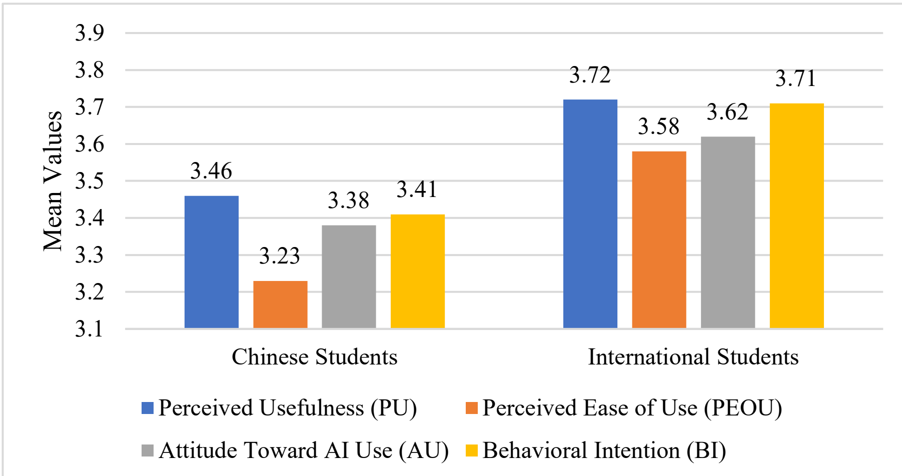
Descriptive Statistics

Participants	Chinese students				International students			
Variables	PU	PEOU	AU	BI	PU	PEOU	AU	BI
M	3.46	3.23	3.38	3.41	3.72	3.58	3.62	3.71
SD	0.81	0.82	0.81	0.83	0.96	0.93	0.99	0.84
Skewness	−0.27	0.10	−0.71	−0.82	−0.97	−0.71	−0.83	−0.87
Kurtosis	1.11	0.89	0.97	0.92	0.87	0.57	0.41	0.67

Note. CR = composite reliability; AVE = average variance extracted; PU = perceived usefulness; PEOU = perceived ease of use; AU = attitude toward use; BI = behavioral intention to use.

Figure 3

Comparison of Mean Values of Constructs Across Chinese and International Students



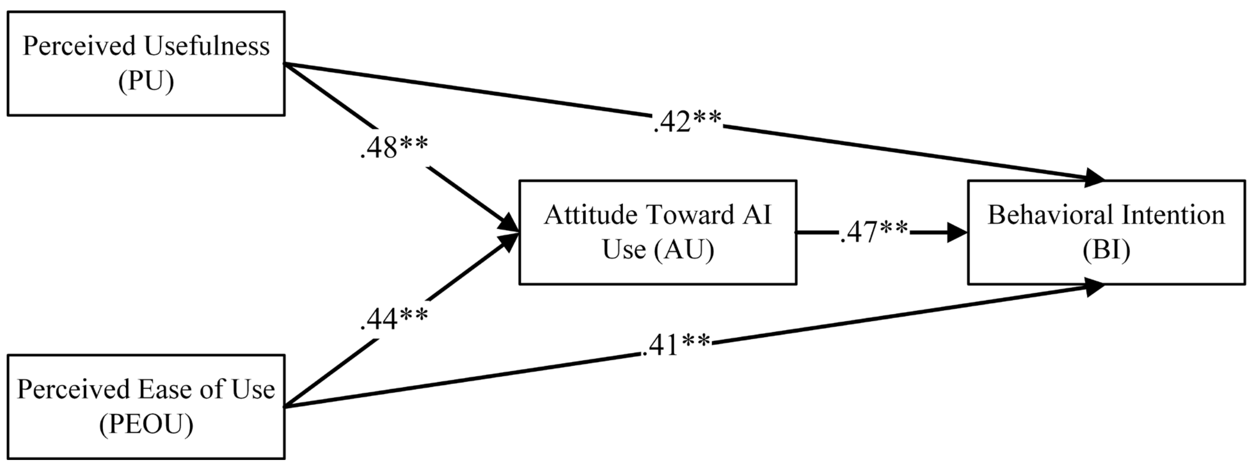
Structural Model Assessment

After ensuring the reliability and validity of the constructs, the study employed the structural model approach to evaluate relationships between the outlined hypotheses (see Figure 4). The results indicate that PU and PEOU significantly impact both Chinese and international students' attitudes and behavioral intentions toward AI use. In the Chinese sample, the influence of PU on attitude was .48, and the influence on their behavioral intentions was .42, supporting H1 and H2 respectively. In the international students' sample, the influence of PU on their attitude was .39, and the influence on their behavioral intentions was .44, supporting H1 and H2 respectively. In the Chinese sample, the influence of PEOU on attitude was .44, and the influence on their behavioral intentions was .41, supporting H3 and H4 respectively. . In the international students sample, the influence of PEOU on their attitude was .54, and the influence on their behavioral intentions was .43, respectively supporting H3 and H4. Results also show that attitude toward AI use significantly influenced the behavioral intentions of both Chinese (.47) and international students (.45) at a significance level of .01, providing support for H5.

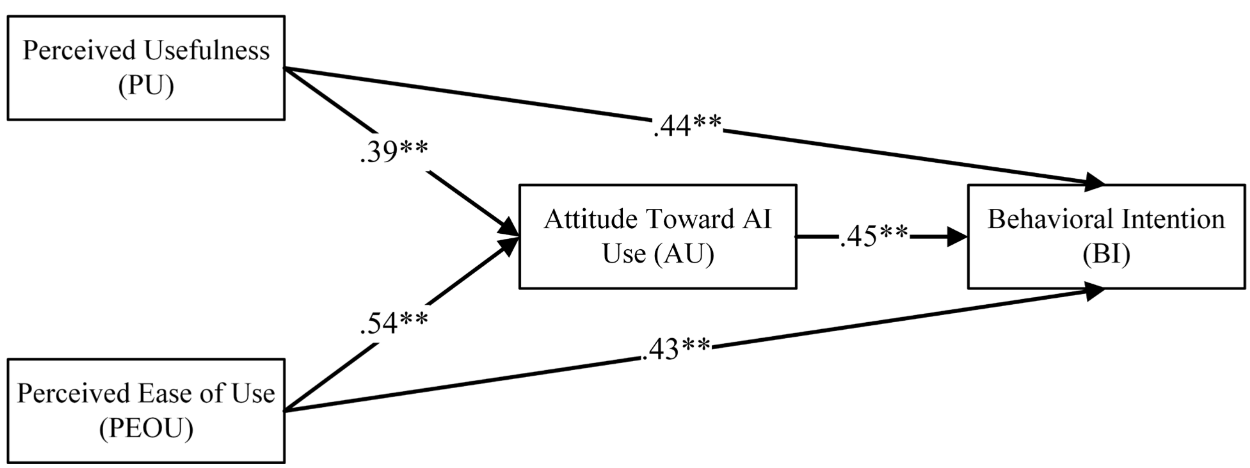
After confirming the reliability and validity of the constructs, the study used the structural model approach (see Figure 4) to explore the relationships outlined in the hypotheses. The results revealed significant insights into the impact of PU and PEOU on the AU and BI of Chinese and international students concerning use of AI.

Figure 4

Structural Model Assessment



(a) Chinese students' sample



(b) International students' sample

**$p < .01$.

In the Chinese student sample, the study found a substantial influence of PU on AU (.48) and BI (.42), providing robust support for H1 and H2. Correspondingly, the group of international students demonstrated a similar pattern, whereby PU had a substantial impact on both AU (.39) and BI (.44), thereby confirming the veracity of H1 and H2. Exploring the robustness of H3 and H4, Chinese students demonstrated a significant impact of PEOU on AU (.44) and BI (.41). A noticeable effect was shown for international participants, on the other hand, with PEOU significantly shaping both AU (.54) and BI (.43), robustly supporting H3 and H4 in this unique cultural setting.

Finally, the study examined the interplay between attitude and behavioral intentions. The results indicated that both Chinese and international students' attitudes toward AI significantly impacted their behavioral intentions (.47 and .45, respectively) at a high significance level (.01), providing robust support for H5. These findings offer a multifaceted understanding of how students from different cultural backgrounds adopt and use AI in their educational practices.

Discussion

In the current educational landscape, AI integration has become a pivotal determinant of students' learning experiences. The success of implementing AI in higher education largely hinges on students' attitudes and perceptions (Al-Adwan et al., 2022; Linardatos & Apostolou, 2023). As new technologies continue to emerge, it becomes crucial to investigate individuals' attitudes toward using them and analyze the factors that shape these perceptions (Akram & Abdelrady, 2023). Additionally, diverse cultural backgrounds may lead to varying preferences regarding technology use in education (Savicki, 2023). Consequently, individuals may hold complex attitudes toward the adoption of technology. Therefore, it is imperative to comprehend these dynamics.

In the context of China, the landscape of higher education has undergone a significant transformation in recent decades due to a substantial increase in the number of international students. This transformative shift motivated the study to provide empirical evidence regarding students' attitudes and behavioral intentions toward AI use across both Chinese and international contexts using the technology acceptance model as the guiding framework. This study not only revealed the intricate dynamics but also highlighted the reliability of the theoretical framework within both Chinese and international contexts, establishing a solid foundation for potential applications in diverse cultural settings. The insights obtained from this study can inform the development of educational interventions to help foster a more enriching educational experience that is tailored to the diverse needs of students in our technology-driven era.

Regarding the prevalence of AI usage, the identification of a notable discrepancy between Chinese and international students reveals an intriguing aspect of the technological landscape in higher education. This coincides with the consensus of Kim and Lee (2023), who highlighted the significant role of sociocultural factors on students' attitudes toward AI. The observed prevalence of, attitudes toward, and behavioral intentions to using AI use among international students reflect the global trend in which students from diverse cultural backgrounds are more likely to employ the latest technologies in their learning (T. Wang et al., 2023). One reason may be that they come from different academic backgrounds and want to have globalized educational experiences (Xiong et al., 2022). International students may perceive AI adoption as essential for staying competitive in a globally connected job market, while Chinese students might see it as an opportunity to align with global educational standards.

Language proficiency also plays a key role in shaping students' interactions with technology. Students with more advanced language skills can better understand instructions and navigate user interfaces, enhancing their overall experience with AI-driven tools (Lee, 2022). Additionally, the fact that technological tools are generally rendered in English makes them easier for international students to use, since they usually possess better English proficiency (Huang et al., 2023). Conversely, Chinese students may have different perspectives on the use of AI in higher education because they may not be as fluent in English and may thus encounter more obstacles when trying to make full use of AI technologies (Galloway & Ruegg, 2020). This linguistic component adds another degree of intricacy to AI technology use, emphasizing the necessity for a detailed comprehension of language-related aspects of technology acceptance.

The observation regarding the significantly elevated levels of both key elements—PU and PEOU—among international students compared with Chinese students marks a noteworthy trend with implications rooted in prior research. Existing literature on technology adoption suggests that positive perceptions of usefulness and ease of use are fundamental predictors of technology acceptance (Park & Kim, 2023). These positive perceptions often lead to increased intentions to use technology and foster a more favorable disposition toward its integration into daily practices (Elfeky & Elbyaly, 2023). Prior studies have also emphasized the importance of positive perceptions in driving technology adoption and use (e.g., Al-Hattami, 2023). For both international and Chinese students to optimize PU and PEOU, it is essential to prioritize the enhancement of user interfaces through a multifaceted approach. Concerned authorities should initiate the process by conducting comprehensive user experience assessments, gathering feedback from diverse student groups to inform targeted improvements. Design features that match international and Chinese students' cultural preferences and educational backgrounds should be implemented, guaranteeing simple access and adaptable design for device accessibility.

The findings reveal a compelling narrative on the intricate relationships between PU and PEOU on AU and BI. The significant role of PU in shaping students' attitudes and intentions toward AI use is consistent with prior studies (Anthony et al., 2022). The strong effect of PU on both AU and BI for Chinese students aligns with studies emphasizing the importance of perceived usefulness in technology adoption (Xiao & Goulias, 2022). Similarly, international students also exhibited a high level of PU, with a few minor differences, indicating that PU of AI is a global driver that transcends cultures.

This study further examined the impact of PEOU on students' attitudes and intentions toward AI use. The findings underline the significance of user-friendly interactions, consistent with prior research (Granić, 2022). We observed a stronger effect of PEOU on AU and BI among international students compared with Chinese students. This difference can be attributed to cultural dissimilarities, which affect how people perceive technology and its ease of use. International students' familiarity with diverse technological tools in their home countries may make AI easier for them to use (Sutrisno & Lubis, 2022). Additionally, different levels of previous technology exposure may also contribute to this difference: international students might have more experience with modern technologies than Chinese students (Steyn & Gunter, 2023).

The intricate connection between attitude and behavioral intentions, as examined in H5, reveals a recurring pattern in existing research. The results of earlier research are reflected in the data, with a significant correlation between both Chinese and international students' attitude toward interacting with AI and their behavioral intentions to use it (Papakostas et al., 2023). This provides evidence for the notion that cultivating a positive attitude is not only a precondition but also a strong indicator of future interactions with AI technology.

Recognizing the influence of students' cultural backgrounds, institutions should design interventions that cater to the unique perspectives of both Chinese and international students, fostering a shared understanding of AI technology. Providing language-specific support and user-friendly interfaces can bridge potential language barriers, ensuring that international students, in particular, feel comfortable and empowered in interacting with AI. Moreover, incorporating cross-cultural workshops and technology training programs tailored to people with varying levels of technological exposure can further enhance positive perceptions and ease of use, contributing to a more inclusive and globally aware educational environment.

Conclusions

This study provides empirical evidence regarding the attitudes and behavioral intentions toward the use of AI among Chinese and international students in higher education. Results shed light on the complex dynamics among PU, PEOU, AU, and BI in a cross-cultural setting in a Chinese university. The findings emphasize the pivotal role of PU and PEOU in influencing both AU and BI of students toward AI use. Moreover, the study revealed notable disparities, particularly in the influence of PEOU on AU and BI. International students exhibited a stronger effect of PEOU on both AU and BI than their Chinese counterparts. These findings underscore the intricate interplay of cultural backgrounds, prior technological exposure, and language considerations in shaping perceptions of AI technology. Tailoring educational strategies to regulate the number of cultural perspectives, providing language-specific support, and ensuring user-friendly interfaces are crucial for fostering positive attitudes and intentions toward using AI, especially among international students. The insights gleaned from this research contribute to a nuanced understanding of the factors influencing technology acceptance in a cross-cultural higher education context, offering practical implications for educators and institutions seeking to optimize AI integration in their pedagogical approaches.

Declarations

Ethics Approval and Consent to Participate

Informed consent to participate was obtained from all individual participants included in the study. The responses were collected after getting approval from the Institutional Review Board in North China University of Water Resources and Electric Power.

Consent for Publication

Informed consent for publication was obtained from all individual participants included in the study.

Availability of Data and Materials

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Competing Interests

The authors declare that they have no competing interests.

Funding

This study is supported by the Education and Teaching Research Project of NCWU (Grant No. 2024XJGXM057), the Education and Teaching Research Project of Henan Province, China (Grant No. 2022SYJXLX051), and the Education and Teaching Research Project of the International Cooperation and Exchange Division, Ministry of Education, China. (Grant No. (2023156).

References

Abdelrady, A. H., & Akram, H. (2022). An empirical study of ClassPoint tool application in enhancing EFL students' online learning satisfaction. *Systems*, *10*(5), Article 154. https://doi.org/10.3390/systems10050154

Akinoso, S. O. (2023). Motivation and ICT in secondary school mathematics using unified theory of acceptance and use of technology model. *Indonesian Journal of Educational Research and Technology*, *3*(1), 79–90. https://doi.org/10.17509/ijert.v3i1.47183

Akram, H., & Abdelrady, A. H. (2023). Application of ClassPoint tool in reducing EFL learners' test anxiety: An empirical evidence from Saudi Arabia. *Journal of Computers in Education*, *10*, 529–547. https://doi.org/10.1007/s40692-023-00265-z

Akram, H., Abdelrady, A. H., Al-Adwan, A. S., & Ramzan, M. (2022). Teachers' perceptions of technology integration in teaching-learning practices: A systematic review. *Frontiers in Psychology*, *13*, Article 920317. https://doi.org/10.1007/s40692-023-00265-z

Akram, H., Aslam, S., Saleem, A., & Parveen, K. (2021). The challenges of online teaching in COVID-19 pandemic: A case study of public universities in Karachi, Pakistan. *Journal of Information Technology Education. Research*, *20*, 263–282. https://doi.org/10.28945/4784

Akram, H., Kamran, M., & Ahmad, N. (2020). An examination of the encountered challenges of Pakistani international students in China: A case study of first-year students. *Pakistan Journal of Social Sciences*, *40*(4), 1567–1576. https://pjss.bzu.edu.pk/index.php/pjss/article/view/959

Akram, H., & Yang, Y. (2021). A critical analysis of the weak implementation causes on educational policies in Pakistan. *International Journal of Humanities and Innovation (IJHI)*, *4*(1), 25-28. https://doi.org/10.33750/ijhi.v4i1.104

Akram, H., Yingxiu, Y., Al-Adwan, A. S., & Alkhalifah, A. (2021). Technology integration in higher education during COVID-19: An assessment of online teaching competencies through technological pedagogical content knowledge model. *Frontiers in Psychology*, *12*, Article 736522. https://doi.org/10.3389/fpsyg.2021.736522

Akram, H., Yingxiu, Y., Aslam, S., & Umar, M. (2021b, June). Analysis of synchronous and asynchronous approaches in students' online learning satisfaction during Covid-19 pandemic. In *2021 IEEE International Conference on Educational Technology (ICET)* (pp. 203-207). IEEE. https://doi.org/10.1109/ICET52293.2021.9563183

Al-Adwan, A. S., Nofal, M., Akram, H., Albelbisi, N. A., & Al-Okaily, M. (2022). Towards a sustainable adoption of e-learning systems: The role of self-directed learning. *Journal of Information Technology Education: Research*, *21*, 245–267. https://doi.org/10.28945/4980

Alam, A. (2021, November 26–27). Possibilities and apprehensions in the landscape of artificial intelligence in education. In *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)* (pp. 7–14). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ICCICA52458.2021.9697272

Alavi, M., Visentin, D. C., Thapa, D. K., Hunt, G. E., Watson, R., & Cleary, M. (2020). Chi-square for model fit in confirmatory factor analysis. *Journal of Advanced Nursing*, *76*(9), 2209–2211. https://doi.org/10.1111/jan.14399

Alfadda, H. A., & Mahdi, H. S. (2021). Measuring students' use of Zoom application in language course based on the technology acceptance model (TAM). *Journal of Psycholinguistic Research*, *50*(4), 883–900. https://doi.org/10.1007/s10936-020-09752-1

Al-Hattami, H. M. (2023). Understanding perceptions of academics toward technology acceptance in accounting education. *Heliyon*, *9*(1), Article E13141. https://doi.org/10.1016/j.heliyon.2023.e13141

Almaiah, M. A., Hajjej, F., Lutfi, A., Al-Khasawneh, A., Shehab, R., Al-Otaibi, S., & Alrawad, M. (2022). Explaining the factors affecting students' attitudes to using online learning (Madrasati platform) during COVID-19. *Electronics*, *11*(7), Article 973. https://doi.org/10.3390/electronics11070973

Al-Mamary, Y. H. S. (2022). Why do students adopt and use learning management systems? Insights from Saudi Arabia. *International Journal of Information Management Data Insights*, *2*(2), Article 100088. https://doi.org/10.1016/j.jjimei.2022.100088

Anthony, B., Kamaludin, A., Romli, A., Raffei, A. F. M., Phon, D. N. A. E., Abdullah, A., & Ming, G. L. (2022). Blended learning adoption and implementation in higher education: A theoretical and systematic review. *Technology, Knowledge and Learning*, *27*, 531–578. https://doi.org/10.1007/s10758-020-09477-z

Barrett, A., Pack, A., Guo, Y., & Wang, N. (2023). Technology acceptance model and multi-user virtual reality learning environments for Chinese language education. *Interactive Learning Environments*, *31*(3), 1665–1682. https://doi.org/10.1080/10494820.2020.1855209

Breitsohl, H. (2019). Beyond ANOVA: An introduction to structural equation models for experimental designs. *Organizational Research Methods*, *22*(3), 649–677. https://doi.org/10.1177/1094428118754988

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, *20*(1), Article 38. https://doi.org/10.1186/s41239-023-00408-3

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *1*, Article 100002. https://doi.org/10.1016/j.caeai.2020.100002

Cheng, E. C., & Wang, T. (2022). Institutional strategies for cybersecurity in higher education institutions. *Information*, *13*(4), Article 192. https://doi.org/10.3390/info13040192

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, *20*(1), Article 22. https://doi.org/10.1186/s41239-023-00392-8

Dai, K., & Hardy, I. (2023). Pursuing doctoral research in an emerging knowledge hub: An exploration of international students' experiences in China. *Studies in Higher Education*, *48*(6), 784–796. https://doi.org/10.1080/03075079.2023.2166917

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Demir, S. (2022). Comparison of normality tests in terms of sample sizes under different skewness and Kurtosis coefficients. *International Journal of Assessment Tools in Education*, *9*(2), 397–409. https://doi.org/10.21449/ijate.1101295

Elfeky, A. I. M., & Elbyaly, M. Y. H. (2023). The use of data analytics technique in learning management system to develop fashion design skills and technology acceptance. *Interactive Learning Environments*, *31*(6), 3810–3827. https://doi.org/10.1080/10494820.2021.1943688

Emerson, R. W. (2021). Convenience sampling revisited: Embracing its limitations through thoughtful study design. *Journal of Visual Impairment & Blindness*, *115*(1), 76–77. https://doi.org/10.1177/0145482X20987707

Epstein, J., Santo, R. M., & Guillemin, F. (2015). A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *Journal of Clinical Epidemiology*, *68*(4), 435–441. https://doi.org/10.1016/j.jclinepi.2014.11.021

Ertel, W. (2018). *Introduction to artificial intelligence*. Springer. https://doi.org/10.1007/978-3-319-58487-4

Galloway, N., & Ruegg, R. (2020). The provision of student support on English medium instruction programmes in Japan and China. *Journal of English for Academic Purposes*, *45*, Article 100846. https://doi.org/10.1016/j.jeap.2020.100846

Granić, A. (2022). Educational technology adoption: A systematic review. *Education and Information Technologies*, *27*(7), 9725–9744. https://doi.org/10.1007/s10639-022-10951-7

Guo, Y., Wang, Y., & Ortega-Martín, J. (2023). The impact of blended learning-based scaffolding techniques on learners' self-efficacy and willingness to communicate. *Porta Linguarum*, *40*(2), 253–273. https://doi.org/10.30827/portalin.vi40.27061

Hair, J., & Alamer, A. (2022). Partial least squares structural equation modeling (PLS-SEM) in second language and education research: Guidelines using an applied example. *Research Methods in Applied Linguistics*, *1*(3), Article 100027. https://doi.org/10.1016/j.rmal.2022.100027

Huang, F., Teo, T., & Zhao, X. (2023). Examining factors influencing Chinese ethnic minority English teachers' technology adoption: An extension of the UTAUT model. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2023.2239304

Jiani, M. A. (2017). Why and how international students choose mainland China as a higher education study abroad destination. *Higher Education*, *74*, 563–579. https://doi.org/10.1007/s10734-016-0066-0

Kayalar, F. (2016). Cross-cultural comparison of teachers' views upon integration and use of technology in classroom. *Turkish Online Journal of Educational Technology-TOJET*, *15*(2), 11–19. https://eric.ed.gov/?id=EJ1096412

Kim, S. W., & Lee, Y. (2023). Investigation into the influence of socio-cultural factors on attitudes toward artificial intelligence. *Education and Information Technologies*. Advance online publication. https://doi.org/10.1007/s10639-023-12172-y

Lai, M. H. (2021). Composite reliability of multilevel data: It's about observed scores and construct meanings. *Psychological Methods*, *26*(1), 90–102. https://doi.org/10.1037/met0000287

Lasfeto, D. B., & Ulfa, S. (2023). Modeling of online learning strategies based on fuzzy expert systems and self-directed learning readiness: The effect on learning outcomes. *Journal of Educational Computing Research*, *60*(8), 2081–2104. https://doi.org/10.1177/07356331221094249

Lee, C. (2022). Intention to use versus actual adoption of technology by university English language learners: What perceptions and factors matter? *Computer Assisted Language Learning*, *35*(8), 2049–2077. https://doi.org/10.1080/09588221.2020.1857410

Lewis, J. R. (2019). Comparison of four TAM item formats: Effect of response option labels and order. *Journal of User Experience*, *14*(4), 224–236. https://uxpajournal.org/tam-formats-effect-response-labels-order/

Linardatos, G., & Apostolou, D. (2023). Investigating high school students' perception about digital comics creation in the classroom. *Education and Information Technologies*, *28*, 10079–10101. https://doi.org/10.1007/s10639-023-11581-3

Makumane, M. A. (2023). Students' perceptions on the use of LMS at a Lesotho university amidst the COVID-19 pandemic. *African Identities*, *21*(2), 209–226. https://doi.org/10.1080/14725843.2021.1898930

Matore, E. M., & Khairani, A. Z. (2020). The pattern of skewness and kurtosis using mean score and logit in measuring adversity quotient (AQ) for normality testing. *International Journal of Future Generation Communication and Networking*, *13*(1), 688–702. http://sersc.org/journals/index.php/IJFGCN/article/view/6571

Matsika, C., & Zhou, M. (2021). Factors affecting the adoption and use of AVR technology in higher and tertiary education. *Technology in Society*, *67*, Article 101694. https://doi.org/10.1016/j.techsoc.2021.101694

Min, Y., Yang, M., Huang, J., & Duangekanong, S. (2023). Influencing factors of behavior intention of Master of Arts students towards online education in Chengdu public universities, China. *Scholar: Human Sciences*, *15*(1), 1–10. https://doi.org/10.14456/shserj.2023.1

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, *2*, Article 100020. https://doi.org/10.1016/j.caeai.2021.100020

Papakostas, C., Troussas, C., Krouska, A., & Sgouropoulou, C. (2023). Exploring users' behavioral intention to adopt mobile augmented reality in education through an extended technology acceptance model. *International Journal of Human–Computer Interaction*, *39*(6), 1294–1302. https://doi.org/10.1080/10447318.2022.2062551

Park, D. Y., & Kim, H. (2023). Determinants of intentions to use digital mental healthcare content among university students, faculty, and staff: Motivation, perceived usefulness, perceived ease of use, and parasocial interaction with AI chatbot. *Sustainability*, *15*(1), Article 872. https://doi.org/10.3390/su15010872

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2021). Partial least squares structural equation modeling. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), *Handbook of market research* (pp. 587–632). Springer International Publishing. https://doi.org/10.1007/978-3-319-57413-4_15

Sathye, S., Prasad, B., Sharma, D., Sharma, P., & Sathye, M. (2018). Factors influencing the intention to use of mobile value-added services by women-owned microenterprises in F iji. *The Electronic Journal of Information Systems in Developing Countries*, *84*(2), e12016. https://doi.org/10.1002/isd2.12016

Savicki, V. (Ed.). (2023). *Developing intercultural competence and transformation: Theory, research, and application in international education*. Routledge. https://doi.org/10.4324/9781003444169

Steyn, C., & Gunter, A. (2023). When an international student stays at home: Defining an international student in distance education. *Journal of Geography in Higher Education*, *47*(1), 56–70. https://doi.org/10.1080/03098265.2021.1991289

Sutrisno, S., & Lubis, A. H. (2022). A survey on online computer practicum during the COVID-19 pandemic: Students' perceptions. *International Journal of Research and Review*, *9*(9), 372–379. https://doi.org/10.52403/ijrr.20220942

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*, 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Taherdoost, H. (2016). Validity and reliability of the research instrument; How to test the validation of a questionnaire/survey in a research. *International Journal of Academic Research in Management (IJARM)*, *5*(3), 28–36. https://doi.org/10.2139/ssrn.3205040

Teo, T. (2009). The impact of subjective norm and facilitating conditions on pre-service teachers' attitude toward computer use: A structural equation modeling of an extended technology acceptance model. *Journal of Educational Computing Research*, *40*(1), 89–109. https://doi.org/10.2190/EC.40.1.d

Vargo, D., Zhu, L., Benwell, B., & Yan, Z. (2021). Digital technology use during COVID-19 pandemic: A rapid review. *Human Behavior and Emerging Technologies*, *3*(1), 13–24. https://doi.org/10.1002/hbe2.242

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Wang, J., & Wang, X. (2019). *Structural equation modeling: Applications using Mplus*. John Wiley & Sons. https://doi.org/10.1002/9781119422730

Wang, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023). Exploring the potential impact of artificial intelligence (AI) on international students in higher education: Generative AI, chatbots, analytics, and international student success. *Applied Sciences, 13*(11), Article 6716. https://doi.org/10.3390/app13116716

Wang, Y. (2023). Probing into the boredom of online instruction among Chinese English language teachers during the COVID-19 pandemic. *Current Psychology*, *43*, 12144–12158. https://doi.org/10.1007/s12144-022-04223-3

Wang, Y. A., & Rhemtulla, M. (2021). Power analysis for parameter estimation in structural equation modeling: A discussion and tutorial. *Advances in Methods and Practices in Psychological Science*, *4*(1). https://doi.org/10.1177/2515245920918253

Xiao, J., & Goulias, K. G. (2022). Perceived usefulness and intentions to adopt autonomous vehicles. *Transportation Research Part A: Policy and Practice*, *161*, 170–185. https://doi.org/10.1016/j.tra.2022.05.007

Xiong, W., Yang, J., & Shen, W. (2022). Higher education reform in China: A comprehensive review of policymaking, implementation, and outcomes since 1978. *China Economic Review*, *72*, Article 101752. https://doi.org/10.1016/j.chieco.2022.101752

Yang, F., Ren, L., & Gu, C. (2022). A study of college students' intention to use metaverse technology for basketball learning based on UTAUT2. *Heliyon*, *8*(9), Article E10562. https://doi.org/10.1016/j.heliyon.2022.e10562

Yilmaz, R., & Yilmaz, F. G. K. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, *4*, Article 100147. https://doi.org/10.1016/j.caeai.2023.100147

Zhang, Y., & Zhu, Y. (2022). Effects of educational robotics on the creativity and problem-solving skills of K-12 students: A meta-analysis. *Educational Studies*. Advance online publication. https://doi.org/10.1080/03055698.2022.2107873

Zhou, C. (2023). Integration of modern technologies in higher education on the example of artificial intelligence use. *Education and Information Technologies*, *28*(4), 3893–3910. https://doi.org/10.1007/s10639-022-11309-9

Zulaiha, D., & Triana, Y. (2023). Students' perception toward the use of open educational resources to improve writing skills. *Studies in English Language and Education*, *10*(1), 174–196. https://doi.org/10.24815/siele.v10i1.25797

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

"To Use or Not to Use?" A Mixed-Methods Study on the Determinants of EFL College Learners' Behavioral Intention to Use AI in the Distributed Learning Context

Hanwei Wu[1], Yunsong Wang[2], and Yongliang Wang[3,*]
[1]School of Foreign Studies, Hunan Normal University, Changsha, China; **[2]***School of Foreign Languages and Cultures, Nanjing Normal University, Nanjing, China***;** **[3]***School of Foreign Studies, North China University of Water Resources and Electric Power, Zhengzhou, China; *corresponding author*

Abstract

Artificial intelligence (AI) offers new possibilities for English as a foreign language (EFL) learners to enhance their learning outcomes, provided that they have access to AI applications. However, little is written about the factors that influence their intention to use AI in distributed EFL learning contexts. This mixed-methods study, based on the technology acceptance model (TAM), examined the determinants of behavioral intention to use AI among 464 Chinese EFL college learners. As to quantitative data, a structural equation modelling (SEM) approach using IBM SPSS Amos (Version 24) produced some important findings. First, it was revealed that perceived ease of use significantly and positively predicts perceived usefulness and attitude toward AI. Second, attitude toward AI significantly and positively predicts behavioral intention to use AI. However, contrary to the TAM assumptions, perceived usefulness does not significantly predict either attitude toward AI or behavioral intention to use AI. Third, mediation analyses suggest that perceived ease of use has a significant and positive impact on students' behavioral intention to use AI through their attitude toward AI, rather than through perceived usefulness. As to qualitative data, semi-structured interviews with 15 learners, analyzed by the software MAXQDA 2022, provide a nuanced understanding of the statistical patterns. This study also discusses the theoretical and pedagogical implications and suggests directions for future research.

Keywords: artificial intelligence, AI, EFL college learner, behavioral intention, distributed learning

“To Use or Not to Use?”: A Mixed-methods Study on the Determinants of EFL College Learners’ Behavioral Intention to Use AI in the Distributed Learning Context

AI uses computational algorithms to perform cognitive tasks or solve complex problems that normally require human intelligence (Chen et al., 2020). In the last 30 years, AI has become a powerful tool to create new paradigms in many domains, including EFL education (Jiang, 2022; Rezai, 2023). AI techniques, such as natural language processing and machine learning, have enabled various applications for language learning, such as ChatGPT (Kohnke et al., 2023a), Pigai (Yang et al., 2023), Duolingo (Shortt et al., 2023), and so forth. These applications provide EFL learners with adaptive materials, instant feedback, and automatic diagnosis (Chassignol et al., 2018; Zhang & Aslan, 2021). Moreover, they allow EFL learners to access academic resources at their convenience, beyond the limits of traditional classes. These benefits have made AI-assisted EFL learning popular in the distributed learning context where flexibility, accessibility, and cost-effectiveness are highlighted (Janbi et al., 2023; Kuddus, 2022; Namaziandost, Razmi, et al., 2021).

Despite the advancement of AI computing technologies, effective learning is not guaranteed by the mere use of AI tools (Castañeda & Selwyn, 2018; Selwyn, 2016). For AI-assisted language learning, EFL learners’ willingness to adopt AI applications is a key factor (Kelly et al., 2023). Nevertheless, the existing literature on this topic tends to overlook EFL learners and concentrate on either the outcomes of a single AI system (Chen & Pan, 2022; Dizon, 2020; Muftah et al., 2023) or the attitudes of EFL teachers (Kohnke et al., 2023b; Ulla et al., 2023). The TAM (Davis, 1989) is a widely used framework that investigates how people’s behavior is affected by their perceptions of technology, encompassing perceived usefulness, perceived ease of use, and attitude toward technology. As AI is an emerging technology, it is suitable for applying the TAM. Therefore, the TAM can provide insights into the factors that shape EFL learners’ intention to use AI in the distributed learning context.

AI applications have the potential to facilitate EFL learning for college students in China, who have more access to electronic devices than K–12 students (Gao et al., 2014). Drawing on the TAM, this mixed-methods study investigated the factors influencing college students’ intention to adopt AI applications for EFL learning. Our findings have implications for various stakeholders. Educators can leverage these findings to enhance college students’ EFL achievement by promoting the use of AI applications, particularly in the distributed learning context where AI applications can provide flexible learning opportunities. AI application developers can use these findings to improve their design to cater to EFL learners and achieve business success.

Literature Review

Artificial Intelligence in EFL Contexts

AI has implications for education, as it can enhance outcomes, personalize instruction, and streamline activities (Bearman et al., 2023; Namaziandost, Hashemifardnia, et al., 2021; Ouyang & Jiao, 2021; Zhang & Aslan, 2021). In EFL contexts, AI applications that employ methods such as machine learning and natural language processing are prevalent. The majority of research has evaluated the effectiveness of AI in advancing EFL learners' writing and speaking skills (Chen & Pan, 2022; Klimova et al., 2023). Researchers have also explored the perceptions of EFL teachers and students on AI and the determinants of their AI adoption (An et al., 2022; Liu & Wang, 2024; Wang et al., 2023).

Effectiveness of AI Practices in EFL Contexts

Based on the role of learners, the extant literature on the effectiveness of AI practices in EFL contexts could be roughly divided into two categories: learner-as-beneficiary (LB) or learner-as-partner (LP). LB research investigates the impact of automatic essay scoring systems on EFL learners' writing skills (Chen & Pan, 2022; Muftah et al., 2023). These systems, such as Aim Writing and WRITER, root in behaviorist principles and provide feedback on spelling, grammar, and sentence structure. The conclusions of previous studies have been inconsistent. For example, Chen and Pan (2022) recommended a hybrid model of Automated Essay Scoring (AES) and instructor feedback, as they found that Aim Writing's feedback, although useful, was insufficient for students. However, in Muftah et al. (2023), the effectiveness of AI-assisted writing was confirmed. Their study indicated that WRITER users improved in all writing skills except organization, compared to traditional method users. This inconsistency may imply that different AES systems have different strengths and weaknesses.

LP research focuses on EFL learners' speaking skills, which follows a constructivist approach and engages EFL learners in dialogue with AI (Ayedoun et al., 2019; Dizon, 2020). This type of research employs technologies such as intelligent personal assistants and dialogue management systems. Previous studies have shown that AI systems can enhance learners' interactional competence. For example, Ayedoun et al. (2019) developed a model of an embodied conversational agent that employs communication strategies and affective backchannels to enhance EFL learners' willingness to communicate. The results showed the embodied conversational agent with both strategies and backchannels increased communication more than the agent with either one alone. Likewise, Dizon (2020) compared the speaking skills of Alexa users and non-users and found that the former group improved more than the latter. A possible explanation for this effectiveness is that AI practices for EFL speaking can motivate learners by providing natural dialogue and alleviate their anxiety by creating a safe environment. Unlike real humans, AI does not judge learners' oral mistakes.

Views of Stakeholders on AI Practices in EFL Contexts

Most studies have reported that EFL teachers have a positive and consistent view of AI and its benefits for their students' learning. For example, Kohnke et al. (2023b) conducted semi-structured interviews with 12 EFL teachers and found that they were familiar with and confident in using AI-driven teaching tools such as Siri and Alexa. Similarly, Ulla et al. (2023) interviewed 17 EFL teachers and explored their perspectives on ChatGPT in their teaching practices. The interviewees expressed positive attitudes toward ChatGPT and recognized its various applications, such as lesson preparation and language activity creation. Other studies have explored the psychological factors that affect EFL teachers' technology adoption. Zhi et al. (2023)

surveyed 214 EFL teachers and found that both emotional intelligence (EI) and self-efficacy were significant and positive predictors of technology adoption, with EI being a stronger predictor.

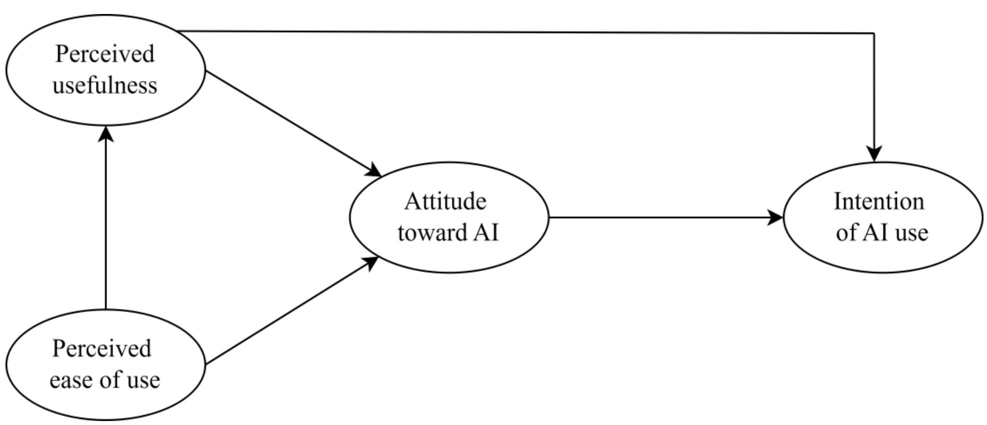
Some studies have examined the factors that influence EFL teachers’ behavioral intention to use AI (An et al., 2022; Liu & Wang, 2024). An et al. (2022) investigated 470 Chinese EFL teachers and found that performance expectancy, social influence, and AI language technological knowledge had direct effects on behavioral intention, while effort expectancy, facilitating conditions, and AI technological pedagogical knowledge had indirect effects. However, only a few studies have investigated the influencing factors of EFL learners’ intention to use AI (An et al., 2023; Liu & Ma, 2023). An et al. (2023) discovered that the intention to use AI among 452 Chinese students in secondary schools was determined by three common factors: their performance expectancy, their interest in AI culture, and their goal attainment expectancy. Despite the significance of their research, Chinese pre-college students could be less benefited from AI in distributed learning contexts, as they are banned from using electronic devices at school. Chinese college students, on the other hand, do not have such restrictions, so they can use AI tools to for EFL learning more easily.

The TAM as the Theoretical Basis

The TAM is a framework that explains users’ technology adoption (Davis, 1989; Venkatesh & Bala, 2008). The TAM suggests that users’ intention to use technology depends on three factors: perceived usefulness, perceived ease of use, and attitude toward use (Davis, 1989). Intention is the main outcome in the TAM, as it predicts actual usage (Davis, 1989; Teo et al., 2017). Perceived usefulness is how users believe technology will improve their performance, while perceived ease of use is how users believe technology will require little effort. Both perceived usefulness and perceived ease of use affect users’ attitude toward use, which is how users feel about using technology. Attitude toward use and perceived usefulness also influence users’ intention (Huang & Liaw, 2005). We used the TAM as the basis for our study to explore how EFL learners perceive and adopt AI for language learning in distributed learning contexts. See Figure 1.

Figure 1

The TAM-Based Model for EFL Learners’ Behavioral Intention to Use AI



Research Gaps and Hypotheses

Previous studies on AI in education have examined its impact on learning outcomes, teacher and student perceptions, and implementation and evaluation challenges. However, these studies mainly focused on EFL teaching contexts and ignored the factors affecting students' intention to use AI for EFL learning. Moreover, most previous studies targeted EFL pre-college learners with limited access to AI applications, leaving a gap in the knowledge of EFL college learners' views and behaviors. This study used the TAM as a theoretical framework to explore how EFL college learners' intention to use AI for language learning is shaped by their perceived ease of use, perceived usefulness, and attitude toward AI. Based on the TAM, we tested the following research hypotheses among EFL college learners in distributed learning contexts:

H1: Perceived ease of use positively predicts perceived usefulness and attitude toward AI.

H2: Perceived usefulness positively predicts attitude toward AI and behavioral intention to use AI.

H3: Attitude toward AI positively predicts behavioral intention to use AI.

H4: Perceived usefulness positively predicts behavioral intention via attitude toward AI.

H5: Perceived ease of use positively predicts behavioral intention to use AI via attitude toward AI.

H6: Perceived ease of use positively predicts behavioral intention to use AI via perceived usefulness and attitude toward AI.

H7: Perceived ease of use positively predicts behavioral intention to use AI via perceived usefulness.

Methodology

Participants

This study adopted an explanatory sequential design (Creswell, 2014) that combined quantitative and qualitative methods to investigate college students' use of AI for EFL learning. College students were chosen as the target population because they do not face the same restrictions as pre-college students, who are prohibited from using electronic devices at school, as discussed earlier. Five colleges in China were selected for the sampling, and 557 students from various majors participated in the survey. After removing 93 duplicate responses, the final sample size was 464, comprising 248 freshmen, 91 sophomores, 116 juniors, and 9 seniors. Most participants were female (74.14%, n = 344) and aged between 18 and 23 (M = 19.18, SD = 1.29). The qualitative data were collected through interviews with 15 students from three colleges (H, N, and J). Table 1 shows the demographic data of the interview participants. A translation and back-translation procedure was applied to ensure the validity of the self-reported scales related to AI constructs (Klotz et al., 2023). The scales were translated from English to Chinese and then back to English by three bilingual researchers and two experts. The participants were given both versions of the scales.

Table 1

Demographic Information of Interviewees

Student	University	Gender	Age	Grade
S1	H	Female	21	junior
S2	N	Male	22	senior
eS3	J	Female	21	sophomore
S4	H	Male	20	junior
S5	H	Female	19	sophomore
S6	J	Female	18	freshman
S7	H	Female	21	senior
S8	H	Female	22	junior
S9	N	Female	21	sophomore
S10	N	Female	19	freshman
S11	N	Male	22	junior
S12	J	Female	22	junior
S13	H	Male	22	junior
S14	N	Female	21	sophomore
S15	J	Female	19	freshman

Instruments

Perceived Ease of Use and Perceived Usefulness

The scales of perceived ease of use and perceived usefulness from Venkatesh et al. (2003) were used to measure students’ perceptions of the AI system. The scales have five items each, rated on a 5-point Likert scale. The items were adapted to the EFL context. For example, one item for perceived ease of use is “Learning to operate the AI system to assist my English learning would be easy for me.” One item for perceived usefulness is “Using the AI system would enhance my effectiveness in English learning.” The scales had Cronbach’s α coefficients of 0.748 and 0.935 in this study.

Attitude Toward AI

The General Attitudes Towards Artificial Intelligence Scale by Schepman and Rodway (2022) was employed to assess students’ attitude toward AI. The scale has 12 items, rated on a 5-point Likert scale. Higher scores indicate a more positive attitude. Some items are reverse-coded to reduce bias. The items, modified for the EFL context, included statements such as “I am interested in using AI systems in my English learning.” The scale had a high Cronbach’s α coefficient of 0.857 in this study.

Behavioral Intention to Use AI

The scale by Ayanwale et al. (2022) was adopted to measure students’ intention to use AI for EFL learning. The scale has five items, modified for the EFL context, such as “I intend to use AI to assist my English learning.” The items are rated on a 6-point Likert scale. A higher score means a stronger intention. The scale had a high Cronbach’s α coefficient of 0.893 in this study.

Semi-Structured Interviews

Semi-structured interviews were conducted with 15 students who had participated in our questionnaire survey. The interviews lasted about 20 minutes each. There were two main objectives: to further verify the roles of the variables from the quantitative stage in determining intention to use AI, and to understand the reasons for the quantitative results. Therefore, we chose an inductive semi-structured interview method for these exploratory purposes. The following interview questions were carefully designed based on a pilot test with two students:

1. Do you think AI tools (such as ChatGPT, Grammarly, Pigai, and so forth) can help you learn English skills (such as listening, speaking, reading, and writing) in distributed learning contexts?
2. Do you use AI tools for your English learning in distributed learning contexts?
3. If you do, what motivates you to use the specific AI and what prevents you from using it? Could you describe it in detail?

Data Collection

With the help of EFL teachers, 464 Chinese college students from different majors were recruited. English is a mandatory course for all colleges in China. A survey tool (http://www.wjx.cn) was used to collect quantitative data online. The participants were asked to give their demographic information and complete four scales. For the qualitative data, 15 random students were interviewed. The AI tools and distributed learning were explained to them before the interview. The questionnaires and interviews were in Chinese for participants’ understanding and expression. A consent form was signed by participants before the study. The study was approved by the Ethics Committee of Hunan Normal University.

Data Analysis

The analysis of quantitative data was done using IBM SPSS Amos (Version 24). The measurement model was verified first, following Kline's (2016) two-stage SEM approach, and the reliability and validity of the constructs were assessed. Several indices were used to evaluate the goodness-of-fit of the models, including chi-square divided by degrees of freedom, comparative fit index, Tucker-Lewis index, and the root-mean-square error of approximation (Hu & Bentler, 1999). The data were checked for skewness, kurtosis, and descriptive statistics to ensure normality. The "maximum likelihood" method was used to test the structural model. The indirect effects were tested using bootstrapping with 5,000 iterations (Shrout & Bolger, 2002). The indirect effects were considered significant if 0 was not in the 95% confidence interval (Hayes, 2013).

The interviews were transcribed and translated by two linguistics professors, resulting in 12,303 Chinese characters and 9,842 English words. The qualitative data were coded and categorized using MAXQDA 2022, following the thematic analysis of Braun and Clarke (2006) with an inductive approach. The agreement was calculated by the first and corresponding authors, who coded the data separately, to check interrater reliability. For example, "portability" was the code for "Dictionaries are too bulky to carry around, so I prefer to use apps to learn vocabulary." The authors agreed on 115 of 123 codes (93.5%) and resolved their disagreement on the remaining codes.

Quantitative Results

Testing the Measurement Model

A confirmatory factor analysis was performed with four latent variables: perceived ease of use, perceived usefulness, attitude toward AI, and intention to use AI. The factor loadings of the items were checked using unstandardized and standardized estimates, and items with non-significant or low loadings were removed. Items ATA8 and ATA9 were deleted for p-values > 0.05, and items ATA1, ATA3, ATA6, ATA10, and PU6 were deleted for standardized loadings < 0.45 (Kline, 2016). The modification indices were then inspected, and the changes that matched the theory were made.

The measurement model fit the data acceptably, as shown by the following indices: χ^2/df = 3.424 (3–5: acceptable, < 3: excellent), CFI = 0.920 (> 0.9: acceptable, > 0.95: excellent), TLI = 0.908 (> 0.9: acceptable, > 0.95: excellent), RMSEA = 0.072 (< 0.08: acceptable), and SRMR = 0.062 (< 0.08: acceptable) (Hu & Bentler, 1999).

The discriminant validity and composite reliability (CR) of each construct are presented in Table 2. All constructs met the criteria of CR > 0.7 and AVE > 0.5, and had maximum shared variance (MSV) values lower than their AVE, demonstrating good reliability and convergent validity. The Fornell-Larcker criterion indicated that all factors were interrelated, with strong correlations among perceived usefulness, perceived ease of use, and attitude toward AI. The discriminant validity was verified by the fact that the square root of AVE for each construct (the bold values in the table) exceeded its correlations with other factors (Fornell & Larcker, 1981).

Table 2

Composite Reliability and Convergent and Discriminant Validity of Factors Determining EFL Students' Intention to Use AI

Factor	CR	AVE	MSV	Fornell-Larcker Criterion			
				1	2	3	4
1. PEU	0.766	0.790	0.712	**0.899**			
2. PU	0.936	0.744	0.712	0.844*	**0.863**		
3. ATA	0.862	0.690	0.675	0.822*	0.708*	**0.831**	
4. IUA	0.896	0.633	0.377	0.560*	0.479*	0.613*	**0.796**

Note. PEU = perceived ease of use; PU = perceived usefulness; ATA = attitude toward AI; IUA = behavioral intention to use AI; CR = composite reliability; MSV = maximum shared variance. Figures in bold show the square root of AVE for each construct.
* $p < .001$.

Testing the Structural Model

The reliability and validity were verified before the data in the measurement model were analyzed. The descriptive statistics are shown in Table 3. The SEM assumptions were met by the data, as the absolute values of skewness and kurtosis of all items were < 2 (Noar, 2003).

Table 3

Descriptive Statistics of all Items of the Constructs

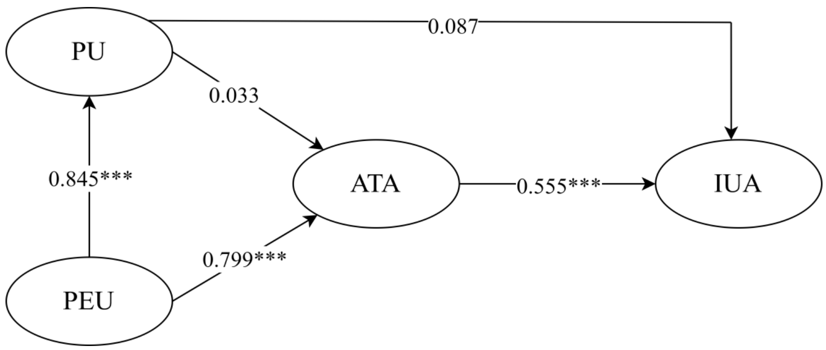
	Min	Max	*M*	*SD*	Skewness	*SE*	Kurtosis	*SE*
PEU	1.000	5.000	2.966–3.776	0.813–0.942	-0.608–-0.090	0.113	-0.291–0.882	0.226
PU	1.000	5.000	3.653–3.683	0.773–0.812	-0.336–-0.483	0.113	0.507–0.939	0.226
ATA	1.000	5.000	3.543–3.866	0.714–0.801	-0.512–0.175	0.113	-0.136–0.741	0.226
IUA	1.000	5.000	4.004–4.362	0.851–0.961	-0.535–-0.323	0.113	0.293–1.559	0.226

Note. $N = 464$. PEU = perceived ease of use; PU = perceived usefulness; ATA = attitude toward AI; IUA = behavioral intention to use AI.

The underlying mechanisms among the four constructs of our study were examined using regression analysis with SEM to address our hypotheses. The structural model of this analysis is depicted in Figure 2.

Figure 2

The Structural Model With Standardized Estimates



Note. PEU = perceived ease of use; PU = perceived usefulness; ATA = attitude toward AI; IUA = behavioral intention to use AI.
*** $p < 0.001$

As shown in Figure 2, H1 was supported by the significant positive effects of perceived ease of use on both perceived usefulness (β = 0.845, $p < .001$) and attitude toward AI (β = 0.799, $p < .001$). Moreover, H3 was confirmed by the significant positive effect of attitude toward AI on behavioral intention to use AI (β = 0.555, $p < .001$). However, perceived usefulness did not significantly affect either behavioral intention to use AI (β= 0.087, $p > .05$) or attitude toward AI (β = 0.033, $p > .05$), thus rejecting H2.

The indirect path analysis results are shown in Table 4. H5 was supported, as the link between perceived ease of use and behavioral intention to use AI was mediated by attitude toward AI (β = 0.444, 95% CI [0.297, 0.656]). However, the other three mediation paths were not supported by the data, as 0 was in their 95% CI. Thus, H4, H6, and H7 were rejected, which contradicts the TAM assumptions. This discrepancy led to the subsequent interviews for deeper understanding.

Table 4

Bootstrapping Analyses of Results of Indirect Effects

Indirect path	*B*	β	*SE*	95% CI	*p*
1. PU → ATA → IUA	0.018	0.018	0.097	[-0.172, 0.145]	.882
2. PEU → ATA → IUA	0.126	0.074	0.071	[-0.044, 0.190]	.303
3. PEU → ATA → IUA	0.758	0.444	0.107	[0.297, 0.656]	.001
4. PEU → PU → ATA → IUA	0.026	0.015	0.083	[-0.150, 0.118]	.872
5. PEU → PU → IUA	0.129	0.074	0.070	[-0.042, 0.193]	.255

Note. PU = perceived usefulness; ATA = attitude toward AI; IUA = behavioral intention to use AI; PEU = perceived ease of use; B = unstandardized path coefficient; CI = confidence interval.

Qualitative Results

Motivations for Using AI in EFL Learning

The interview data were analyzed to investigate the non-significant results of H2, H4, H6, and H7. It was shown in the analysis that all interviewees had a positive attitude toward AI for EFL learning. For the motivations, 71 codes, 11 categories, and 3 themes were extracted. The thematic map is presented in Figure 3.

Figure 3

Thematic Map of Motivations for Using AI for EFL Learning

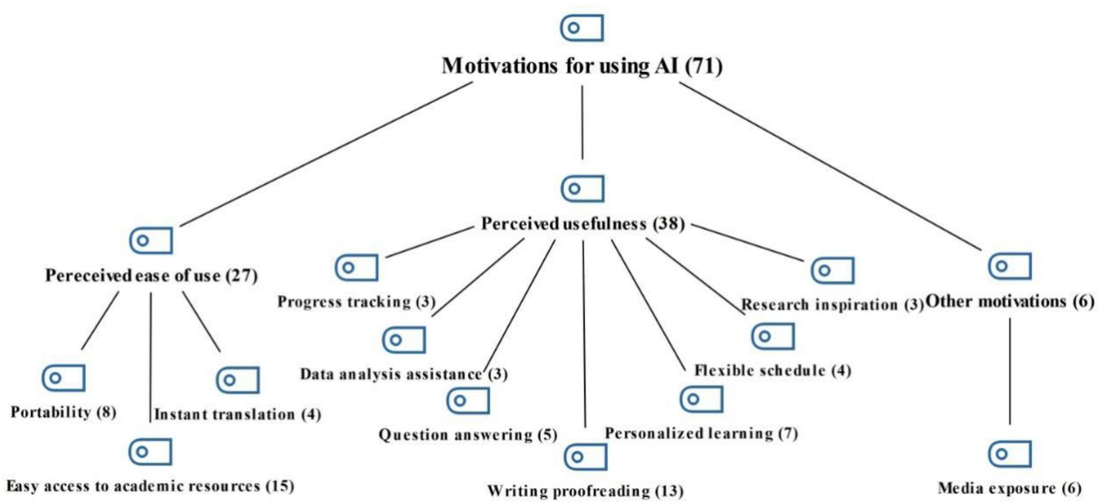


Figure 3 illustrates the dominant themes: perceived ease of use (27 codes) and perceived usefulness (38 codes). Perceived ease of use had 3 categories: easy access to academic resources (15), portability (8), and instant translation (4). Perceived usefulness had more categories, such as writing proofreading (13), personalized learning (17), question answering (5), flexible schedule (4), data analysis (3), research inspiration (3), and progress tracking (3).

We give one example from the interviews for each main theme; underlined text shows key words we used to extract categories. Student 13 described the connection between AI and proofreading, one category of motivation found under perceived usefulness: “First of all, regarding writing, ChatGPT and Grammarly can help me improve my writing skills, fix grammar errors, and increase clarity. Grammarly, in particular, can also offer real-time feedback, making my writing more consistent.” On the other hand, student 5 described the connection between using AI and issues of portability, another category of motivation found under perceived ease of use: “Dictionaries are too bulky to carry around, so I prefer to use apps like Baicizhan and

Bubei Danzi to learn vocabulary more frequently” (Baicizhan and Bubei Danci are two popular mobile applications for learning English in China). Besides the two main themes, another motivation for using AI to facilitate EFL learning was media exposure (6), which influenced students’ awareness and interest in AI. Overall, the two major themes indicate the main motivations for college students to use AI for EFL learning, which aligns with the TAM framework.

Barriers to Using AI for EFL Learning

From the interview data, we identified 53 codes, 8 categories, and 2 themes related to the barriers that prevent college students from using AI to support their EFL learning. Figure 4 shows the thematic map of this analysis.

Figure 4

Thematic Map of Barriers to Using AI for EFL Learning

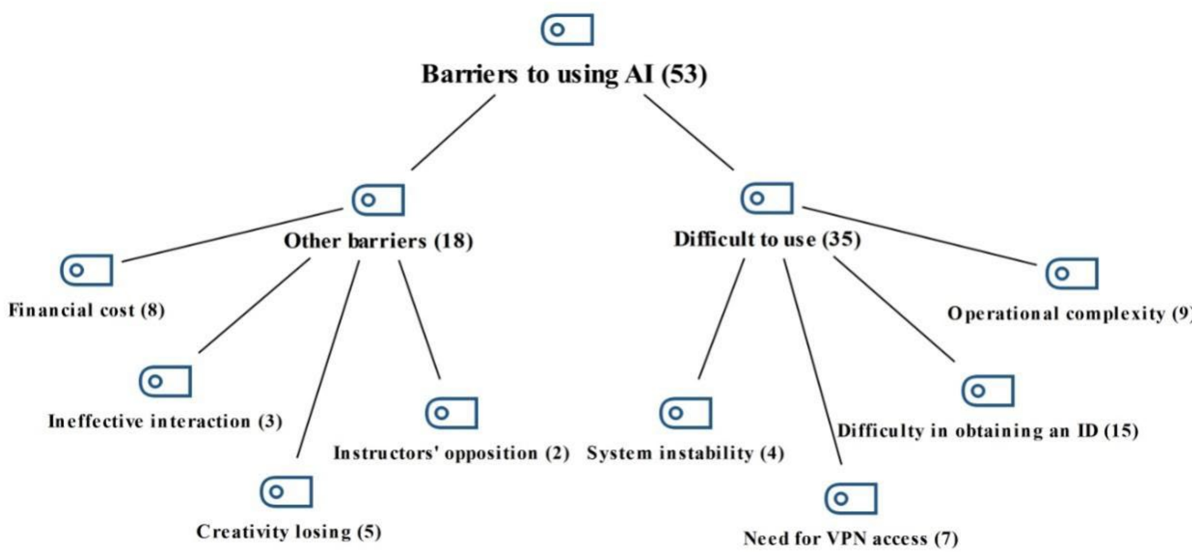


Figure 4 shows the dominant theme of difficult to use (35 codes), which included four categories: difficulty in obtaining an ID (15), operational complexity (9), need for VPN access (7), and system instability (4). Other barriers were financial cost (8), creativity loss (5), ineffective interaction (3), and instructors’ opposition (2). We give one example from the interviews for each main theme; underlined text shows key words we used to extract categories. Student 11 explained how operational complexity makes AI applications difficult to use. “Some AI applications have technical difficulties that prevent me from using them. For example, some applications have complex interface designs, which require me to spend a lot of time learning various functions and settings.” Student 9 spoke about the barriers that result from financial considerations. “Some AI applications for learning English are paid, and they are not very cost-effective when I use them. I cannot stick to them as much as I imagine, so it is not worth buying a membership.”

While we found evidence that ease of use and usefulness motivate students to use AI for EFL learning, our analysis of the barriers to using AI for EFL learning indicates that these factors are not enough to keep students using AI applications. The students said they would stop using AI for EFL learning if they had problems or frustrations with the apps, regardless of their usefulness. This means that ease of use and usefulness start AI-assisted EFL learning, but difficulty to use or low ease of use stop it. This finding matches our quantitative results, which stressed the key role of ease of use in affecting the behavioral intention to use AI in distributed learning contexts.

Discussion

Drawing on the TAM, we employed a mixed-method approach to examine the mechanisms underlying the relationship between perceived ease of use, perceived usefulness, attitude toward AI, and behavioral intention to use AI among EFL college learners in distributed learning contexts. In the quantitative analysis, we first verified the reliability and validity of the instruments used in the study. Afterward, we conducted SEM to examine the proposed hypotheses.

The quantitative results of our study support H1 (perceived ease of use → perceived usefulness, perceived ease of use → attitude toward AI), H3 (attitude toward AI → behavioral intention to use AI), and H5 (perceived ease of use → attitude toward AI → behavioral intention to use AI) by showing that perceived ease of use has a positive effect on both perceived usefulness and attitude toward AI. This indicates that EFL college learners who perceive AI as easy to use are more likely to view AI as useful and have a favorable attitude toward AI. Moreover, EFL college learners with a positive attitude toward AI have a higher intention to use AI. Additionally, the mediated path from perceived ease of use to behavioral intention to use AI through attitude toward AI implies that EFL college learners who perceive AI as easy to use develop a more positive attitude toward AI, which in turn encourages them to use AI for their language learning.

These findings corroborate the TAM, which asserts that perceived ease of use is a key factor for users’ perceived usefulness and attitude toward a technology, and that attitude toward a technology influences behavioral intention to use it (Ursavaş, 2022). Moreover, these results are consistent with the existing literature in general education, which has confirmed the mediating role of attitude toward a specific AI tool in the link between perceived ease of use and behavioral intention to use that AI tool among both college students (Gado et al., 2022; Li, 2023) and teachers (Huang & Teo, 2020; Siyam, 2019). However, few studies have examined these relationships in AI-assisted EFL learning contexts, thus demonstrating the originality and importance of our study. In conclusion, our study validates both the TAM and previous studies in general education, particularly regarding the mediating effect of attitude toward AI on the use of AI among EFL college learners.

The quantitative results of this study are surprising, as they do not support H2 (perceived usefulness → attitude toward AI, perceived usefulness → behavioral intention to use AI), H4 (perceived usefulness → attitude toward AI → behavioral intention), H6 (perceived ease of use → perceived usefulness → attitude toward AI → behavioral intention to use AI), and H7 (perceived ease of use → perceived usefulness → behavioral intention to use AI). It was found that perceived usefulness had no significant effect on attitude toward AI or behavioral intention to use AI. Therefore, attitude toward AI does not mediate the relationship between perceived usefulness and behavioral intention to use AI. Furthermore, perceived usefulness does not mediate the relationship between perceived ease of use and behavioral intention to use AI, either by itself or together with attitude toward AI. This suggests that college students’ perception of AI’s usefulness for EFL learning is not associated with their attitude or intention to use AI.

The results of this study contradict the TAM, which suggests that perceived usefulness influences attitude and behavioral intention to use technology directly or indirectly (Ursavaş, 2022). Moreover, the results diverge from the findings of Wang et al. (2022) and Liu and Ma (2023), who verified the mediating effect of perceived usefulness, either alone or together with attitude, on the relationship between perceived ease of use and behavioral intention to use a specific AI tool among EFL learners. A possible reason for this discrepancy is that users may have varying degrees of familiarity with AI in general and specific AI tools. For instance, in Liu and Ma’s (2023) study, EFL learners’ perception was limited to a particular AI tool, ChatGPT, which is one of the most advanced and powerful AI applications in the world, with high levels of ease of use and usefulness. Consequently, most of its users, including EFL learners, may have a favorable impression of it due to their positive user experience and media exposure.

However, EFL learners’ evaluation of AI in general, revealed in our qualitative analysis, depends on their user experience of the specific AI tools that they are familiar with. Different from ChatGPT, these AI tools vary in quality and performance. Specifically, college students were motivated to adopt AI for EFL learning mainly by perceived ease of use and perceived usefulness. However, these factors were not enough to ensure their continued use of AI applications. The students reported that they would stop using AI for EFL learning if they encountered difficulties or dissatisfaction with the applications, even if they recognized their usefulness. Consequently, they would seek other AI tools for EFL learning. This qualitative finding suggests that perceived ease of use and perceived usefulness are the initial facilitators of AI-assisted EFL learning, but perceived difficulty to use or a low level of perceived ease of use are major barriers to long-term use. In other words, given that they all agree that AI can improve their EFL learning (Betal, 2023; Divekar et al., 2022), the key determinant of their intention to use AI is not its usefulness but its ease of use.

This qualitative finding supports our quantitative finding that perceived usefulness does not play a significant role in predicting EFL college learners’ attitude toward AI and their intention to use AI for language learning. Although it contradicts the assumptions of the TAM, this finding is still reasonable. Currently, there have been a variety of AI applications available for college students to facilitate their language learning, such as ChatGPT (Kohnke et al., 2023a), Duolingo (Shortt et al., 2023), Grammarly (Barrot, 2022), Pigai (Yang et al., 2023), and so forth. With such a wide range of options, it’s plausible that college students tend to opt for the one that they consider most user-friendly. Besides, it’s noteworthy that this finding may imply that the TAM has its robust soundness in terms of its components when it is applied to a single technology. However, in a specific learning context with various technological tools available, a tool’s ease of use may be the most decisive factor for users’ intention to use it.

Conclusion

This mixed-methods study investigated the factors influencing Chinese EFL college learners’ intention to adopt AI. The quantitative results reveal several key findings. First, students’ perceived ease of use of AI has a positive effect on their perceived usefulness and attitude toward AI. Second, students’ attitude toward AI is a positive predictor of their intention to use AI. However, in contrast to the TAM propositions, students’ perceived usefulness of AI does not significantly influence their attitude toward AI or their intention to use AI. Third, students’ perceived ease of use of AI positively influences their intention to use AI through their attitude toward AI, rather than through their perceived usefulness of AI. According to the qualitative results, perceived ease of use and perceived usefulness are the main factors that facilitate college learners’ behavioral intention to use AI for EFL learning. However, only perceived difficulty to use or a lack of perceived ease of use hinder their sustained use of AI, confirming the crucial role of perceived ease of use as revealed by the quantitative results.

Implications and Limitations

This study contributes to both theory and practice. On the theoretical level, our study reveals that perceived ease of use is the main determinant of college learners’ intention to use AI for EFL learning, which contradicts the TAM proposition that perceived usefulness is more influential. This indicates the distinctiveness of the EFL college learning context where various AI-assisted language applications are accessible and the necessity of TAM adaptations. On the practical level, our findings suggest that educators can foster EFL learners’ use of AI applications by providing them with training and guidance. This can lower the perceived difficulty and boost the confidence of learners in using AI applications, particularly in distributed learning environments where learners have greater flexibility and autonomy over time and space. Moreover, AI application developers can enhance their products by improving the usability and user-friendliness of AI applications for EFL learning. They can also solicit feedback from college learners to satisfy target users’ preferences.

This study has several limitations that should be acknowledged. First, it is a cross-sectional study that cannot establish the causal relationships among the three constructs over time. Future research could use a cross-lagged panel design to examine the longitudinal causality among the constructs in this study (Derakhshan et al., 2023). Second, the sample is not sufficiently representative, as it only included participants from two provinces of China. Future studies could expand the sample size and diversity by recruiting participants from more provinces. Third, the study only focused on college learners. However, other populations, such as primary school learners and workers, may also greatly benefit from AI-assisted language learning. These populations could be selected as participants in future research to verify the generalizability of our findings. Fourth, the study only tested the constructs in the TAM, ignoring other factors in similar theories such as effort expectancy and performance expectancy in the unified theory of acceptance and use of technology (Venkatesh et al., 2003). Future studies could explore more variables that could affect EFL learners’ behavioral intention to use AI.

Acknowledgments

This study was sponsored by Teacher Education Project of Foundation of Henan Educational Committee, entitled "Using Online Learning Resources to Promote English as a Foreign Language Teachers' Professional Development in the Chinese Middle School Context" (Grant No.: 2022-JSJYYB-027).

References

An, X., Chai, C. S., Li, Y., Zhou, Y., Shen, X., Zheng, C., & Chen, M. (2022). Modeling English teachers' behavioral intention to use artificial intelligence in middle schools. *Education and Information Technologies*, *28*, 5187–5208. https://doi.org/10.1007/s10639-022-11286-z

An, X., Chai, C. S., Li, Y., Zhou, Y., & Yang, B. (2023). Modeling students' perceptions of artificial intelligence assisted language learning. *Computer Assisted Language Learning*, 1–22. https://doi.org/10.1080/09588221.2023.2246519

Ayanwale, M. A., Sanusi, I. T., Adelana, O. P., Aruleba, K. D., & Oyelere, S. S. (2022). Teachers' readiness and intention to teach artificial intelligence in schools. *Computers and Education: Artificial Intelligence*, *3*, Article 100099. https://www.doi.org/10.1016/j.caeai.2022.100099

Ayedoun, E., Hayashi, Y., & Seta, K. (2019). Adding communicative and affective strategies to an embodied conversational agent to enhance second language learners' willingness to communicate. *International Journal of Artificial Intelligence in Education*, *29*(1), 29–57. https://doi.org/10.1007/s40593-018-0171-6

Barrot, J. S. (2022). Integrating technology into ESL/EFL writing through Grammarly. *RELC Journal*, *53*(3), 764–768. https://doi.org/10.1177/0033688220966632

Bearman, M., Ryan, J., & Ajjawi, R. (2023). Discourses of artificial intelligence in higher education: A critical literature review. *Higher Education*, *86*(2), 369–385. https://doi.org/10.1007/s10734-022-00937-2

Betal, A. (2023). Enhancing second language acquisition through artificial intelligence (AI): Current insights and future directions. *Journal for Research Scholars and Professionals of English Language Teaching*, *7*(39). https://doi.org/10.54850/jrspelt.7.39.003

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Castañeda, L., & Selwyn, N. (2018). More than tools? Making sense of the ongoing digitizations of higher education. *International Journal of Educational Technology in Higher Education*, *15*, Article 22. https://doi.org/10.1186/s41239-018-0109-y

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial intelligence trends in education: A narrative overview. *Procedia Computer Science*, *136*, 16–24.

https://doi.org/10.1016/j.procs.2018.08.233

Chen, H., & Pan, J. (2022). Computer or human: A comparative study of automated evaluation scoring and instructors' feedback on Chinese college students' English writing. *Asian-Pacific Journal of Second and Foreign Language Education*, *7*, Article 34. https://doi.org/10.1186/s40862-022-00171-4

Chen, X., Xie, H., & Hwang, G.-J. (2020). A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers. *Computers and Education: Artificial Intelligence*, *1*, Article 100005. https://doi.org/10.1016/j.caeai.2020.100005

Creswell, J. W. (2014). *A concise introduction to mixed methods research*. SAGE Publications.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Derakhshan, A., Wang, Y., Wang, Y., & Ortega-Martín, J. L. (2023). Towards innovative research approaches to investigating the role of emotional variables in promoting language teachers' and learners' mental health. *International Journal of Mental Health Promotion*, *25*(7), 823–832. https://doi.org/10.32604/ijmhp.2023.029877

Divekar, R. R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen, Y., Zhu, H., Hendler, J. A., & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation. *Computer Assisted Language Learning*, *35*(9), 2332–2360. https://doi.org/10.1080/09588221.2021.1879162

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language, Learning and Technology*, *24*(1), 16–26. https://doi.org/10125/44705

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/10.2307/3151312

Gado, S., Kempen, R., Lingelbach, K., & Bipp, T. (2022). Artificial intelligence in psychology: How can we enable psychology students to accept and use artificial intelligence? *Psychology Learning & Teaching*, *21*(1), 37–56. https://doi.org/10.1177/14757257211037149

Gao, Q., Yan, Z., Zhao, C., Pan, Y., & Mo, L. (2014). To ban or not to ban: Differences in mobile phone policies at elementary, middle, and high schools. *Computers in Human Behavior*, *38*, 25–32. https://doi.org/10.1016/j.chb.2014.05.011

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression based approach*. Guilford Press.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, F., & Teo, T. (2020). Influence of teacher-perceived organisational culture and school policy on Chinese teachers’ intention to use technology: An extension of technology acceptance model. *Educational Technology Research and Development*, *68*(3), 1547–1567. https://doi.org/10.1007/s11423-019-09722-y

Huang, H.-M., & Liaw, S.-S. (2005). Exploring users’ attitudes and intentions toward the web as a survey tool. *Computers in Human Behavior*, *21*(5), 729–743. https://doi.org/10.1016/j.chb.2004.02.020

Janbi, N., Katib, I., & Mehmood, R. (2023). Distributed artificial intelligence: Taxonomy, review, framework, and reference architecture. *Intelligent Systems With Applications*, *18*, Article 200231. https://doi.org/10.1016/j.iswa.2023.200231

Jiang, R. (2022). How does artificial intelligence empower EFL teaching and learning nowadays? A review on artificial intelligence in the EFL context. *Frontiers in Psychology*, *13*, Article 1049401. https://doi.org/10.3389/fpsyg.2022.1049401

Kelly, S., Kaye, S.-A., & Oviedo-Trespalacios, O. (2023). What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telematics and Informatics*, *77*, Article 101925. https://doi.org/10.1016/j.tele.2022.101925

Klotz, A. C., Swider, B. W., & Kwon, S. H. (2023). Back-translation practices in organizational research: Avoiding loss in translation. *Journal of Applied Psychology, 108*(5), 699–727. https://doi.org/10.1037/apl0001050

Klimova, B., Pikhart, M., Benites, A. D., Lehr, C., & Sanchez-Stockhammer, C. (2023). Neural machine translation in foreign language teaching and learning: A systematic review. *Education and Information Technologies*, *28*(1), 663–682. https://doi.org/10.1007/s10639-022-11194-2

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (6th ed.). Guilford Press.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023a). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023b). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence*, *5*, Article 100156. https://doi.org/10.1016/j.caeai.2023.100156

Kuddus, K. (2022). Artificial intelligence in language learning: Practices and prospects. In A. Mire, S. Malik, & A. K. Tyagi (Eds.), *Advanced analytics and deep learning models* (pp. 3–18). Wiley. https://doi.org/10.1002/9781119792437.ch1

Li, K. (2023). Determinants of college students’ actual use of AI-based systems: An extension of the technology acceptance model. *Sustainability*, *15*(6), Article 5221. https://www.mdpi.com/2071-1050/15/6/5221

Liu, G., & Ma, C. (2024). Measuring EFL learners’ use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, *18*(2), 125–138. https://doi.org/10.1080/17501229.2023.2240316

Liu, G. L., & Wang, Y. (2024). Modeling EFL teachers’ intention to integrate informal digital learning of English (IDLE) into the classroom using the theory of planned behavior. *System*, *120*, Article 103193. https://doi.org/10.1016/j.system.2023.103193

Muftah, M., Al-Inbari, F. A. Y., Al-Wasy, B. Q., & Mahdi, H. S. (2023). The role of automated corrective feedback in improving EFL learners’ mastery of the writing aspects. *Psycholinguistics*, *34*(2), 82–109. https://doi.org/10.31470/2309-1797-2023-34-2-82-109

Namaziandost, E., Hashemifardnia, A., Bilyalova, A. A., Fuster-Guillén, D., Palacios Garay, J. P., Diep, L. T. N., Ismail, H., Sundeeva, L. A., Hibana, & Rivera-Lozada, O. (2021). The effect of WeChat-based online instruction on EFL learners’ vocabulary knowledge. *Education Research International*, *2021*, Article 8825450. https://doi.org/10.1155/2021/8825450

Namaziandost, E., Razmi, M. H., Atabekova, A., Shoustikova, T., & Kussanova, B. H. (2021). An account of Iranian intermediate EFL learners’ vocabulary retention and recall through spaced and massed distribution instructions. *Journal of Education*, *203*(2), 275–284. https://doi.org/10.1177/00220574211031949

Noar, S. M. (2003). The role of structural equation modeling in scale development. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(4), 622–647. https://doi.org/10.1207/S15328007SEM1004_8

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, *2*, Article 100020. https://doi.org/10.1016/j.caeai.2021.100020

Rezai, A. (2023). Investigating the association of informal digital learning of English with EFL learners’ intercultural competence and willingness to communicate: A SEM study. *BMC Psychology*, *11*, Article 314(2023). https://doi.org/10.1186/s40359-023-01365-2

Schepman, A., & Rodway, P. (2022). The General Attitudes towards Artificial Intelligence Scale (GAAIS): Confirmatory validation and associations with personality, corporate distrust, and general trust. *International Journal of Human-Computer Interaction*, *39*(13), 2724–2741. https://doi.org/10.1080/10447318.2022.2085400

Selwyn, N. (2016). *Is technology good for education?* Polity Press.

Shortt, M., Tilak, S., Kuznetcova, I., Martens, B., & Akinkuolie, B. (2023). Gamification in mobile-assisted language learning: A systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning*, *36*(3), 517–554. https://doi.org/10.1080/09588221.2021.1933540

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*(4), 422–445.

Siyam, N. (2019). Factors impacting special education teachers’ acceptance and actual use of technology. *Education and Information Technologies*, *24*(3), 2035–2057. https://doi.org/10.1007/s10639-018-09859-y

Teo, T., Huang, F., & Hoi, C. K. W. (2017). Explicating the influences that explain intention to use technology among English teachers in China. *Interactive Learning Environments*, *26*(4), 460–475. https://doi.org/10.1080/10494820.2017.1341940

Ulla, M. B., Perales, W. F., & Busbus, S. O. (2023). ‘To generate or stop generating response’: Exploring EFL teachers’ perspectives on ChatGPT in English language teaching in Thailand. *Learning: Research and Practice*, *9*(2), 168–182. https://doi.org/10.1080/23735082.2023.2257252

Ursavaş, Ö. F. (2022). Technology acceptance model: History, theory, and application. In Ö. F. Ursavaş (Ed.), *Conducting technology acceptance research in education: Theory, models, implementation, and analysis* (pp. 57–91). Springer International Publishing. https://doi.org/10.1007/978-3-031-10846-4_4

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision Sciences*, *39*(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Wang, Y., Yu, L., & Yu, Z. (2022). An extended CCtalk technology acceptance model in EFL education. *Education and Information Technologies*, *27*, 6621–6640. https://doi.org/10.1007/s10639-022-10909-9

Wang, Y. L., Wang, Y. X., Pan, Z. W., & Ortega-Martín, J. L. (2023). The predicting role of EFL students’ achievement emotions and technological self-efficacy in their technology acceptance. *The Asia-Pacific Education Researcher*, *2023*. https://doi.org/10.1007/s40299-023-00750-0

Yang, H., Gao, C., & Shen, H.-z. (2023). Learner interaction with, and response to, AI-programmed automated writing evaluation feedback in EFL writing: An exploratory study. *Education and*

Information Technologies, *29*, 3837–3858. https://doi.org/10.1007/s10639-023-11991-3

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research and future directions. *Computers and Education: Artificial Intelligence*, *2*, Article 100025. https://doi.org/10.1016/j.caeai.2021.100025

Zhi, R., Wang, Y., & Wang, Y. (2023). The role of emotional intelligence and self-efficacy in EFL teachers' technology adoption. *The Asia-Pacific Education Researcher*, *2023*. https://doi.org/10.1007/s40299-023-00782-6

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

How AI Literacy Affects Students' Educational Attainment in Online Learning: Testing a Structural Equation Model in Higher Education Context

Jingyu Xiao[1], Goudarz Alibakhshi[2,*], Alireza Zamanpour[3], Mohammad Amin Zarei[4], Shapour Sherafat[5], and Seyyed-Fouad Behzadpoor[6]

*[1]Guangdong University of Foreign Studies, Guangzhou, China; [2]Allameh Tabataba'i University, Iran; [3]Islamic Azad University Science and Research Branch, Tehran, Iran; [4]Allameh Tabataba'i University; [5]University of Tehran, Iran; [6]Azarbaijan Shahid Madani University, Tabriz, Iran; *Corresponding author*

Abstract

Artificial intelligence (AI) has contributed to various facets of human lives for decades. Teachers and students must have competency in AI and AI-empowered applications, particularly when using online electronic platforms such as learning management systems (LMS). This study investigates the structural relationship between AI literacy, academic well-being, and educational attainment of Iranian undergraduate students. Using a convenience sampling approach, we selected 400 undergraduate students from virtual universities equipped with LMS platforms and facilities. We collected data using three instruments—an AI literacy scale, an academic well-being scale, and educational attainment scale—and analyzed the data using Smart-PLS3 software. Results showed that the hypothetical model had acceptable psychometrics (divergent and convergent validity, internal consistency, and composite reliability). Results also showed that the general model had goodness of fit. The study thus confirms the direct effect of AI on academic well-being and educational attainment. By measuring variables of academic well-being, we also show that AI literacy in China and Iran significantly affects educational attainment. These findings have implications for students, teachers, and educational administrators of universities and higher education institutes, providing knowledge about the educational uses of AI applications.

Keyword: AI, AI applications, academic well-being, AI literacy, educational attainment, undergraduate students

Introduction

Open and distributed learning, often referred to as distance education or online learning, is a flexible educational approach that transcends traditional classroom boundaries (Bozkurt et al., 2015). Unlike conventional face-to-face instruction, open and distributed learning uses digital technologies to deliver educational content and facilitate interactions among learners and instructors across geographical locations and time zones (Alavi et al., 2002). This approach allows learners to engage with course materials, participate in discussions, and complete assignments remotely, typically through online platforms or learning management systems (LMS; Alavi et al., 2002).

Complementing this online learning approach, the rapid evolution of technology, particularly the advent of artificial intelligence (AI), has become a transformative force in shaping various facets of our daily lives. The proliferation of smart devices and applications embedded with AI has ushered in an era in which individuals are transitioning from being mere AI immigrants to being proficient AI natives. This paradigm shift has had a multifaceted impact on society, influencing how people work, learn, and interact.

The swift advancement of e-learning platforms is a notable global trend in higher education. Even before the COVID-19 pandemic, universities worldwide were increasingly exploring and adopting online teaching and learning methods. However, the global health crisis significantly accelerated this transition, pushing many educational institutions to pivot rapidly to online modalities. Hodges et al. (2020) observe this transformative shift, noting how the pandemic has compelled universities to make online education the primary mode of instruction. Consequently, the integration of online learning platforms within higher education has become a significant global phenomenon and reshaped the landscape of academic delivery. The accelerated adoption of online education has profound implications for students, educators, and institutions. As universities grapple with the challenges and opportunities presented by online learning, it becomes imperative to discern the factors contributing to students' academic success in virtual classrooms. The study by Hodges et al. (2020) underscores the urgency of this examination.

Simultaneously, the rapid integration of AI into various aspects of our lives necessitates a critical evaluation of our readiness to navigate the challenges posed by the "AI era" (Davenport & Ronanki, 2018). The term "AI natives" reflects a new reality wherein individuals are not merely users but proficient navigators of AI-driven technologies. This shift prompts reevaluating the skills and competencies required in the modern world. Consequently, competence in effectively using AI is increasingly recognized as essential. Researchers such as Kandlhofer et al. (2016) and Tarafdar et al. (2019) emphasize the urgent need to enhance people's capacity to interact with and harness the potential of AI.

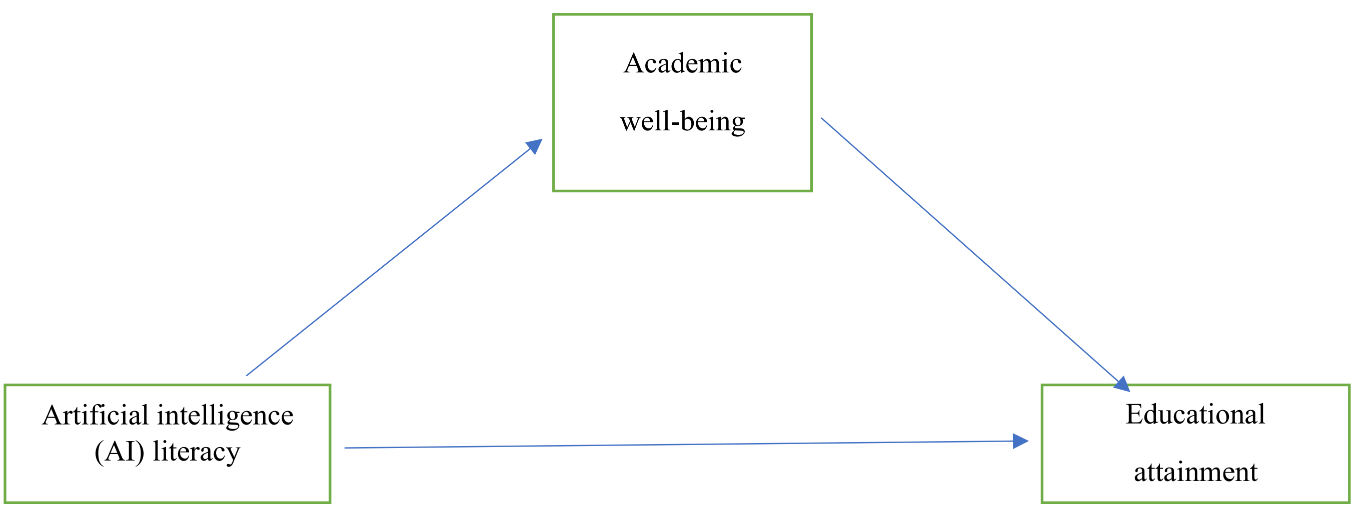
Emerging AI systems, exemplified by technologies like IBM's Watson, demonstrate remarkable capabilities in learning and self-improvement. This advancement has led to their use in knowledge-based tasks that were traditionally exclusive to human white-collar workers. Tasks once deemed resistant to automation are now within the purview of AI systems (Wladawsky-Berger, 2017). The intelligence of AI technologies is rapidly evolving, positioning them as semi-autonomous decision-makers across a growing array of intricate contexts (Davenport & Kirby, 2016).

While Jarrahi (2018) and Stembert and Harbers (2019) have emphasized the benefits of high AI competence in human–AI interaction, other scholars have stressed the need to improve AI competence. Long and Magerko (2020) have attempted to delineate core competencies using AI technology. However, a comprehensive framework or tool for assessing such competence still needs to be developed. To fill this gap, we suggest using "AI literacy" to describe people's proficiency with AI. AI literacy is the ability to correctly recognize, operate, and assess AI-related products using moral guidelines. Comparable to other forms of literacy—such as digital (Ala-Mutka, 2011; Calvani et al., 2008) and computer literacy (Hoffman & Blake, 2003)—AI literacy emphasizes proficiency and appropriate use rather than requiring knowledge of the underlying theory.

Research on AI literacy is pivotal for three main reasons. Firstly, AI literacy contributes to understanding ongoing research in human–AI interaction. An individual's level of literacy regarding a product shapes their mental model (Rosling & Littlemore, 2011), which is crucial in interaction processes (Norman, 2013). Research on AI literacy may elucidate variations in people's behavior during AI interactions. Secondly, unlike previous studies that measured participants' AI competence through prior experience and usage frequency (Lee & Choi, 2017; Luo et al., 2019; Metelskaia et al., 2018), Thirdly, AI literacy provides a more comprehensive measurement of competence that considers the multifaceted aspects of AI use. However, the structural relationship between AI literacy, academic resilience, academic well-being, and educational attainment of university students in online classes with learning management systems (LMS) still needs to be explored. This study seeks to validate the structural equation modeling illustrating the interplay of academic resilience, academic well-being, and educational attainment of Iranian undergraduate students in online classes with an LMS, as depicted in Figure 1.

Figure 1

Hypothetical Model of the Study



Based on the model, we propose the following five hypotheses:

H1: AI literacy has a significant effect on undergraduate students' academic well-being.

H2: Undergraduate students' academic well-being significantly affects their educational attainment.

H3: Undergraduate students' AI literacy has a significant direct effect on their educational attainment.

H4: Undergraduate students' AI literacy has a significant indirect effect on their educational attainment through the mediating role of academic well-being.

H5: The hypothetical structural model of undergraduate students' AI literacy, educational attainment, and academic well-being has a goodness of fit.

Literature Review

Theoretical Framework

One of the defining features of open and distributed learning is its accessibility, as it offers educational opportunities to individuals who may face barriers to attending traditional brick-and-mortar institutions (Bozkurt et al., 2015; Lea & Nicoll, 2013). This accessibility is particularly beneficial for non-traditional students—such as working professionals, parents, individuals with disabilities, or people residing in remote areas—who may find it challenging to pursue higher education through conventional means (Chen et al., 2021). By breaking down geographical constraints and temporal limitations, open and distributed learning democratizes access to education, promoting lifelong learning and continuous skill development (Alavi et al., 2002).

Furthermore, open and distributed learning fosters a learner-centered approach, allowing individuals to personalize their learning experiences according to their preferences, pace, and learning styles (Kop et al., 2011). Through asynchronous and synchronous communication tools, learners can engage with course content at their convenience, collaborate with peers, and receive timely feedback from instructors (Matheos & Archer, 2004). This flexibility empowers learners to take ownership of their education, cultivate self-discipline, and enhance digital literacy skills, which are increasingly vital in today's technology-driven society (Matheos & Archer, 2004).

In the context of AI literacy, open and distributed learning presents unique challenges and opportunities (Mühlenbrock et al., 1998). On the one hand, the digital nature of open and distributed learning means students must be proficient in navigating online platforms, interacting with AI-driven tools and resources, and critically evaluating AI-generated content. Learners must develop AI literacy skills to discern between reliable and misleading information, understand the ethical implications of AI technologies, and harness AI tools effectively to support their learning objectives (Mühlenbrock et al., 1998). On the other hand, open and distributed learning platforms can use AI-driven algorithms to personalize learning experiences, adapt instructional content to individual needs, and provide targeted interventions to support struggling learners (Mühlenbrock et al., 1998). AI technologies embedded within LMS platforms can analyze learner data, identify patterns of engagement and performance, and offer personalized recommendations for optimizing learning outcomes (Mühlenbrock et al., 1998). Thus, open and distributed learning environments serve as fertile grounds for exploring the intersection of AI literacy, educational attainment, and academic well-being among diverse student populations, including Iranian university students.

Artificial Intelligence and Education

The inception of AI in education dates back to the 1970s, marked by initial attempts to replace teachers with supercomputers. Pivotal research between 1982 and 1984 demonstrated that students receiving both human and AI instruction outperformed those exposed solely to traditional teaching methods (Hao, 2019; Kay, 2012). Integrating AI into education offers unique prospects for hands-on learning and enhancing technology-based educational settings. However, despite the potential of AI for education, many professionals in technology and education seek guidance in establishing effective procedures and frameworks (Li, et al., 2024; Kay, 2012).

As suggested by Mollman (2022), the release in November 2022 of the ChatGPT AI tool, based on the generative pre-trained transformer (GPT) language model, significantly increased public awareness of AI. This sophisticated chatbot tool broadened its user base to over a million within a week of its launch, showcasing its rapid adoption and success. Beyond technical capabilities, ChatGPT has the potential to revolutionize various societal aspects, including the generation of scholarly articles and assistance in complex tasks such as writing academic papers (Liu et al., 2021). The emergence of AI, represented by ChatGPT, poses challenges and opportunities for educators and students, emphasizing the need for specialized digital skills and literacy in the era of information technologies (Budzianowski & Vulić, 2019; Cote & Milliner, 2018).

AI involves the development of machines capable of tasks traditionally requiring human intelligence, ranging from learning and reasoning to problem-solving and pattern recognition (Davenport & Kalakota, 2019; Terra et al., 2023). It). Its applications extend to healthcare, where AI demonstrates equal or superior performance in tasks such as disease diagnosis (Gómez-Trigueros et al., 2019).

Wang and Wang (2022) have conducted a study addressing the need for a standardized tool to measure artificial intelligence anxiety. They developed an AI anxiety scale through a rigorous process, confirming the reliability and validity of the instrument with a sample of 301 respondents and contributing to the understanding of artificial intelligence anxiety and associated behaviors. In addition, Wang and Wang (2022) introduced the concept of AI literacy, identifying core constructs: awareness, use, evaluation, and ethics. They developed a 12-item scale through a rigorous process, confirming its adequacy. The study revealed significant associations between AI literacy, digital literacy, attitudes toward robots, and daily AI usage, providing insights into human–AI interaction and application design.

AI constitutes a dynamic and interdisciplinary field that spans computer science, information science, mathematics, psychology, sociology, linguistics, and philosophy, as highlighted by Russell and Norvig (2010). This convergence of diverse disciplines accentuates the nuanced distinctions between AI and digital literacy. While digital literacy encompasses proficiency in conventional digital technologies, AI literacy involves a more intricate understanding of systems endowed with biological and social attributes, as elucidated by Poria et al. (2017) and Tao and Tan (2005).

Users' perceptions of the AI landscape are notably distinct from their perceptions of conventional digital technology. They tend to attribute more human-like and social qualities to AI entities, reflecting a cognitive shift in how they see and interact with technology. This shift departs from conventional digital literacy, where users engage primarily with tools that lack the social and biological dimensions they perceive as inherent in AI systems (Poria et al., 2017). When users engage with AI, particularly in human–robot interaction, their cognitive processes often lean toward social logic rather than machine logic, as Vossen and Hagemann (2010) have observed. This underscores a fascinating aspect of AI literacy—users navigate and interpret AI systems using mental models that incorporate social and relational dynamics, mirroring human interactions. In contrast, users typically approach conventional digital technologies with a more transactional and utilitarian mindset (Tao & Tan, 2005).

In essence, the interdisciplinary nature of AI not only reflects the convergence of diverse scientific and philosophical domains but also gives rise to a distinct cognitive landscape. The delineation between AI literacy and digital literacy is more than just a matter of technical proficiency. It encompasses a profound shift in how individuals conceptualize and engage with technology, particularly regarding its perceived social and biological dimensions. This shift in mental models emphasizes the need for a nuanced approach to literacy education that encompasses the multifaceted dimensions of contemporary technological landscapes (Davenport & Kalakota, 2019).

Consequently, users employ distinct criteria to evaluate AI and digital products. AI literacy cannot be directly equated with digital literacy, as demonstrated by a digitally literate high school student unfamiliar with AI concepts. This makes digital literacy instruments unsuitable for assessing AI literacy. However, the digital literacy framework can guide AI literacy development. The next section explores essential digital literacy concepts for a more nuanced understanding of AI literacy (Wang & Wang, 2022).

Academic Well-being and Educational Attainment

A comprehensive assessment of a student's academic experiences—comprising their engagement in learning, sense of belonging, and satisfaction with school—is referred to as their academic well-being. Academic well-being has been determined to be a significant element influencing students' academic achievement (Seligman et al., 2009). A person's level of education is referred to as their educational attainment, and it is frequently used as a gauge of their success in school. This literature review aims to identify the factors that affect academic well-being and explore the relationship between well-being and educational attainment. Academic well-being and educational attainment are positively correlated: a study conducted by Suldo et al. (2008) found that students' grades and test scores were higher when they reported higher levels of academic well-being, which included positive emotions, a solid academic self-concept, and supportive relationships. In addition, a study by Salmela-Aro and Upadyaya (2014) found that students who reported higher academic well-being were likelier to complete secondary school and pursue higher education. Several factors influence the relationship between academic well-being and educational attainment. One of the critical factors is social support: that is, how much students feel encouraged by their parents, teachers, and peers to pursue their academic goals. Students who have robust social networks are more likely to succeed academically. Moreover, research indicates that a positive school climate—marked by elevated expectations, nurturing relationships, and avenues for active participation—is pivotal in fostering academic well-being and the achievement of educational goals (Marsh, 1990). Student motivation, particularly intrinsic motivation, has also been identified as a positive correlate of intellectual well-being and educational attainment (Salmela-Aro & Upadyaya, 2014). Finally, academic self-concept, representing an individual's belief in their capabilities, is a crucial factor influencing intellectual well-being and educational achievement (Yeager & Dweck, 2012).

In light of the comprehensive literature review encompassing the theoretical underpinnings of open and distributed learning, the intersections of AI literacy, and the dynamics of academic well-being and educational attainment, this study aims to address a critical gap in understanding the complex relationships among these variables within the context of Iranian university students' experiences in online classes with LMS. While previous research has extensively explored the individual components of AI literacy, academic well-being, and educational attainment, very few empirical investigations have examined the interplay between these constructs, particularly within the framework of open and distributed learning. Moreover, the evolving landscape of AI technologies, coupled with the unprecedented challenges posed by the COVID-19 pandemic, underscores the need for a nuanced examination of how undergraduate students navigate AI-driven online learning environments and the implications of this for their academic well-being and educational outcomes.

Methodology

Sampling and Procedure

The participants for this research were selected from Guangdong University of Foreign Studies in China Payame Noor University, and the Tehran University e-Learning Center and encompassed individuals from diverse academic disciplines, in Iran. Participants were selected through a random sampling technique to ensure a representative cross-section of the universities' student populations. In determining the sample size for this investigation, we adhered to a robust methodology involving power analysis and sample size calculation.

Considering a preset significance level and power, the power analysis sought to determine the minimum sample size necessary for identifying statistically significant differences in the study's results. The effect size—obtained from earlier studies on similar subjects—was included in this computation, and the significance level was fixed at 0.5. A standard level in social science research, 0.80, was chosen to determine the study's power. The study included 400 participants, 220 (55%) of whom identified as female and 180 (45%) as male. The age range of the participants was 20–35 years, with a mean of–28 years (SD = 5). Remarkably, half of the subjects were between the ages of 20 and 23 (50%). Subjects aged 23–26 accounted for 25% of participants, and those 26–30 made up 13%. A smaller group (12%) of participants belonged to the 30–35 year age range. The intentional variation in the sample concerning age, gender, and level of education improves the applicability of the study's conclusions to a larger group of young adults. Table 1 presents the participants' demographic data.

Table 1

Demographic Profile of Participants

		Number of participants	Percentage
Gender	Female	220	55
	Male	180	45
Age	20–23	200	50
	23–26	100	25
	26–30	50	12.5
	30–35	50	12.5
Level	Basic sciences	120	30
	Humanities	140	35
	Engineering	140	35

This study implemented a cross-sectional survey design using structural equation modeling (SEM) analysis as the research methodology. Students from a specific educational institution were selected for the study and willingly filled out the survey. During the administration of the survey, students were required to complete a series of self-report tasks carefully crafted to assess their level of academic achievement, academic well-being, and AI literacy. Using the cross-sectional survey design, we gathered information from the student body at a particular moment. This method made it easier to investigate the connections between the variables being studied, which provided important information about the dynamics occurring in the given setting. The study's main goal was to examine the variables impacting the student cohort's educational attainment and academic well-being.

The chosen students completed a comprehensive survey. After the data collection process was finished, SEM analysis was conducted to examine the complex relationships between the variables of interest. SEM is a sophisticated statistical technique that makes it possible to evaluate complex relationships by considering both direct and indirect effects. We evaluated the proposed model's overall fit and methodically tested the proposed relationships using SEM. This methodology thoroughly and systematically investigates the complex relationships between academic achievement, AI literacy, and student well-being.

Instruments

In this study, three instruments were used to thoroughly assess all variables of the study The primary tool was the AI literacy scale, created by Wang et al. (2023), with a total of 12 items. This scale condenses usage, evaluation, awareness, and ethics into five distinct factors. The scale's items are carefully designed to measure participants' proficiency in these areas, giving rise to a comprehensive understanding of AI literacy.

The second tool used was the academic well-being scale, a self-report tool for assessing students' overall academic well-being. The academic work score is based on three primary dimensions: academic engagement, positive emotions, and sense of purpose. It uses a five-point scoring system, with a maximum score of 100. Higher scores indicate higher levels of academic well-being on this scale, providing a quantitative measure to evaluate and compare participants' well-being across various dimensions.

The third tool was the educational attainment scale, which focused explicitly on grade point average (GPA). This scale involved the meticulous consideration of students' GPA for the present semester and the cumulative GPA encompassing all courses and semesters. Notably, in Iranian higher education, GPA is computed based on the grades acquired in individual classes. The standard GPA scale in this educational setting spans from 0 to 20, with 20 representing the pinnacle of achievement. Using this instrument, we could obtain a comprehensive overview of participants' academic performance using a standardized metric widely recognized in the educational domain.

Data Collection

Data for this study were acquired by administering a comprehensive survey questionnaire encompassing self-report measures targeting AI literacy, academic well-being, and educational attainment. The survey, designed for electronic delivery, was disseminated to the student cohort during a specified time frame. The anonymity and privacy of participants were meticulously safeguarded. The survey was structured to be anonymous, assuring respondents that their input would remain confidential. This method created an atmosphere in which students felt free to respond honestly and completely. Participants were more inclined to give frank feedback when worries about identification or consequences were removed.

An electronic invitation with detailed instructions and a hyperlink to the survey was sent to the chosen students as part of the administration process. Students had to give their informed consent after being made aware of the study's goal and the importance of their involvement in supporting academic research. Students were informed about potential rewards for their participation, including gift cards and participation certificates, to encourage their involvement further. Students were free to finish the survey whenever it was convenient for them in the allotted time. They were equipped with their electronic devices and provided with Internet connectivity, and asked to react thoughtfully, offering their viewpoints and firsthand accounts of academic perseverance, student satisfaction, academic welfare, and educational achievement. This methodology aimed to obtain a nuanced and rich dataset that reflected the varied experiences and perspectives of the students who took part in the study.

Data Analysis

A thorough and systematic statistical analysis was performed on the gathered dataset to investigate the complex relationships between essential variables, such as academic resilience, student engagement, well-being, educational attainment, and personality type. Several crucial steps were involved in the analytical process to obtain insightful results. First, we carried out a thorough data cleaning to guarantee the accuracy and completeness of the survey answers. Second, we maintained the dataset's integrity by fixing erroneous or missing data points using methods such as imputation or removal. Third, we then calculated descriptive statistics to summarize the important variables. Central tendencies, variabilities, and data distributions were expressed using metrics such as means and standard deviations. We carried out a reliability analysis to evaluate the internal consistency of the self-report measures. For every scale in the survey, we calculated the Cronbach's alpha coefficient and used it as a metric to assess the consistency and dependability of the measurements. After that, we performed a bivariate correlation analysis to examine the connections between the variables, determining the degree and direction of correlations using metrics such as Pearson's correlation coefficient. SEM was used to test the proposed relationships between variables. This sophisticated analytical approach facilitated the evaluation of both direct and indirect effects, yielding a comprehensive understanding of the intricate relationships within the model. To assess the adequacy of the proposed model, we scrutinized model fit indices, including chi-square, comparative appropriate index, root mean square error of approximation, and standardized root mean square residual.

Results

The study outcomes are delineated in two segments: an assessment of the model and findings related to examining the student's personality types. Appraising the theoretical model involved applying multivariate and SEM analysis through Smart-PLS3 software. This selection was driven by the non-normal distribution of the data and the relatively modest sample size. Figures 2 and 3 present the model's results, illustrating non-standard coefficients (indicating significant coefficients) and standard coefficients (reflecting effect coefficients), respectively.

Figure 2

Evaluation of Hypothetical Model Based on Non-Standard Coefficients (Significant Coefficients)

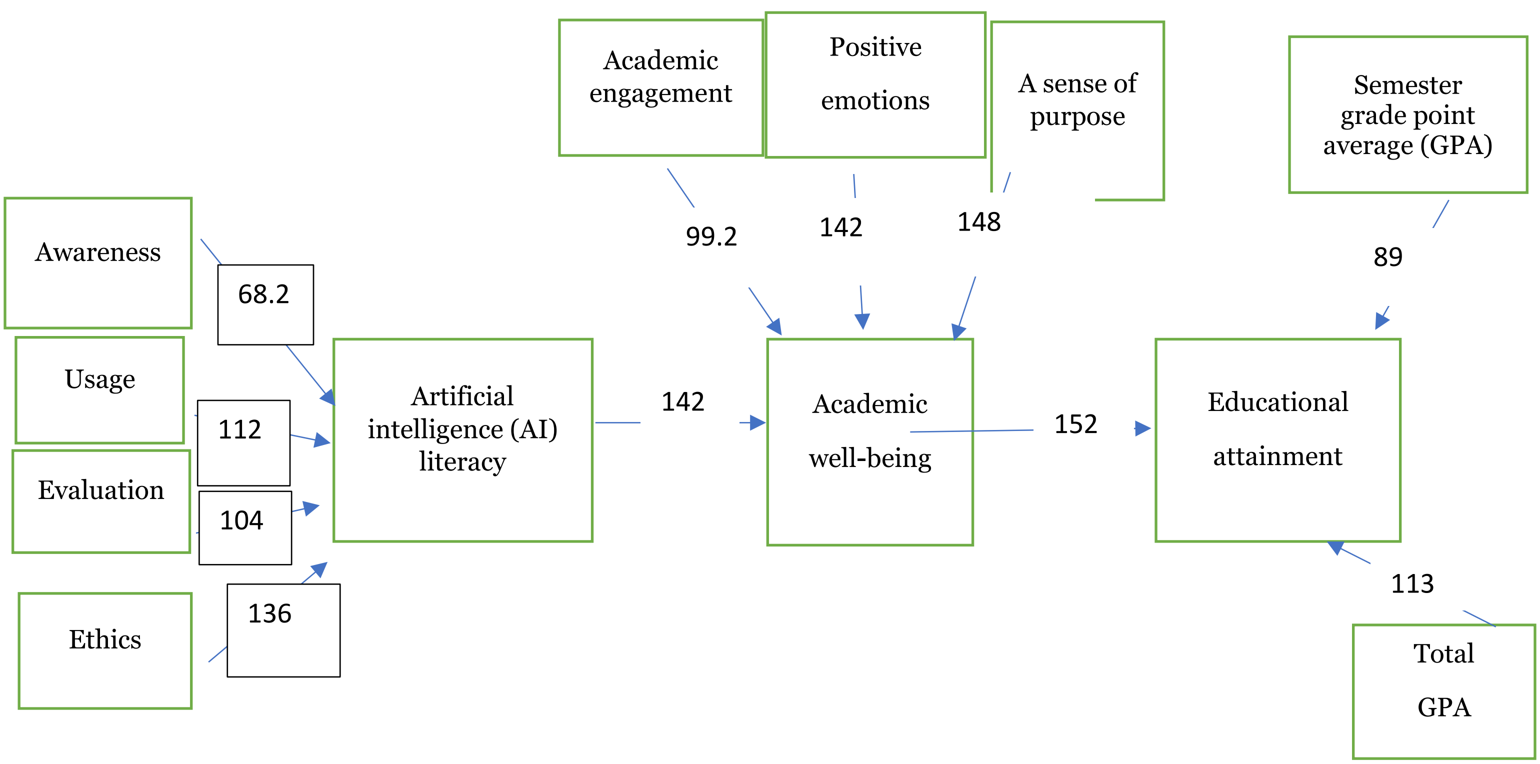
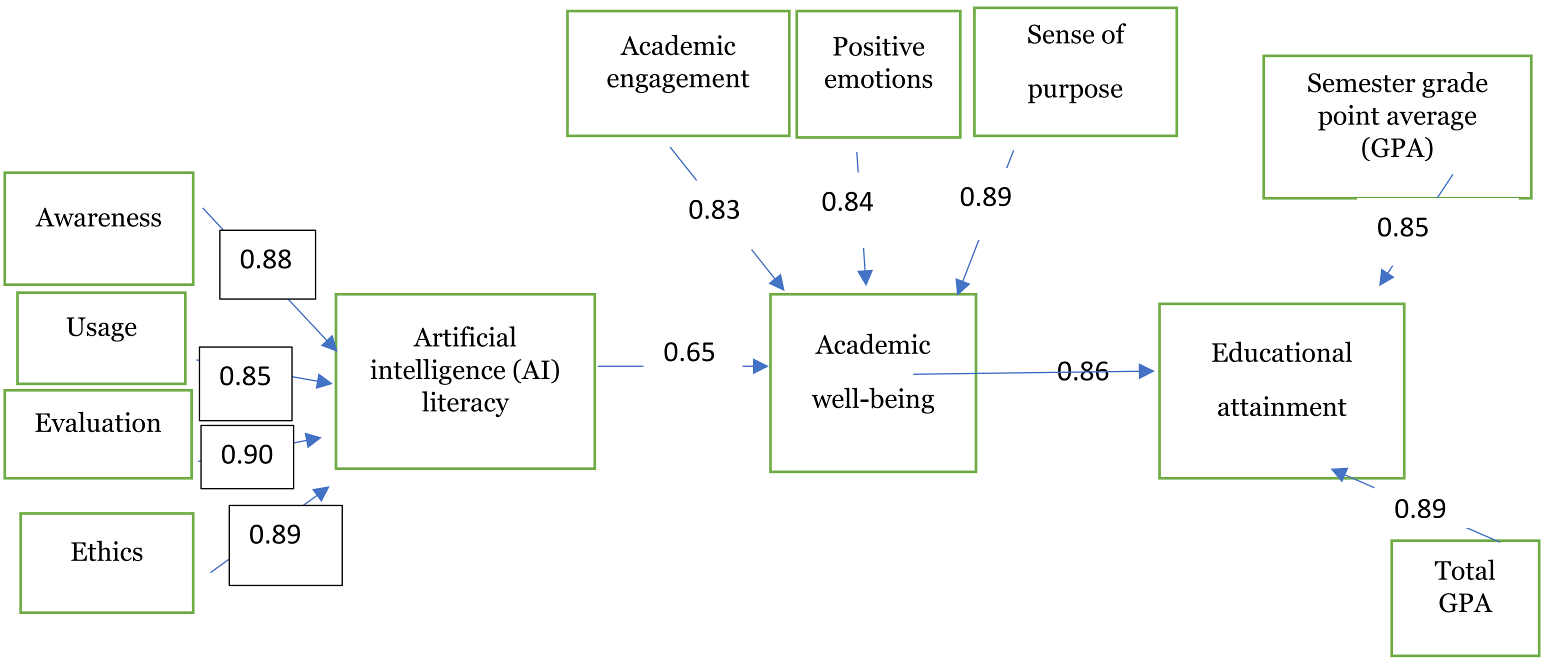


Figure 3

Evaluation of Hypothetical Model Based on Standard Coefficients



In general, when evaluating structural equation models with Smart-PLS3 software, researchers examine three models: the outer, inner, and entire empirical models. Similar to the measurement model in SEM, the external model shows the relationships between manifest or observed variables (here represented as indicators) and latent or unobservable variables (both independent and dependent). Examining the relationships between latent or unobservable variables, the inner model relates to the structural model or path analysis in SEM. Furthermore, the overall empirical model evaluates the model's global suitability. As a result, we tested and evaluated the general practical model concerning students' educational attainment and the empirical structural model (path) and measurement. Table 2 summarizes the findings on the measurement model (outer).

Table 2

Psychometrics (Validity and Reliability) of the Hypothetical Model

Construct	Indicator	Validity assessment			Reliability assessment		
		Convergent validity		Divergent validity			
		FL	AVE	Fornell & Larker Index	Cronbach's alpha	Rho-A	Composite reliability
Artificial intelligence (AI) literacy	Awareness	0.826	0.740	0.870	0.920	0.940	0.910
	Usage	0.714					
	Evaluation	0.820					
	Ethics	0.923					
Academic well-being	Academic engagement	0.842	0.920	0.920	0.930	0.936	0.948
	Positive emotions	0.907					
	Sense of purpose	0.983					
Educational attainment	Semester grade point average (GPA)	0.84	0.880	0.9	0.940	0.9	0.930
	Total GPA	0.86					

Note. FL=Factor Loading, AVE=Average variance explained

Table 2 and Figure 3 show that the cross-loadings of the independent and mediating variables (personality, academic resilience, and academic well-being) on educational attainment, mediated by learner enjoyment, were statistically significant at the 0.01 level ($p < 0.01$) and exceeded the acceptable threshold of 0.7. This implies acceptable correlations between the latent variables (independent and mediating variables) and the observed variables (components). Additionally, all research constructs extracted average variance extracted (AVE) values greater than the acceptable cutoff of 0.5, demonstrating the constructs' convergent validity. Furthermore, as indicated by Table 2, the square root of the AVE for each construct outperforms its correlation with other constructs. As a result, the model's acceptable discriminant validity is confirmed. The latent or independent and mediating variables are more strongly associated with their components than other constructs. Lastly, the internal consistency (reliability) assessment, as indicated in the results of Table 2, discloses that Cronbach's alpha coefficient for the research constructs (independent and mediating variables) exceeds the acceptable threshold of 0.7. Moreover, the composite reliability and the homogenous Rho-A coefficient also surpass the acceptable threshold of 0.7 for all constructs, indicating satisfactory composite reliability.

Internal Model Evaluation (Structural or Path Model)

Path Coefficients of the Structural Model (Significance Coefficients)

This section evaluates all pathways delineated in the internal model, which illustrates relationships between constructs according to research hypotheses by applying the *t*-test for statistical significance. Consequently, should the confidence level of the test exceed 1.96 and 2.58, the pathways will be considered statistically validated at the 95% and 99% confidence levels, respectively. As illustrated in Figure 2, all hypotheses have been duly acknowledged, with their respective *t*-values demonstrating statistical significance at the 99% and 95% confidence levels ($p < 0.01$ and $p < 0.05$, respectively).

Examination of the Coefficient of Determination

The coefficient of determination, signifying the overall explained variance of the dependent variable (educational attainment) predicated on independent and mediating variables in the structural model, stands at 0.983, reflecting an exceptionally high level. Furthermore, the independent variables in the study have demonstrated the capacity to elucidate and forecast 96.1% of the variance in the mediating variable.

Examination of Research Hypotheses

The stated hypotheses were tested, and the results are presented in Table 3.

Table 3

Evaluation of the Stated Hypotheses

Research hypotheses	Effect sizes			Results
	Non-standard coefficients		Standard coefficients	
	t	*p*		
H1: AI literacy has a significant effect on undergraduate students' academic well-being.	9.823	p < 0/01	0.145	Accepted
H2: Undergraduate students' academic well-being significantly affects their educational attainment.	1.941	p < 0/05	0.067	Accepted
H3: Undergraduate students' AI literacy has a significant direct effect on their educational attainment.	3.711	$p < 0.01$	0.677	Accepted
H4: Undergraduate students' AI literacy has a significant indirect effect on their educational attainment through the mediating role of academic well-being.	21.070	p < 0/01	0.778	Accepted
H5: The hypothetical structural model of undergraduate students' AI literacy, educational attainment and academic well-being has a goodness of fit.	21.070	p < 0/01	0.778	Accepted

Conclusively, through a comprehensive assessment of the model employing the RNS Theta index, registering a value of 0.623 out of 100, we can infer that the generated model bears a relatively close resemblance to the theoretical model. Thus, the overarching hypothesis of the research, grounded in the structural and measurement model elucidating the impact of AI literacy on the educational attainment of students with the mediating variable of well-being, is substantiated.

Discussion

The findings of the study resonate strongly with the theoretical framework of open and distributed learning, as they highlight the transformative potential of online education in breaking down traditional barriers to learning (Bozkurt et al., 2015; Alavi et al., 2002). The accessibility provided by open and distributed learning aligns with the study's emphasis on democratizing education, particularly for non-traditional students facing geographical or temporal constraints (Lea & Nicoll, 2013; Chen et al., 2021). Moreover, the learner-centered approach facilitated by online platforms reflects the empowerment of students to personalize their learning experiences, a key feature emphasized in the theoretical framework (Kop et al., 2011; Matheos & Archer, 2004). The integration of AI literacy within open and distributed learning environments represents both an opportunity and a challenge, echoing the study's exploration of the intersection between AI literacy and educational attainment (Mühlenbrock et al., 1998). By using AI-driven algorithms to personalize learning experiences and support struggling learners, open and distributed learning platforms exemplify the innovative potential of technology to enhance educational outcomes, consistent with the study's findings on the role of AI in education. Overall, the study underscores the significance of open and distributed learning in promoting lifelong learning, skill development, and academic well-being among diverse student populations, aligning closely with the principles espoused in the theoretical framework.

Moreover, the findings of this study align with the broader discourse on the evolving landscape of education, particularly in the context of the integration of AI and the changing nature of literacy. The distinction between emergency remote teaching and online learning, as highlighted by Hodges et al. (2020), becomes crucial in understanding the implications of AI literacy for academic well-being and educational attainment.

In a similar vein, the study underlines the significant impact of AI literacy on the academic well-being of undergraduate students. Insights from Davenport and Kalakota (2019) regarding the potential of artificial intelligence in healthcare resonate with the idea that AI literacy may contribute positively to students' overall well-being in an academic setting. As educational environments increasingly incorporate AI-driven tools and resources, students with higher AI literacy skills may experience enhanced well-being, potentially stemming from increased efficacy in navigating and using AI technologies (Lee & Choi, 2017).

Moreover, the influence of AI literacy on academic well-being aligns with Ala-Mutka's (2011) conceptualization of digital competence and Yeager and Dweck's (2012) acknowledgment of the multifaceted nature of student success. The ability to effectively engage with AI technologies is a modern facet of digital competence, impacting students' overall well-being and shaping their experiences in educational settings.

The study also establishes a significant connection between academic well-being and educational attainment. This finding is consistent with the broader literature emphasizing the intricate relationship between psychological well-being and academic success (Suldo et al., 2008; Salmela-Aro & Upadyaya, 2014). Positive emotions, dedication, and absorption, as captured by the Schoolwork Engagement Inventory (Salmela-Aro & Upadyaya, 2014), are crucial components of academic well-being that contribute to sustained effort and focus, ultimately influencing educational outcomes.

The direct effect of AI literacy on educational attainment underscores the role of AI literacy as a determinant of academic success. This is in line with the work of Jarrahi (2018) on AI's impact on organizational decision-making, suggesting that proficiency in AI literacy may translate into more informed and adequate decision-making in academic contexts. The significant indirect effect of AI literacy on educational attainment, mediated by intellectual well-being, adds nuance to the relationship. This aligns with the idea that students with higher AI literacy benefit not only directly from their skills but also indirectly through the positive impact on their academic well-being. The holistic nature of the influence of AI literacy on educational outcomes is supported by Long and Magerko (2020) and their insights into AI literacy competencies.

Integrating AI literacy into educational curricula becomes imperative, as Kandlhofer et al. (2016) suggest, to prepare students for the evolving technological landscape. The findings of the study by Kandlhofer et al. (2016) support the idea that AI literacy is not merely a technical skill but a factor influencing students' well-being and educational achievements. Considering the challenges Stembert and Harbers (2019) identify in designing with AI, educational institutions need to foster a balance that accounts for the human aspects of AI integration. Our study contributes to understanding how AI literacy intersects with academic well-being and educational attainment among undergraduate students (Hao, 2019; Kay, 2012). The implications extend to educational practices, emphasizing the need for AI literacy education and holistic support for students to thrive in an AI-driven educational landscape. Future research could delve into specific AI literacy components and explore interventions that promote both AI literacy and well-being in educational settings (Liu et al., 2021; Mollman, 2022).

The implications of this study also extend to various stakeholders in the field of education, including language teachers, materials developers, and policymakers. For language teachers, the findings highlight the importance of integrating AI literacy into language instruction to equip students with the necessary skills to navigate and critically evaluate AI-driven language learning tools and resources. Language teachers can use AI technologies to personalize instruction, adapt materials to individual student needs, and provide targeted interventions to support language learning outcomes. Additionally, language teachers can use online platforms and LMS to create interactive and engaging learning experiences that foster student engagement and autonomy in language learning. Materials developers can use AI-driven algorithms to develop adaptive language learning materials that cater to learners' diverse needs and preferences, ensuring accessibility and inclusivity in language education. Furthermore, policymakers can use the findings to advocate for the integration of AI literacy and online learning platforms into language education policies and curricula, promoting lifelong learning and continuous skill development in language learners. Policymakers can also invest in teacher training programs to enhance language teachers' proficiency in using AI technologies effectively in language instruction, thereby fostering innovation and excellence in language education. Overall, the study underscores the transformative potential of AI and online learning in language education and highlights the need for collaborative efforts among language teachers, materials developers, and policymakers to harness these technologies for the benefit of language learners.

Conclusions

This study sheds light on the complex relationships between academic success, AI literacy, and educational attainment and offers insightful information about the changing educational environment. The results highlight the significant influence of AI literacy on the academic well-being of undergraduate students. They are consistent with the broader discussion on how artificial intelligence can be used to improve a range of fields. Students with higher levels of AI literacy are likely to be more adept at navigating the growing number of AI-driven tools being integrated into educational environments. This will likely lead to increased success and well-being in the digital age. The literature that has already been written about the complex relationship between psychological well-being and academic success is consistent with the observed relationship between academic well-being and educational attainment. Good feelings, commitment, and immersion—essential elements of academic health—greatly impact long-term effort and concentration, which in turn shapes learning outcomes. In line with current discussions on AI's impact on decision-making, the direct relationship between educational attainment and AI literacy highlights the latter's critical role as a predictor of academic success.

The findings of this study have broad implications for educational approaches, highlighting the need to include AI literacy in curricula as a critical component that influences students' academic performance and well-being in addition to being a technical skill. The study emphasizes the critical use of a balanced approach—one that considers the integration of human elements—when designing with AI. The significance of comprehensive AI literacy education and holistic support for students to thrive in an AI-driven educational landscape is highlighted by this research, which adds to our understanding of the relationship between AI literacy and academic well-being and educational attainment. Subsequent investigations may examine particular elements of AI literacy and interventions that work in concert to advance both AI literacy and student well-being in educational environments.

Acknowledgments

The authors acknowledge that the manuscript is original. Some statistical terms (such as structural equation modeling, Confirmatory Factor Analysis average variance explained) might be detected as by plagiarism checkers. We also acknowledge that we used the AI-empowered applications Poe and Grammarly to check the accuracy of the language of the manuscript.

References

Ala-Mutka, K. (2011). *Mapping digital competence: Towards a conceptual understanding* (Report no. JRC67075). Institute for Prospective Technological Studies. https://www.researchgate.net/publication/340375234

Alavi, M., Marakas, G. M., & Yoo, Y. (2002). A comparative study of distributed learning environments on learning outcomes. *Information Systems Research*, *13*(4), 404–415. https://doi.org/10.1287/isre.13.4.404.72

Bozkurt, A., Akgun-Ozbek, E., Yilmazel, S., Erdogdu, E., Ucar, H., Guler, E., Sezgin, S., Karadeniz, A., Sen-Ersoy, N., Goksel-Canbek, N., Deniz Dincer, G., Ari, S., & Aydin, C. H. (2015). Trends in distance education research: A content analysis of journals 2009–2013. *International Review of Research in Open and Distributed Learning*, *16*(1), 330–363. https://doi.org/10.19173/irrodl.v16i1.1953

Budzianowski, P., & Vulić, I. (2019). Hello, it's GPT-2—how can I help you? Towards the use of pre-trained language models for task-oriented dialogue systems. *arXiv*:1907.05774. https://doi.org/10.48550/arXiv.1907.05774

Calvani, A., Cartelli, A., Fini, A., & Ranieri, M. (2009). Models and instruments for assessing digital competence at school. *Journal of E-Learning and Knowledge Society*, *4*(3). https://doi.org/10.20368/1971-8829/288

Chen, M., Gündüz, D., Huang, K., Saad, W., Bennis, M., Feljan, A. V., & Poor, H. V. (2021). Distributed learning in wireless networks: Recent progress and future challenges. *IEEE Journal on Selected Areas in Communications*, *39*(12), 3579–3605. https://doi.org/10.1109/JSAC.2021.3118346

Cote, T., & Milliner, B. (2018). A survey of EFL teachers' digital literacy: A Report from a Japanese University. *Teaching English with Technology, 18*(4), 71–89. https://tewtjournal.org/download/6-a-survey-of-efl-teachers-digital-literacy-a-report-from-a-japanese-university-by-travis-cote-and-brett-milliner/

Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal, 6*(2), 94–98. https://doi.org/10.7861/futurehosp.6-2-94

Davenport, T.H., & Ronanki, R. (2018) Artificial Intelligence for the Real World. *Harvard Business Review, 96*, 108-116. https://doi.org/10.1007/s11747-019-00696-0

Gómez-Trigueros, I. M., Ruiz-Bañuls, M., & Ortega-Sánchez, D. (2019). Digital literacy of teachers in training: Moving from ICTs (information and communication technologies) to LKTs (learning and knowledge technologies). *Education Sciences, 9*(4), 274. https://doi.org/10.3390/educsci9040274

Hao, K. (2019, August 2). China has started a grand experiment in AI education. It could reshape how the world learns. *MIT Technology Review*.

https://www.technologyreview.com/2019/08/02/131198/china-squirrelhas-started-a-grandexperiment-in-ai-education-it-could-reshape-how-the/

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020, March 27). The difference between emergency remote teaching and online learning. *EDUCAUSE Review*. https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Hoffman, M., & Blake, J. (2003). Computer literacy: Today and tomorrow. *Journal of Computing Sciences in Colleges, 18*(5), 221–233.

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human–AI symbiosis in organizational decision making. *Business Horizons, 61*(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

Kandlhofer, M., Steinbauer, G., Hirschmugl-Gaisch, S., & Huber, P. (2016). Artificial intelligence and computer science in education: From kindergarten to university. In *IEEE Frontiers in Education Conference (FIE)* (pp. 1–9). IEEE. http://doi.org/10.1109/FIE.2016.7757570

Kay, J. (2012). AI and Education: Grand Challenges. *IEEE Intelligent Systems, 27*(5), 66–69. https://doi.org/10.1109/MIS.2012.92.

Kop, R., Fournier, H., & Mak, J. S. F. (2011). A pedagogy of abundance or a pedagogy to support human beings? Participant support on massive open online courses. *International Review of Research in Open and Distributed Learning*, *12*(7), 74–93. https://doi.org/10.19173/irrodl.v12i7.1041

Lea, M. R., & Nicoll, K. (Eds.). (2013). *Distributed learning: Social and cultural approaches to practice*. Routledge. https://doi.org/10.4324/9781315014500

Lee, S., & Choi, J. (2017). Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *International Journal of Human-* https://doi.org/10.1016/j.ijhcs.2017.02.005

Li, G., Zarei, M.A., Alibakhshi, G. *& Labbafi, A. (2024)*. Teachers and educators' experiences and perceptions of artificial-powered interventions for autism groups. *BMC Psycholo*gy, **12**, (2024). https://doi.org/10.1186/s40359-024-01664-2

Liu, X., Zheng, Y., Du, Z., Ding, M., Qian, Y., Yang, Z., & Tang, J. (2021). GPT understands, too. *arXiv:2013.10385*. https://doi.org/10.48550/arXiv.2103.10385

Long, D., & Magerko, B. (2020, April). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). Association for Computing Machinery. https://aiunplugged.lmc.gatech.edu/wp-content/uploads/sites/36/2020/08/CHI-2020-AI-Literacy-Paper-Camera-Ready.pdf

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, *38*(6), 937–947. https://doi.org/10.1287/mksc.2019.1192

Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology*, *82*(4), 623–636. https://doi.org/10.1037/0022-0663.82.4.623

Matheos, K., & Archer, W. (2004). From distance education to distributed learning: Surviving and thriving. *Online Journal of Distance Learning Administration*, *7*(4), 1–13.

Metelskaia, I., Ignatyeva, O., Denef, S., & Samsonowa, T. (2018). A business model template for AI solutions. In *Proceedings of the International Conference on Intelligent Science and Technology* (pp. 35–41). Association for Computing Machinery.

Mollman, S. (2022). ChatGPT gained 1 million users in under a week. https://www.yahoo.com/video/chatgpt-gained-1-million-followers-224523258.html?

Mühlenbrock, M., Tewissen, F., & Hoppe, U. (1998). A framework system for intelligent support in open distributed learning environments. *International Journal of Artificial Intelligence in Education*, *9*, 256–274.

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic Books.

Poria S, Cambria E, Bajpai R & Hussain A (2017) A review of affective computing: From unimodal analysis to multimodal fusion, *Information Fusion, 37*, 98-125. http://doi.org/10.1016/j.inffus.2017.02.003

Rosling, A., & Littlemore, K. (2011). Improving student mental models in a new university Information setting. In *Digitisation Perspectives* (pp. 89–101). Brill.

Russell, S. J., & Norvig, P. (2010). *Artificial intelligence: A modern approach*. Pearson.

Salmela-Aro, K., & Upadyaya, K. (2014). School burnout and engagement in the context of demands–resources model. *British Journal of Educational Psychology*, *84*(1), 137–151. DOI: 10.1111/bjep.12018

Seligman, M. E. P., Ernst, R. M., Gillham, J., Reivich, K., & Linkins, M. (2009). Positive Education: Positive psychology and classroom interventions. *Oxford Review of Education, 35*(3), 293–311. DOI:10.1080/03054980902934563

Stembert, N., & Harbers, M. (2019). Accounting for the human when designing with AI: Challenges identified. In *CHI'19-extended abstracts, Glasgow, Scotland UK—May 04-09, 2019*.

Suldo, S. M., Shaunessy, E., & Hardesty, R. (2008). Relationships among stress, coping, and mental health in high-achieving high school students. *Psychology in the Schools, 45*(4), 273–290. https://doi.org/10.1002/pits.20300

Tao, J., & Tan, T. (2005). Affective computing: A review. In *International Conference on Affective Computing and Intelligent Interaction* (pp. 981–995). Springer.

Tarafdar, M., Beath, C. M., & Ross, J. W. (2019, June 11). Using AI to enhance business operations. *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/using-ai-to-enhance-business-operations/

Terra, M., Baklola, M., Ali, S., & El-Bastawisy, K. (2023). Opportunities, applications, challenges and ethical implications of artificial intelligence in psychiatry: A narrative review. *Egyptian Journal of Neurology, Psychiatry and Neurosurgery, 59*(1), 80–90. https://doi.org/10.1186/s41983-023-00681-z

Vossen, G., & Hagemann, S. (2010). *Unleashing Web 2.0: From concepts to creativity*. Elsevier.

Wang, B., Rau, P. L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. *Behavior and Information Technology, 42*(9), 1324–1337. https://doi.org/10.1080/0144929X.2022.2072768

Wang, Y. Y., & Wang, Y. S. (2022). Development and validation of an artificial intelligence anxiety scale: An initial application in predicting motivated learning behavior. *Interactive Learning Environments*, *30*(4), 619–634. https://doi.org/10.1080/10494820.2019.1674887

Wladawsky-Berger, I. (2017, February 22). The emerging, unpredictable age of AI. *MIT Initiative on the Digital Economy*. http://ide.mit.edu/news-blog/blog/emergingunpredictable-age-ai

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47*(4), 302–314. https://doi.org/10.1080/00461520.2012.722805

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

Can Artificial Intelligence Give a Hand to Open and Distributed Learning? A Probe into the State of Undergraduate Students' Academic Emotions and Test Anxiety in Learning via ChatGPT

Sha Gao
Zhejiang Industry & Trade Vocational College, Wenzhou 325003, China

Abstract

Artificial Intelligence (AI), as an innovation in technology, has greatly affected human life. AI applications such as ChatGPT have been used in different fields, particularly education. However, the use of AI applications to enhance undergraduate students' academic emotions and test anxiety has not been appropriately investigated. This study addresses the effects of undergraduate students' test anxiety and academic emotions. A total of 160 undergraduate students majoring in different fields of study were selected through convenience sampling and divided into control and experimental groups. Both groups received test anxiety and academic emotions scales at the onset of the treatment. The students assigned to the experimental group were trained to use ChatGPT and monitored for learning and doing their assignments outside the classroom during the semester. The two groups received the scales at the end of the semester, which lasted 16 weeks. Independent samples *t*-tests were used for analyzing the data. Results revealed that using AI-empowered applications significantly reduced the students' test anxiety and negative academic emotions but enhanced their positive academic emotions. Students can use ChatGPT as an auxiliary instrument to overcome their negative emotions and enhance their educational attainment. Findings affect teachers, educational technologists, educational psychologists, and students.

Keywords: AI-empowered applications, undergraduate students, academic emotions, test anxiety

Introduction

Distance learning has rapidly expanded in a short period due to the power of the Internet and high-speed communication. This expansion has been fueled by the shift to using smartphones, virtual reality, and augmented reality as tools for advancing blended learning structures (Clark, 2020). Virtual education, a paradigm of distance learning, has reached significant development levels nationally and internationally. The Ministry of National Education in Colombia defines virtual education as a teaching method that overcomes spatial and temporal constraints (Ngah et al., 2022).

According to this model, scientific activity is related to the teacher and involves the student. Virtual education transfers information in virtual classrooms and continuously supports the production and reproduction of information, contributing to knowledge generation (Dhir et al., 2017). Virtual education can be said to refer to a method of education where the teacher and student are separated from each other in terms of time, place, or both, unlike traditional methods presented in face-to-face classrooms, laboratories, etc. (Roa et al., 2022]).

Virtual learning generally refers to education in a learning environment where the teacher and student are separated by time, place, or both. The content of these courses is transmitted through information technology programs, multimedia resources, the Internet, video conferences, and so on (Dung, 2020). According to Garrison et al. (2003), virtual education occurs over a network via an Internet environment within a formal structure; a set of multimedia technologies is used in its creation. Johnson et al. (2023) stated that virtual education employs network technology to design, select, manage, and expand education. Khan (2003) viewed virtual learning as an innovative approach that uses Web facilities to provide education remotely. Virtual learning is not just about delivering educational content; it focuses on the learning process and knowledge generation, making use of information technology and computers to create learning experiences (Horton, 2006).

In the context of open and distributed learning (ODL), the proliferation of virtual education platforms and technologies has reshaped the landscape of education so that it transcends traditional boundaries of time and space (Blake, 2009). ODL, characterized by its flexibility and accessibility, has become increasingly synonymous with virtual education, offering learners opportunities to engage with educational content remotely (Rumble, 1989). Leveraging advancements in information and communication technologies, ODL environments enable students to access resources, interact with instructors, and collaborate with peers regardless of geographical constraints (Fleming & Hiple, 2004). As such, the integration of artificial intelligence (AI) into these platforms holds immense potential to further enhance the efficacy and inclusivity of ODL experiences.

The expansion of virtual space has had extensive effects on human life. Nowadays, one distressing factor for students is exam anxiety, a common phenomenon in schools (Yazdani & Asadi, 2022). Piroozmanesh and Imanipour (2018) described *anxiety* as a widespread, unpleasant, and ambiguous feeling of fear and apprehension with an unknown origin that puts individuals in a state of agitation and stress in related circumstances. It includes uncertainty, indecision, and physiological arousal, affecting individuals' mental well-being (Salmalian et al., 2020).

Exam anxiety is considered a significant inhibitory factor in successful assessment and learning for students, imposing substantial costs on societies (King et al., 1991; Lufi & Awwad, 2013). Moreover, it is one of the most prominent psychological issues among students, consistently emphasized by various theorists and researchers. Exam anxiety has been studied extensively since the early 20th century and has always been a severe issue in the field of education. Exam anxiety is a psychological reaction to an evaluative situation that leaves individuals doubtful and reduces their coping abilities in that situation (Basaknezhad et al., 2013). It results from cognitive and physiological responses triggered in testing situations or similar evaluative conditions (Ghafourian et al., 2020).

The second area that national technology, particularly distance learning, might affect is academic emotions. As a critical determinant of university students' access to social and economic opportunities, students' academic achievement depends on many cognitive, affective, and educational variables. Therefore, identifying the main factors that promote and correlate with academic achievement is necessary (Hayat et al., 2018). Different factors affect students' performance in educational contexts (Yavorsky, 2017), including academic emotions related to motivational, cognitive, physiological, and behavioral processes (Pekrun et al., 2011). Pekrun's (2006) control-value theory provides researchers with a comprehensive framework to study the impacts of different emotions that students experience in academic contexts. This theory assumes that expectancy-value theories of transactional approaches, attributional theories, and performance models affect students' emotions and achievement. Different emotions, particularly academic emotions, are also believed to be associated with students' academic outcomes (Talib & Sansgiry, 2012). Yu and Dong (2010) maintained that students' academic emotions affect their academic achievements, manifested while doing daily homework, learning in classrooms, and taking exams. Researchers have classified academic emotions in different ways. However, the most commonly used typology of academic emotions reveals that positive and negative emotions are either activating or deactivating; that is, there are positive/negative activating emotions and positive/negative deactivating emotions.

This study holds significance in contributing to the existing body of knowledge by shedding light on the specific relationship between AI-empowered technology educational applications (apps) and undergraduate students' academic emotions and test anxiety. The findings of this research have practical implications for educators, app developers, and policymakers in shaping the future of AI in education. A comprehensive understanding of how these technologies impact students' emotional well-being can guide the development of better tailored and more effective educational tools. Additionally, insights from this study can inform strategies to create a supportive learning environment, ultimately enhancing the overall educational experience for undergraduate students. Although undergraduate students' test anxiety and academic emotions have been studied in different ways, the effect of AI-empowered educational applications on undergraduate Chinese students' test anxiety and academic emotions has not been well explored. To fill in the gap, the following research questions were raised:

1. Does undergraduate students' use of AI-empowered educational applications significantly affect their test anxiety?

2. Does undergraduate students' use of AI-empowered educational applications significantly affect their academic emotions?

Literature Review

Open and Distributed Learning

ODL constitutes a multifaceted educational paradigm characterized by its use of technology to extend learning opportunities beyond traditional classroom confines. Central to ODL is the concept of "openness," advocating for unrestricted access to educational resources and materials (Bates, 2015). This approach aims to democratize education by removing barriers such as geographical, temporal, and socioeconomic constraints, thereby fostering inclusivity and equity in learning (Bates, 1997).

In ODL, the theoretical framework of distributed learning plays a pivotal role. Distributed learning emphasizes the dispersion of learning activities across various modalities, platforms, and contexts (Siemens, 2005). It recognizes the dynamic interplay between formal and informal learning environments, acknowledging that learning occurs not only within structured educational settings but also through interactions with peers and mentors and through real-world experiences (Anderson, 2016). By embracing the principles of distributed learning, ODL endeavors to create interconnected learning ecosystems that facilitate seamless navigation between different learning environments and modalities, promoting lifelong learning and adaptability in the digital age (Bozkurt et al., 2015).

AI and Higher Education

Twenty-first-century higher education is rapidly changing due to globalization, technological advancements, and student demographics (Dieguez et al., 2021). Online learning platforms have become widely accessible, enabling universities to offer fully online courses and degree programs, expanding access to education, and providing flexibility in learning (Neumann & Baumann, 2021). The growing diversity of the educational field, with students from various backgrounds, highlights the significance of global citizenship and intercultural understanding. Universities play a significant role in promoting innovation and research as technological advancements speed up (Amornkitpinyo et al., 2021), encouraging industry–academia cooperation and focusing on commercialization and entrepreneurship. The emphasis is shifting toward skills-based learning patterns for practical, career-focused skills, as evidenced by recent recruitment trends favoring graduates with particular skills (Koçak et al., 2021).

To enhance the quality of higher education, the industry is exploring various strategies to meet stakeholders' requirements (Khan et al., 2022). AI integration is one particularly hopeful solution (Chedrawi et al., 2019). As technology advances, AI in education has enormous potential to change the teaching and learning environment (Bahado-Singh et al., 2019). AI is significantly improving the quality of higher education in several ways (Ali & Choi, 2020). AI-powered learning strategies evaluate students' performance, pinpoint their advantages and disadvantages and offer individualized learning experiences. With the help of these strategies, students can acquire knowledge and produce more valuable results in the real world (Aldosari, 2020).

Chatbots, virtual assistants, and adaptive learning systems are examples of AI-based technologies that provide immersive and exciting learning environments while enabling students to actively investigate complicated ideas (Chaudhry et al., 2023; Pradana et al., 2023). AI helps with assessment and feedback in grading assignments, tracking student participation, giving quicker and more accurate feedback, and freeing up teachers' time for other teaching responsibilities (Essien et al., 2020). AI chatbots provide quick, individualized support by evaluating student data to identify individuals who may be at risk of academic failure and enabling early interventions for academic success. Various AI apps and platforms, including Bit.ai, Mendeley, Turnitin, elink.io, and Coursera are tools that support higher education research by analyzing large datasets, generating insights, and identifying patterns challenging for human researchers to detect (Wenge, 2021). We expect even more cutting-edge AI applications to emerge in education due to continued technological advancement, giving students individualized, engaging, and productive learning experiences (Li et al., 2021).

The exciting development of AI dramatically improves both the effectiveness and engagement of instructors in postsecondary education. Adopting AI helps educators free up time for more meaningful activities by automating administrative duties like tracking attendance and grading assignments (Bisen et al., 2021). Additionally, AI allows educators to pinpoint areas in which they can grow by offering individualized opportunities for professional development (Minkevics & Kampars, 2021). Solutions are needed for enduring problems in modern higher education, such as limited inclusivity and unequal access (Odhiambo, 2016). Traditional teaching methods hinder active participation and critical thinking skills (Kistyanto et al., 2022). The inability of traditional assessment techniques to capture thorough understanding makes using AI necessary. AI algorithms analyze individual learning patterns, tailor coursework, and predict at-risk students, enabling timely interventions (Rudolph et al., 2023). Content delivery is revolutionized by AI-driven systems that adjust to students' learning styles, pace, and knowledge gaps.

Adopting AI in higher education empowers the system by addressing challenges and enhancing the quality of education. Ongoing research aims to understand faculty members' awareness of AI's applicability and impact on learning experiences, work engagement, and productivity in higher education. This research provides insights for institutional policymakers to facilitate the adoption of new technologies and overcome specific challenges. Despite the increasing integration of technology and AI in education, there is a notable gap in understanding how AI-powered educational apps specifically influence the academic emotions and test anxiety of undergraduate students. While various studies have explored the general impact of technology on education and student emotions, focused research on the unique effects of AI-powered educational apps is needed. Understanding the dynamics between these technologies and students' emotional experiences can provide valuable insights into the efficacy of AI app in promoting positive emotions and reducing test anxiety.

Studies on Academic Emotions

Lei and Cui (2016) defined *academic emotions* as "students' emotional experiences related to the academic processes of teaching and learning, including enjoyment, hopelessness, boredom, anxiety, anger, and pride" (p. 1541). Based on arousal and enjoyment concepts, academic emotions have been divided into three categories: positive low arousal, negative low arousal, and negative high arousal (Artino & Jones, 2012). It is also argued that achievement emotions include prospective emotions, such as fear of failure, and retrospective emotions, such as shame, which learners experience after they receive feedback on their achievements (Pekrun et al., 2017).

Academic accomplishment serves as a commonly employed criterion for evaluating the effectiveness of educational systems, teachers, schools, and the success or failure of students. Consequently, scholars in this field have conducted empirical investigations to explore the causal link between students' academic emotions and academic achievements, as evidenced by a body of practical studies (Cocoradă, 2016; Kim & Hodges, 2012). However, the findings from these studies have been inconsistent. In general, positive emotions are anticipated to forecast favorable outcomes in academic settings, including high grades and commendable performance in both local and large-scale educational assessments (Villavicencio & Bernardo, 2013). Conversely, it is hypothesized that negative emotions will correlate with adverse consequences, such as lower grades and compromised performance in classroom activities and standardized examinations (Shen et al., 2023; Villavicencio, 2011).

Results of the meta-analysis undertaken by Lei and Cui (2016) showed support for Dong and Yu's (2010) Chinese version of the Academic Emotions Questionnaire, which was employed to evaluate the academic emotions of adolescents. Academic emotions have been linked to various variables, including cognitive activity, learning motivation, and strategies. Lei and Cui's (2016) meta-analysis revealed positive correlations between positive high arousal, positive low arousal, and academic achievement and negative correlations between negative high arousal, negative low arousal, and academic achievement. The study suggested that factors such as a student's age, regional location, and gender could moderate the effects of epistemic cognition on academic achievement.

Positive correlations have been shown between positive high arousal, positive low arousal, and academic achievement and negative correlations between negative high arousal, negative low arousal, and academic achievement.

Currently, scholars, both domestically and internationally, are directing their attention toward analyzing academic emotions in distance learners, resulting in noteworthy research outcomes (Pekrun, 2006). Cerniglia et al (2021) delved into the impact of screen time on emotion regulation and student performance. The study involved over 400 children over 4 years of age, examining their use of smartphones and tablets. The research analyzed the correlation between these behaviors, emotions, and academic performance, concurrently evaluating students' abilities and academic achievements. Similarly, Schlesier et al (2019) investigated the influence of early childhood emotions on academic preparation and social-emotional issues. Emotion regulation, the process of managing emotional arousal and expression, is crucial in determining children's adaptation to the school environment.

Moreover, Chen and Li (2012) integrated connectionist learning theory to devise an innovative distance education model. This model introduced educational content aligned with emotional education objectives and implemented the Mu class teaching mode, establishing a distance learning community and humanized network courses to address emotional shortcomings in the distance education process. Ensuring effectiveness, Pekrun et al. (2017) developed a hybrid virtual reality intelligent classroom system, incorporating television broadcasting and interactive space technology, to create a networked teaching environment. Teachers used diverse media, including video, audio, and text, to foster engagement and enhance communication between educators and students during the network teaching phase.

Artino and Jones (2012) introduced an emotion recognition algorithm based on facial expression scale-invariant feature transformation. This algorithm captures the facial expressions of distance learners, employing Scale Invariant Feature Transform (SIFT) feature extraction and expression recognition to address emotional gaps in the learning phase of distance education. Simultaneously, Turner and Schallert (2001) developed a learner emotion prediction model for an intelligent learning environment using a fuzzy cognitive map. This model facilitates extracting and predicting distance learners' emotional states, allowing real-time adjustments to the teaching approach based on anticipated emotions. Wang and Che (2005) contributed to the field by introducing the Distance Learner Emotion Self-Assessment scale, defining essential emotion variables, and establishing a distance learner emotion early warning model.

Drawing inspiration from the valuable contributions of the scholars mentioned earlier, Zembylas and McGlynn (2012) examined the academic emotions experienced by adults in online education. This investigation involved analyzing diverse influencing factors and exploring an environmental factor model within the online learning community, specifically focusing on academic emotional tendencies.

Building upon the insights derived from these scholars, our objective was to delve into the academic emotions of distance learners. To do this, we analyzed online learning behavior data to uncover meaningful findings in this domain.

Methodology

Sampling and Procedure

A cohort of 180 undergraduate students from y Zhejiang Industry & Trade Vocational College in China was randomly chosen during the second semester of 2023. The students' majors ranged from engineering and humanities to social sciences and law. The primary goals at the study's outset were to assess test anxiety and academic emotions. Following that, 80 students were randomly selected to participate in a 4-hour workshop covering the use of an AI-powered app (ChatGPT) in education. The participants ranged in age from 20 to 30, with a mean age of 24 (*SD* = 3.12). Interestingly, 25% of the students were over 27, 25% were between the ages of 22 and 27, and 50% were between 20 and 22. Over 16 weeks, this intervention group was closely observed via weekly Skype sessions to record and monitor their interactions with ChatGPT. Additionally, they were instructed to work on their homework and assignments using ChatGPT. The remaining 100 pupils were in the control group and did not receive special assistance. Scales measuring test anxiety and academic emotions were given to all 180 participants 2 weeks before the term final exams; 160 completed questionnaires were returned at the end of the study, resulting in a large dataset that could be analyzed. Statistical tests such as *t*-tests were used to identify significant differences between the control group and the experimental group receiving AI training. It is acknowledged that all ethical requirements were met, guaranteeing that the study complied with regulations about informed consent, confidentiality, and withdrawal rights.

Instruments

The researcher used the Westside Test Anxiety Scale, created by O'Driscoll and McAleese (2023), containing 10 items, to measure test anxiety. Participants chose answers using a 5-point Likert-type scale, with scores ranging from 1 to 5. The internal consistency of this scale was robust, with a Cronbach's alpha of 0.856, a mean of 32.5, and a standard deviation of 3.25. Pekrun (2006) developed and validated the Academic Emotion Questionnaires (AEQ) to assess academic emotions. The AEQ, which consists of 75 statements and 8 different emotions, was scored using a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree). These feelings had both positive and negative aspects. Among the positive feelings were joy (9 items), hope (5 items), and pride (8 items). Anger (10 points), boredom (11 points), shame (11 points), fear (11 points), and hopelessness (10 points) were the negative emotions. The eight categories of academic emotion in the current study had Cronbach's alpha coefficients ranging from 0.82 to 0.89, suggesting a high degree of internal consistency.

Data Analysis

Data were analyzed in different ways. First, descriptive statistics were estimated, including means and standard deviations for all pretests and the posttest. Then, the groups' scores on all variables were submitted to different independent samples *t*-tests. Moreover, Cohen's *d* for each *t*-test was calculated to determine the effect size for the treatment.

Results

Research Question 1

The first research question investigated the effect of AI-empowered educational applications on undergraduate students' test anxiety. The results of independent samples *t*-tests showed that the groups' mean scores on the test anxiety at the onset of the study were not statistically significant. Still, they had different test anxiety at the end of the semester. Results are presented in Table 1.

Table 1

Groups' Test Anxiety Before and After Treatment

Test group		Descriptive statistics		*t*-test			
		M	*SD*	*t*	*df*	*p*	Effect size
Pretest	Control group	32.5	3.25	1.29	158	.52	0.079
	Experimental group	32.23	3.56				
Posttest	Control group	31.26	2.89	16.26	158	.001	1.24
	Experimental group	27.25	3.5				

As seen in Table 1, the difference between the control group (M = 32.5, SD = 3.25) and experimental group (M = 32.23, SD = 3.56) on the pretest was not statistically significant (t = 1.29, df = 158, p = .52, d = 0.079). This suggests that at the onset of the study, the two groups had comparable levels of test anxiety. However, in contrast, the posttest results revealed a notable distinction between the control group (*M* = 31.26, *SD* = 2.89) and the experimental group (*M* = 27.25, *SD* = 3.5). The *t*-test for the posttest demonstrated a highly significant difference between the groups ($t = 16.26$, $df = 158$, $p < .001$, $d = 1.24$), indicating that the experimental intervention had a substantial impact on reducing test anxiety in the experimental group compared with the control group.

Research Question 2

The group's scores on the academic emotions test administered before and after the treatment are presented as follows.

As shown in Table 2, the mean scores of the control and experimental groups' scores on all academic emotions are similar. The results of independent sample *t*-tests verified that the differences between the groups' scores on all emotions were not statistically significant ($p > .05$).

Table 2

Groups' Academic Emotions Scores Before Treatment

Em**otion group**		Descriptive statistics		*t*-test			
		M	*SD*	*t*	*df*	*p*	Effect size
Enjoyment	Control group	3.71	0.60	1.29	158	.52	0.079
	Experimental group	3.62	0.56				
Hope	Control group	3.50	0.45	0.96	158	.46	0.054
	Experimental group	3.48	0.46				
Pride	Control group	3.30	0.62	0.83	158	.33	0.063
	Experimental group	3.33	0.58				
Anger	Control group	2.50	0.40	0.87	158	.29	0.058
	Experimental group	2.53	0.39				
Anxiety	Control group	2.60	0.54	1.12	158	.16	0.061
	Experimental group	2.63	0.53				
Hopelessness	Control group	2.56	0.46	1.23	158	.42	0.042
	Experimental group	2.60	0.39				
Shame	Control group	2.80	0.42	1.32	158	.36	0.08
	Experimental group	2.78	0.52				
Boredom	Control group	2.62	0.40	0.98	158	.41	0.06
	Experimental group	2.58	0.38				

In addition, the groups' scores on the academic emotions administered after the treatment were compared through independent samples *t*-tests. Results are presented in Table 3.

Table 3

Groups' Academic Emotions Scores After Treatment

Emotion group		Descriptive statistics		***t*-test**			
		M	*SD*	*t*	*df*	*p*	Effect size
Enjoyment	Control group	3.61	0.50	1.29	158	.001	1.53
	Experimental group	4.25	0.32				
Hope	Control group	3.60	0.40	0.96	158	.001	1.69
	Experimental group	4.20	0.30				
Pride	Control group	3.33	0.52	0.83	158	.001	1.87
	Experimental group	4.10	0.26				
Anger	Control group	2.50	0.44	0.87	158	.001	1.24
	Experimental group	1.78	0.56				
Anxiety	Control group	2.40	0.63	1.12	158	.001	1.36
	Experimental group	1.80	0.53				
Hopelessness	Control group	2.43	0.46	1.23	158	.001	1.40
	Experimental group	1.80	0.39				
Shame	Control group	2.60	0.42	1.32	158	.001	1.60
	Experimental group	2.00	0.52				
Boredom	Control group	2.52	0.40	0.98	158	.001	1.57
	Experimental group	21.90	0.38				

As seen in Table 3, the difference between the control group (M = 3.61, SD = 0.50) and the experimental group (M = 4.25, SD = 0.32) in enjoyment is statistically significant (t = 1.29, df = 158, p = .001, d = 1.53). Implementing the AI-empowered application had a substantial positive impact on the level of enjoyment experienced by the experimental group. Likewise, for hope, the observed difference between the control group (M = 3.60, SD =0.40) and the experimental group (M = 4.20, SD = 0.30) was statistically significant (t = 0.96, df = 158, p = .001, d = 1.69), underscoring the substantial improvement in hope facilitated by the AI-empowered application. A noteworthy difference was found in pride between the control group (M = 3.33, SD = 0.52) and the experimental group (M = 4.10, SD = 0.26) that was statistically significant (t = 0.83, df = 158, p = .001, d = 1.87), signifying the considerable enhancement in pride resulting from the experimental treatment. Moreover, the differences in anger (t = 0.87, df = 158, p = .001, d = 1.24), anxiety (t = 1.12, df = 158, p = .001, d = 1.36), hopelessness (t = 1.23, df = 158, p = .001, d = 1.40), shame (t = 1.32, df = 158, p = .001, d = 1.60), and boredom (t = 0.98, df = 158, p = .001, d = 1.57) all highlight statistically significant reductions in emotional states for the experimental group compared with the control group.

Discussion and Conclusion

The results indicate a significant reduction in test anxiety among undergraduate students after using an AI-powered educational app. At the start of the study, the first analysis revealed no statistically significant differences in test anxiety scores between the experimental and control groups. By the end of the semester, though, there was a noticeable difference in the two groups' test anxiety levels, with the experimental group reporting significantly lower test anxiety than the control group. This observed decrease in test anxiety among the experimental group aligns with existing research highlighting the potential benefits of incorporating AI and technology in educational settings. Kim and Hodges (2012) investigated the effects of an emotional control treatment on academic emotions, motivation, and achievement in an online mathematics course, emphasizing the interconnectedness of emotions and learning outcomes. Similarly, Ghafourian et al. (2020) used brain signal analysis to assess exam anxiety in healthy individuals, providing insight into possible physiological components of anxiety.

According to a review of the literature on AI in education found in several sources, including Picard and Healey's (1997) groundbreaking work on affective computing, technology can be used to personalize learning, adjust to the needs of each student, and improve overall engagement—all of which may help lower anxiety. Additionally, Wang et al. (2015) investigated how students' emotional experiences in computer-based learning environments are influenced by their cognitive-affective states, offering insights into how technology shapes students' learning experiences.

Another study by D'Mello and Graesser (2012) on the dynamics of affective states during complex learning lends credence to the idea that technology—including AI—can significantly impact how students feel about themselves. Furthermore, Calvo and D'Mello (2010) provided an interdisciplinary viewpoint on affect detection, which is pertinent when talking about how AI affects students' emotional states in learning environments.

Conversely, it is imperative to recognize that the present investigation may not be in direct accordance with all cited sources. Some references might not specifically address AI interventions, instead concentrating on other facets of technology in education or the study of emotions. Additionally, since the field is developing, more recent sources—like the instructional technology research of Moreno and Mayer (2005)—may present differing viewpoints regarding the efficacy of AI applications in lowering test anxiety.

The study's findings also show that AI-powered educational apps have a complex effect on students' emotional experiences, affecting positive and negative academic emotions. According to the findings, positive academic emotions including hope, pride, and enjoyment have significantly improved, and negative academic emotions like anxiety, shame, helplessness, anger, and boredom have decreased considerably. These outcomes align with existing research that underscores the potential of technology, including AI, to positively influence students' emotional states in educational settings.

Kim and Hodges (2012) underscored the interplay between emotions and learning outcomes by investigating the impact of an emotion control intervention on academic emotions, motivation, and achievement in an online mathematics course. Similarly, Ghafourian et al. (2020) advanced our understanding of emotional experiences in academic contexts through their analysis of exam anxiety via brain signals, revealing potential physiological components of anxiety.

As articulated by Picard and Healey (1997), the literature on affective computing provides a theoretical framework for understanding how technology can detect and respond to human emotions. This framework emphasizes the importance of considering affective dimensions in the learning process and can inform discussions on AI in education. Additionally, Baker et al. (2010) examined cognitive-affective states during interactions with computer-based learning environments, offering insights into how technology influences students' emotional experiences during the learning process. D'Mello and Graesser (2012) further supported the notion that technology, including AI, can significantly affect students' self-perceptions through their study on the dynamics of affective states during complex learning.

Calvo and D'Mello's (2010) review article provided an interdisciplinary perspective on affect detection, highlighting the relevance of technology in understanding and addressing emotional states in educational settings. The overarching goal of fostering a positive and supportive learning environment aligns with the observed reduction in negative academic emotions such as anxiety, shame, helplessness, anger, and boredom.

The literature reviewed, including the study by Pekrun (2006) on the control-value theory of achievement emotions, emphasizes the importance of addressing negative emotions to enhance overall academic performance and well-being.

The findings of this study resonate strongly with the principles of ODL, which seeks to employ technology to extend learning opportunities beyond traditional classroom boundaries, thereby democratizing education and promoting inclusivity and equity (Bates, 2015). The observed reduction in test anxiety among undergraduate students using AI-powered educational apps aligns with the essence of ODL, as it demonstrates how technology can remove barriers such as geographical, temporal, and socioeconomic constraints, making learning more accessible and accommodating diverse learner needs. Furthermore, the study's emphasis on the dynamic interplay between formal and informal learning environments and the promotion of lifelong learning and adaptability through distributed learning principles mirrors the approach taken in implementing AI-powered interventions to enhance students' emotional experiences and learning outcomes. Thus, the study's findings underscore the transformative potential of ODL paradigms facilitated by technology in promoting positive educational experiences and outcomes.

The implications of this study hold significant value for language teachers, developers of educational materials, and policymakers alike. For language teachers, the findings emphasize the potential of AI-powered educational applications in mitigating test anxiety and fostering positive academic emotions among undergraduate students. Integrating such technologies into language-learning environments can create more supportive and engaging settings conducive to enhanced learning outcomes. Materials developers can leverage these findings to design and adapt educational materials that incorporate AI-driven interventions, catering to diverse learner needs and promoting a more inclusive and effective learning experience. Policymakers can use this research to inform decisions regarding the integration of technology in education and the allocation of resources to support the development and implementation of AI-powered educational tools. By recognizing the benefits demonstrated in this study, language teachers, materials developers, and policymakers can collaborate to harness the potential of AI in education, ultimately improving the quality and accessibility of language-learning experiences.

To conclude, the results of this study align with the broader body of research indicating the possibility for AI-driven educational interventions to positively impact students' emotional experiences, particularly in terms of lowering exam anxiety. Nonetheless, it is critical to consider each study's unique characteristics, approaches, and settings in addition to recent developments in this dynamic field of study. Future studies should examine the complex implications of AI in the classroom and how it affects students' emotional health. Further investigation into the complex impacts of AI in education and its capacity to promote a healthy emotional environment is also necessary to improve learning outcomes and student well-being.

Acknowledgments

I would like to confirm that for editing, proofreading, and checking the language accuracy of the manuscript, I used AI-related applications. I would also like to thank all participants.

References

Aldosari, S. A. M. (2020). The future of higher education in the light of artificial intelligence transformations. *International Journal of Higher Education*, *9*(3), 145-151. https://files.eric.ed.gov/fulltext/EJ1248453.pdf

Ali, S. S., & Choi, B. J. (2020). State-of-the-art artificial intelligence techniques for distributed smart grids: A review. *Electronics*, *9*(6), Article 1030. https://doi.org/10.3390/electronics9061030

Amornkitpinyo, T., Yoosomboon, S., Sopapradit, S., & Amornkitpinyo, P. (2021). The structural equation model of using cloud learning for higher education students in the 21st century. *Journal of e-Learning and Knowledge Society*, *17*(1), 72–80. https://doi.org/10.20368/1971-8829/1135300

Anderson, T. (2016). Theories for learning with emerging technologies. In *Handbook of emerging technologies for learning* (pp. 89–103). Springer.

Artino, A. R., & Jones, K. D. (2012). Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning. *The Internet and Higher Education*, *15*(3), 170–175. https://doi.org/10.1016/j.iheduc.2012.01.006

Bahado-Singh, R. O., Vishweswaraiah, S., Aydas, B., Mishra, N. K., Guda, C., & Radhakrishna, U. (2019). Deep learning/artificial intelligence and blood-based DNA epigenomic prediction of cerebral palsy. *International Journal of Molecular Sciences*, *20*(9), 2075. doi: 10.3390/ijms20092075

Basaknezhad, S., Saeidi, R., & Mehrabizadeh, M. (2013). The study of perceived threat of test, cognitive test anxiety, looming maladaptive style, attributions to performance and fear of negative evaluation predictors of test anxiety in female high school students. *Journal of Educational Sciences*, *20*(1), 189–202. https://education.scu.ac.ir/article_10110.html?lang=en

Bates, A. W. (1997). The impact of technological change on open and distance learning. *Distance Education*, *18*(1), 93–109. https://doi.org/10.1080/0158791970180108

Bates, A. W. (2015). *Teaching in a digital age: Guidelines for designing teaching and learning*. BCcampus. https://openlibrary-repo.ecampusontario.ca/jspui/handle/123456789/276

Bisen, I. E., Arslan, E. A., Yildirim, K., & Yildirim, Y. (2021). Artificial intelligence and machine learning in higher education. In Z. Gulzar & A. Leema (Eds.), *Machine learning approaches for improvising modern learning systems* (pp. 1–17). IGI Global. https://doi.org/10.4018/978-1-7998-5009-0.ch001

Blake, R. J. (2009). The use of technology for second language distance learning. *The Modern Language Journal*, *93*, 822–835. https://doi.org/10.1111/j.1540-4781.2009.00975.x

Bozkurt, A., Akgun-Ozbek, E., Yilmazel, S., Erdogdu, E., Ucar, H., Guler, E., Sezgin, S., Karadeniz, A., Sen-Ersoy, N., Goksel-Canbek, N., Dincer, G.D., Ari, S., & Aydin, C. H. (2015). Trends in distance education research: A content analysis of journals 2009–2013. *The International Review of Research in Open and Distributed Learning*, *16*(1), 330–363. https://doi.org/10.19173/irrodl.v16i1.1953

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and applications. *IEEE Transactions on Affective Computing*, *1*(1), 18–37. https://doi.org/10.1109/T-AFFC.2010.1

Cerniglia, L., Cimino, S., & Ammaniti, M. (2021). What are the effects of screen time on emotion regulation and academic achievements? A three-wave longitudinal study on children from 4 to 8 years of age. *Journal of Early Childhood Research*, *19*(2), 145-160. https://doi.org/10.1177/1476718X20969

Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT—A case study. *Cogent Education*, *10*(1), Article 2210461. https://doi.org/10.1080/2331186X.2023.2210461

Chedrawi, C., Howayeck, P., & Tarhini, A. (2019). CSR and legitimacy in higher education accreditation programs, an isomorphic approach of Lebanese business schools. *Quality Assurance in Education*, *27*(1), 70–81. https://doi.org/10.1108/QAE-04-2018-0053

Chen, J. J., & Li, S. F. (2012). The paths of junior school students' achievement attribution and academic emotions forecasting their academic achievement. *Chinese Journal of Clinical Psychology*, *2012*(3), 392–394. https://caod.oriprobe.com/articles/29870105/The_Paths_of_Junior_School_Students__Achievement_Attribution_and_Acade.htm

Clark, J. T. (2020). Distance education. In E. Iadanza (Ed.), *Clinical engineering handbook* (2nd ed., pp. 410–415). Academic Press. https://doi.org/10.1016/B978-0-12-813467-2.00063-8

Cocoradă, E. (2016). Achievement emotions and performance among university students. *Bulletin of the Transilvania University of Braşov, Series VII: Social Sciences and Law, 9*(2-Suppl), 119–128. https://www.ceeol.com/search/article-detail?id=510917

Dhir, S. K., Verma, D., Batta, M., & Mishra, D. (2017). E-learning in medical education in India. *Indian Pediatrics*, *54*(10), 871–877. https://doi.org/10.1007/s13312-017-1152-9

Dieguez, T., Loureiro, P., & Ferreira, I. (2021, November 8–9). Entrepreneurship and leadership in higher education to develop the needed 21st century skills. In F. Bezzina (Ed.), *ECMLG 2021: 17th European Conference on Management, Leadership and Governance* (p. 143). Academic Conferences.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157. https://doi.org/10.1016/j.learninstruc.2011.10.001

Dong, Y., & Yu, G. L. (2010). Effects of adolescents' academic emotion on their academic achievements. *Journal of Psychological Science*, *33*(4), 934–937.

Dung, D. T. H. (2020). The advantages and disadvantages of virtual learning. *IOSR Journal of Research & Method in Education*, *10*(3), 45–48. https://doi.org/10.9790/7388-1003054548

Essien, A., Chukwukelu, G., & Essien, V. (2020). Opportunities and challenges of adopting artificial intelligence for learning and teaching in higher education. In M. Ali & T. Wood-Harper (Eds.), *Fostering communication and learning with underutilized technologies in higher education* (pp. 67–78). IGI Global. https://doi.org/10.4018/978-1-7998-4846-2.ch005

Fleming, S., & Hiple, D. (2004). Distance education to distributed learning: Multiple formats and technologies in language instruction. *CALICO Journal*, *22*(1), 63–82. http://www.jstor.org/stable/24149444

Garrison, D. R., Anderson, T., & Archer, W. (2003). A theory of critical inquiry in online distance education. In M. Moore and G. Anderson (Eds.), *Handbook of distance education* (pp. 113–127). Erlbaum. http://hdl.handle.net/2149/764

Ghafourian, P., Ghoshuni, M., & Vosoogh, I. (2020). Evaluation of exam anxiety in healthy subjects using brain signals analysis. *The Neuroscience Journal of Shefaye Khatam*, *8*(3), 61–69. https://doi.org/10.29252/shefa.8.3.61

Hayat, A. A., Salehi, A., & Kojuri, J. (2018). Medical student's academic performance: The role of academic emotions and motivation. *Journal of Advances in Medical Education and Professionalism*, *6*(4), 168–175. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6191829/

Horton, W. (2006). *E-learning by design*. Wiley Publishing, Inc.

Johnson, C. C., Walton, J. B., Strickler, L., & Elliott, J. B. (2023). Online teaching in K-12 education in the United States: A systematic review. *Review of Educational Research*, *93*(3), 353–411. https://doi.org/10.3102/00346543221105550

Khan, B. H. (2003). National virtual education plan: Enhancing education through e-learning in developing countries. *Quarterly of the Commonwealth Educational Media Centre for Asia*, *9*(1), 2–5.

Khan, S., Zaman, S. I., & Rais, M. (2022). Measuring student satisfaction through overall quality at business schools: A structural equation modeling. *South Asian Journal of Social Review*, *1*(2), 34–55. https://doi.org/10.57044/SAJSR.2022.1.2.2210

Kim, C., & Hodges, C. B. (2012). Effects of an emotion control treatment on academic emotions, motivation and achievement in an online mathematics course. *Instructional Science*, *40*, 173–192. https://doi.org/10.1007/s11251-011-9165-6

King, N. J., Ollendick, T. H., & Gullone, E. (1991). Test anxiety in children and adolescents. *Australian Psychologist*, *26*(1), 25–32. https://doi.org/10.1080/00050069108258829

Kistyanto, A., Rahman, M. F. W., Adhar Wisandiko, F., & Setyawati, E. E. P. (2022). Cultural intelligence increase student's innovative behavior in higher education: The mediating role of interpersonal trust. *International Journal of Educational Management*, *36*(4), 419–440. https://doi.org/10.1108/IJEM-11-2020-0510

Koçak, O., Koçak, Ö. E., & Younis, M. Z. (2021). The psychological consequences of COVID-19 fear and the moderator effects of individuals' underlying illness and witnessing infected friends and family. *International Journal of Environmental Research and Public Health*, *18*(4), 1836. https://doi.org/10.3390/ijerph18041836

Lei, H., & Cui, Y. (2016). Effects of academic emotions on achievement among mainland Chinese students: A meta-analysis. *Social Behavior and Personality: An International Journal*, *44*(9), 1541–1554. https://doi.org/10.2224/sbp.2016.44.9.1541

Li, J., Li, J., Yang, Y., & Ren, Z. (2021). Design of higher education system based on artificial intelligence technology. *Discrete Dynamics in Nature & Society*, *2021*, Article 3303160. https://doi.org/10.1155/2021/3303160

Lufi, D., & Awwad, A. (2013). Using the Minnesota Multiphasic Personality Inventory-2 to develop a scale to identify test anxiety among students with learning disabilities. *Learning Disability Quarterly*, *36*(4), 242–249. https://doi.org/10.1177/0731948712471199

Minkevics, V., & Kampars, J. (2021). Artificial intelligence and big data driven IS security management solution with applications in higher education organizations. In P. Chemouil (Ed.), *17th International Conference on Network and Service Management, CNSM 2021* (pp. 340–344). Institute of Electrical and Electronics Engineers. https://doi.org/10.23919/CNSM52442.2021.9615575

Moreno, R., & Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology*, *97*(1), 117–128. https://doi.org/10.1037/0022-0663.97.1.117

Neumann, M., & Baumann, L. (2021, October 13–16). Agile methods in higher education: Adapting and using eduScrum with real world projects. In *2021 IEEE Frontiers in Education Conference (FIE)* (pp. 1–8). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/FIE49875.2021.9637344

Ngah, A. H., Kamalrulzaman, N. I., Mohamad, M. F. H., Rashid, R. A., Harun, N. O., Ariffin, N. A., & Osman, N. A. A. (2022). The sequential mediation model of students' willingness to continue online learning during the COVID-19 pandemic. *Research and Practice in Technology Enhanced Learning*, *17*(1), Article 13. https://doi.org/10.1186/s41039-022-00188-w

Odhiambo, G. (2016). Higher education in Kenya: An assessment of current responses to the imperative of widening access. *Journal of Higher Education Policy & Management*, *38*(2), 196–211. https://doi.org/10.1080/1360080X.2016.1150551

O'Driscoll, D., & McAleese, M. (2023). The protective role of self-compassion on test anxiety among adolescents. *Pastoral Care in Education*, *41*(2), 211-224. https://doi.org/10.1080/02643944.2022.2054021

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, *36*(1), 36–48. https://doi.org/10.1016/j.cedpsych.2010.10.002

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development*, *88*(5), 1653–1670. https://doi.org/10.1111/cdev.12704

Picard, R. W., & Healey, J. (1997). Affective wearables. *Personal Technologies*, *1*, 231–240. https://doi.org/10.1007/BF01682026

Piroozmanesh, A., & Imanipour, M. (2018). The effect of formative evaluation on test anxiety of nursing students. *Journal of Medical Education Development*, *10*(28), 18–26. https://doaj.org/article/7b6baa79dda7411c8f54af86e3cf3dc4

Pradana, M., Elisa, H. P., & Syarifuddin, S. (2023). Discussing ChatGPT in education: A literature review and bibliometric analysis. *Cogent Education*, *10*(2), Article 2243134. https://doi.org/10.1080/2331186X.2023.2243134

Roa, C. C., Matutes, K. C. B., & Ombajen, J. D. (2022). Learning management: A review on emerging technologies for distance education. *World Wide Journal of Multidisciplinary Research and Development*, *8*(03), 121–125. https://doi.org/10.17605/OSF.IO/WVB45

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, *6*(1), 342–363. https://doi.org/10.37074/jalt.2023.6.1.9

Rumble, G. (1989). "Open learning," "distance learning," and the misuse of language. *Open Learning: The Journal of Open, Distance and e-Learning*, *4*(2), 28–36. https://doi.org/10.1080/0268051890040206

Salmalian, H., Maleki Pirbazari, M., & Salehi, S. (2020). The relationship early maladaptive schemas of students with their social anxiety and academic burnout (a canonical correlation). *Journal of Educational Sciences*, *27*(1), 183–202. https://doi.org/10.22055/edus.2020.32778.3005

Schlesier, J., Roden, I., & Moschner, B. (2019). Emotion regulation in primary school children: A systematic review. *Children and Youth Services Review*, *100*, 239-257. https://doi.org/10.1016/j.childyouth.2019.02.044

Shen, B., Wang, Y., Yang, Y., & Yu, X. (2023). Relationships between Chinese university EFL learners' academic emotions and self-regulated learning strategies: A structural equation model. *Language Teaching Research*. Advance online publication. https://doi.org/10.1177/13621688221144832

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, *2*(1), 3–10.

Talib, N., & Sansgiry, S. S. (2012). Determinants of academic performance of university students. *Pakistan Journal of Psychological Research*, *27*(2), 365–378. https://psycnet.apa.org/record/2014-02064-007

Turner, J. E., & Schallert, D. L. (2001). Expectancy–value relationships of shame reactions and shame resiliency. *Journal of Educational Psychology*, *93*(2), 320–329. https://doi.org/10.1037/0022-0663.93.2.320

Villavicencio, F. T. (2011). Critical thinking, negative academic emotions, and achievement: A meditational analysis. *The Asia-Pacific Education Researcher*, *20*(1), 118–126. https://animorepository.dlsu.edu.ph/faculty_research/1277

Villavicencio, F. T., & Bernardo, A. B. (2013). Positive academic emotions moderate the relationship between self-regulation and academic achievement. *British Journal of Educational Psychology*, *83*(2), 329-340.

Wang, A. P., & Che, H. (2005). A research on the relationship between learning anxiety, learning attitude, motivation and test performance. *Psychological Development and Education*, 2005(1), 55–59, 86. https://caod.oriprobe.com/articles/8589649/A_Research_On_the_Relationship_Between_Learning_Anxiety_Learning_Attit.htm

Wang, R., Ryu, H., & Katuk, N. (2015). Assessment of students' cognitive-affective states in learning within a computer-based environment: Effects on performance. *Journal of Information and Communication Technology, 14*, 153-176. DOI:10.32890/jict2015.14.0.8161

Wenge, M. (2021). Artificial intelligence-based real-time communication and AI-multimedia services in higher education. *Journal of Multiple-Valued Logic and Soft Computing*, *36*(1–3), 231–248. https://dblp.org/rec/journals/mvl/Ma21.html

Yavorsky, K. (2017). *Academic emotion and self-efficacy impacting sense of math class belonging in college students* (Publication no. 2415) [MA thesis, Rowan University]. Rowan University Theses and Dissertations. https://rdw.rowan.edu/etd/2415

Yazdani, Z., & Asadi, F. (2022). Investigating the relationship between virtual education and exam anxiety mediated by academic engagement. *Technology of Instruction and Learning*, *5*(15), 57–76. https://doi.org/10.22054/jti.2023.72129.1361

Zembylas, M., & McGlynn, C. (2012). Discomforting pedagogies: Emotional tensions, ethical dilemmas and transformative possibilities. *British Educational Research Journal*, *38*(1), 41-59. https://www.jstor.org/stable/23211443

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

Artificial Intelligence in Education: A Bibliometric Study on Its Role in Transforming Teaching and Learning

Gürhan Durak[1], Serkan Çankaya[2], Damla Özdemir[1], and Seda Can[1]
[1]Balıkesir University; [2]Izmir Democracy University

Abstract

This study aimed to present a comprehensive bibliometric analysis of 1,726 academic studies from among those indexed by the Web of Science database platform between 2013 and 2023, to provide a general framework for the concept of artificial intelligence in education (AIEd). Trends in publications and citations across countries, institutions, academic journals, and authors were identified, as well as collaborations among these elements. Several bibliometric analysis techniques were applied, and for each analysis, the motivations behind the execution and method of producing findings were documented. Our findings showed that the number of studies on the concept of AIEd has increased significantly over time, with the U.S. and China being the most common countries of origin. Institutions in the U.S. stand out from those around the world. Pioneering journals in education have also emerged as prominent in the field of AIEd. On the other hand, collaboration between authors has been limited. The study was supplemented with keyword analysis to reveal thematic AIEd concepts and to reflect changing trends. For those exploring artificial intelligence in education, our insights on popular topics offer valuable guidance toward greater understanding of the latest advancements and key research areas.

Keywords: artificial intelligence, bibliometric analysis, bibliographic coupling, co-authorship analysis, co-citation analysis, co-occurrence analysis

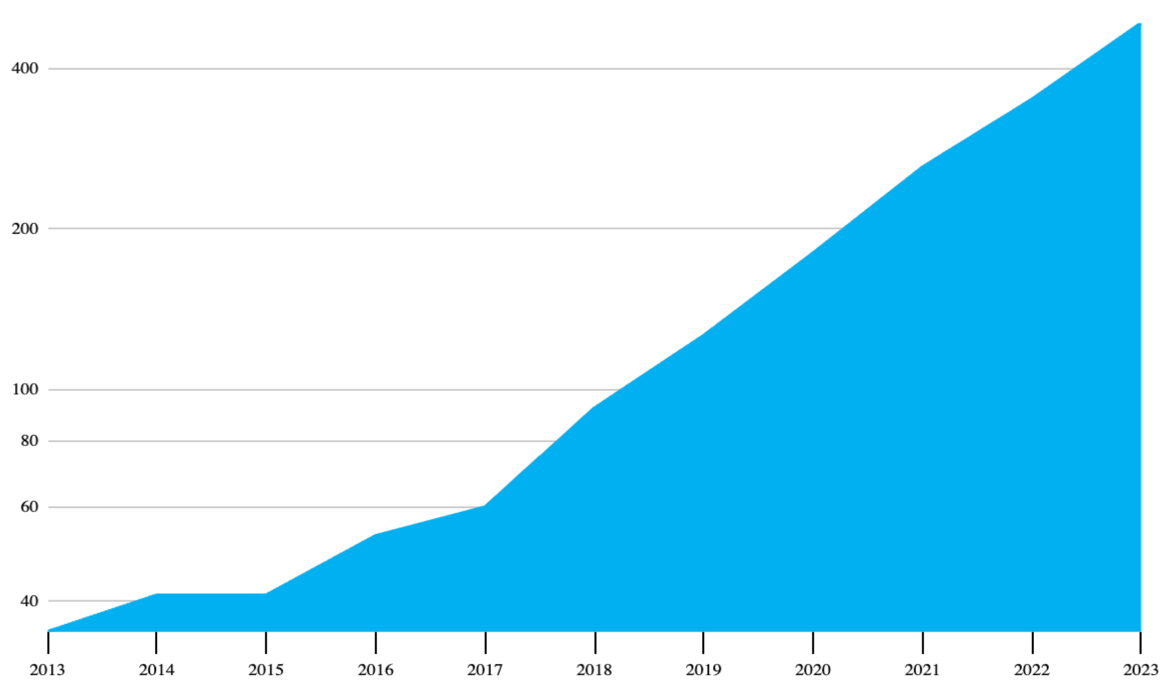
Artificial Intelligence in Education: A Bibliometric Study on Its Role in Transforming Teaching and Learning

Artificial intelligence (AI) represents a broad domain within computer science dedicated to the development of smart machines capable of undertaking tasks akin to those performed by humans (Bartneck et al., 2021; Joiner, 2018). The objective is to imbue computers with intelligence, enabling them to think and learn through programmed algorithms or to emulate human thought processes and actions. With recent technological advancements, it is reasonable to say that AI has acquired the capability to execute operations far beyond the rapid processing capacities of the human mind (Khanam et al., 2021). Understanding AI is crucial if we aim to integrate it meaningfully into society, and this involves a deep dive into what we mean by AI, its developmental trajectory, and its current standing (Bozkurt, 2023).

One of the transformative impacts of AI in our lives has manifested in the education sector. AI has swiftly emerged as an instrumental force there, paving the way for a new era of personalized learning, enhanced engagement, and data-driven insights to foster an enriched educational experience and optimize academic outcomes (Chaudhry & Kazim, 2022). According to the Horizon Report 2023, the following two concepts are at the top of the key technologies and practices section: (a) AI-enabled applications for predictive personal learning and (b) generative AI (Pelletier et al., 2023). Figure 1 illustrates the number of studies conducted in the field of AI in education (AIEd) in the Web of Science (WoS) database platform over the last 10 years.

Figure 1

Number of Academic Publications on AIEd in the Web of Science



According to Figure 1, the concept of AIEd has become the center of attention for many researchers in recent years, and there has been a great increase in the number of studies. This increasing number of studies has made it difficult to follow the research in the field. The bibliometric analysis method can be used for the follow-up and detailed analysis of research in a particular field and can be enriched with various graphics. The aim of this study was to explore AIEd, using the bibliometric analysis method to evaluate research developments comprehensively and methodically. This study sought to provide a thorough understanding of the AIEd research field through a bibliometric analysis on articles indexed in WoS. Concentrating on AIEd, this study addressed the following research questions. In the AIEd literature in the WoS database platform:

1. What is the distribution of leading countries and institutions?
2. What are the patterns of research connections and collaborations among countries and institutions?
3. What are the leading journals and authors?
4. What patterns of citation networks can be observed among leading journals?
5. What does analysis reveal about the nature of author collaborations and the impact of co-citation among prominent researchers in this field?
6. Which topics are most prevalent, and how are they interconnected, as indicated by the analysis of commonly used keywords?

The Importance of Bibliometric Analysis of Artificial Intelligence in Education

Bibliometric analysis plays a key role in providing an in-depth understanding of scientific research. The importance of the parts of the bibliometric analysis method is explained as follows. The number of publications shows the growth rate of the research, while citation and co-citation analyses reveal the most influential studies and authors in the field. Keyword analyses identify the key topics on which research has focused. Journal analyses show which journals have dominated each field, while geographic distribution reveals which regions or countries have been more active. Research trends indicate which topics are on the rise or ignored, and network analysis visualizes the relationships between different authors, institutions, and topics. In other words, bibliometric analysis is a method that can be used to comprehensively assess the current state, impact, and potential future directions of scientific research.

A bibliometric analysis of AIEd offers a systematic, quantitative, and insightful examination of the scholarly landscape, elucidating prevalent trends, key contributors, and emergent areas of interest within this interdisciplinary domain (Donthu et al., 2021; Ho, 2008). As the education sector grapples with the challenges and promises of AI integration, a bibliometric analysis provides evidence-based insights to educators, developers, and policymakers (Argente et al., 2023). By showcasing where we have been and indicating where we might go, such an analysis serves as both a historical record and a strategic compass, ensuring that AI's incorporation into education is thoughtful, research-informed, and optimized for pedagogical efficacy (Gavira-Marin et al., 2018; Yin, 2013). A bibliometric analysis of AIEd is not merely an academic exercise; it is a critical tool in comprehending the intricacies of a rapidly evolving research domain (Ellegaard & Wallin, 2015; Moral-Munoz et al., 2020). By providing clarity, direction, and insight, such a study enriches the scholarly community's collective understanding and paves the way for impactful and informed innovations in the intersection of AI and education (Martins et al., 2022).

In general, bibliometric analyses has been important in evaluating the current and future status of scientific research. In addition, since the results are presented objectively, they have been free from researchers' biases. However, search criteria have not been given enough importance in most such studies, raising doubts about whether the publications included because of the search criteria fully reflected the relevant concept. One of the most powerful aspects of this study was that it analyzed all systematic review, content analysis, and bibliometric analysis studies published in the relevant field, and analyzed the keywords used in those studies. The search criteria for this study were developed in an appropriate way which was explained in detail in method section.

Literature Review

Table 1 presents the bibliometric analysis and systematic reviews related to AIEd in the literature. The table provides brief information about (a) author(s) of the systematic review studies; (b) the databases where the publications were obtained for systematic review in these studies; (c) the total number of publications reviewed in these studies; and (d) the number of citations for the studies. This curated list serves as a testament to the increasing prominence and relevance of AIEd.

Table 1

Bibliometric and Review Studies in the Field of AIEd

Author(s)	Database	Number of publications reviewed	Number of citations
Zawacki-Richter et al. (2019)	EBSCO Education Source, WoS, Scopus	146	1,302
Hinojo-Lucena et al. (2019)	WoS, Scopus	132	181
Prahani et al. (2022)	Scopus	457	26
Tang et al. (2023)	WoS	86	154
Durso & Arruda (2022)	Brazilian Digital Library of Dissertations and Theses	63	3
Hwang & Tu (2021)	WoS, Scopus	129	127
Sapci & Sapci (2020)	PubMed, IEEE, CINAHL, Plus, ScienceDirect	76	94

Liang et al. (2023)	WoS	71	51
Baek & Doleck (2020)	WoS	135	18
Salas-Pilco & Yang (2022)	WoS, IEEE Xplore, Scielo, CAPES	31	52
Chiu (2021)	WoS, Scopus, ERIC	45	36
Celik et al. (2022)	ProQuest, ERIC, WoS	44	86
Salas-Pilco et al. (2022)	WoS, ScienceDirect, IEEE	30	34
Xu & Ouyang (2022)	WoS, Science Direct, Scopus, IEEE, EBSCO, ACM, Taylor & Francis, Wiley	63	60
Mohamed et al. (2022)	ScienceDirect, Scopus, Springer Link, ProQuest, EBSCO Host	20	7
García-Martínez et al. (2023)	WoS, Scopus	25	8
Pua et al. (2021)	WoS, Scopus, Google Scholar	135	3
Kaban (2023)	Wos	1,153	0
Jia et al. (2023)	Wos, Scopus	76	0

Note. Citation values taken from Google Scholar in December 2023.

Based on the data in Table 1, the following themes are evident in recent publications exploring the multifaceted applications and implications of AIEd.

General Trends in Higher Education

Zawacki-Richter et al.'s (2019) study delved into the role of educators in the rapidly developing field of AI applications in higher education. Similarly, Hinojo-Lucena et al. (2019) conducted a bibliometric study the same year, evaluating AI's influence on higher education through a thorough analysis of scientific literature. Finally, Prahani et al. (2022) and Baek and Doleck (2020) comprehensively examined the general trends and impacts of AI in higher education. These studies provided overarching insights into AI's growth in academia and its potential implications for educators and students.

Focus on Specific Educational Domains

Several studies narrowed down AI's application to specialized educational domains. Complementing these, Jia et al. (2023) presented a bibliometric analysis and content analysis that examined the significant role of AI in science education at the primary and secondary levels, and its growing influence over the past decade. For instance, Hwang and Tu (2021) mapped AI's roles and research trends in mathematics education. Liang et al. (2023) explored the fusion of AI with language education. Chiu (2021) ventured into the intersection of emerging technologies, including AI, in the context of chemical education. Finally, the studies by García-Martínez et al. (2023) and Pua et al. (2021) were included in this category as they examined the impact and trends of AI in specific educational fields. These contributions demonstrated the versatile nature of AI and its adaptability to cater to various academic disciplines.

Regional and Specific Case Analyses

Salas-Pilco and Yang (2022) conducted a targeted review of AI applications in Latin American higher education. Another work by Durso and Arruda (2022) delved into AI's impact on distance education within Brazilian studies. Similar to the other studies mentioned above, it can be said that Kaban's (2023) study focused on specific regions or situations. By focusing on particular regions or cases, these papers offered unique perspectives, addressing localized challenges and potentials of AI in education.

AI's Role in Analyzing and Enhancing Pedagogy

Celik et al. (2022) and Salas-Pilco et al. (2022) investigated the promises, challenges, and roles AI plays in teacher education, and how it has intertwined with learning analytics. Their findings emphasized the transformative power of AI in reshaping pedagogical strategies and aiding educators in their teaching processes.

Method

This study used bibliometric analysis to investigate the vast landscape of academic literature related to AIEd. This section delineates the methodological framework, databases included, criteria for including and excluding publications, and the analytical tools employed to interpret the data.

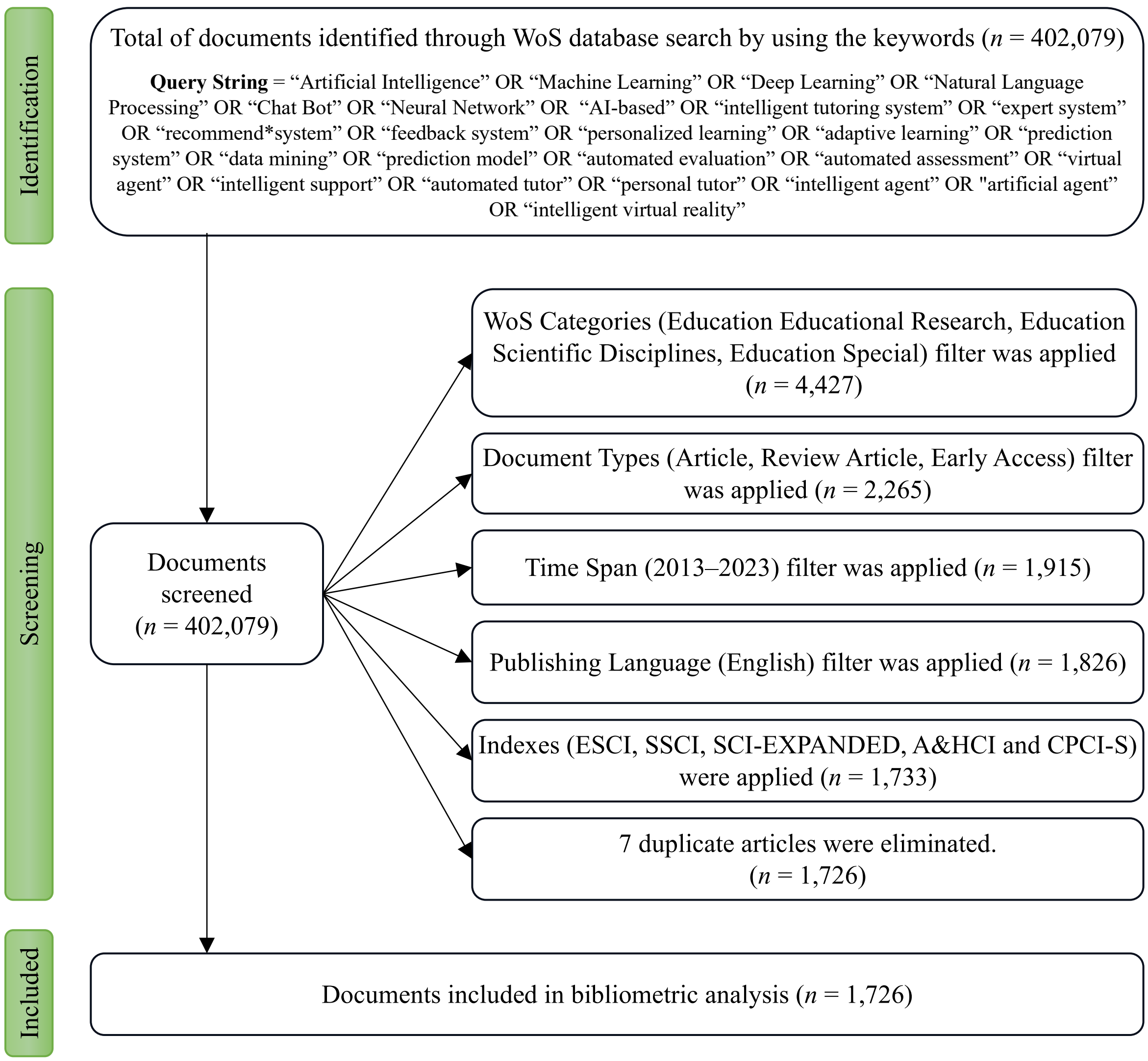
Determination of Studies

This bibliometric analysis was carried out using WoS, a highly esteemed and comprehensive research platform that covers a wide range of disciplines. Using WoS for a bibliometric study ensured a methodical approach to quantitatively analyze the academic literature in terms of publications and citations. Conducting a bibliometric study using the WoS database ensured access to high-quality, peer-reviewed journals and publications, which offered a credible and reliable overview of the research landscape. Additionally, WoS offered robust citation tracking; this enabled researchers to effectively trace the impact and evolution of research trends in the field. Given its comprehensive nature and the emphasis on citation data, the WoS database platform was particularly well suited for bibliometric analyses, ensuring a rigorous examination of the topic within the context of established academic scholarship.

Figure 2 outlines the data collection process and the key search terms that were identified after a thorough examination of bibliometric analysis and systematic review studies in the domain of AI using the PRISMA method. These search terms encompassed a wide spectrum of AI-related concepts and tools, and also offered a robust framework for extracting relevant publications from academic databases.

Figure 2

PRISMA Flowchart for Bibliometric Analysis



To maintain clarity and precision in our bibliometric analysis, it was essential to define clear criteria for including and excluding research papers. The inclusion criteria were adopted to ensure that the selected publications aligned with the study's objectives and maintained a consistent standard of quality and relevance.

Data Analysis

This study employed the VOSviewer program to analyze information from studies obtained from the WoS database, categorizing the data into various types. The findings included graphs that were generated using a total of five different networks of analyses, explained in detail in the findings section:

- Bibliographic coupling assessed the overlap in references between AIEd papers, indicating research connections.
- Co-authorship analysis focused on analyzing the collaborations between authors and countries in AIEd research.
- Citation analysis was employed to determine the most frequently referenced journals in the field of AIEd.
- Co-citation analysis identified frequently co-cited AIEd studies, revealing influential research relationships.
- Co-occurrence analysis in which keywords from the articles were selected and classified to illustrate the most popular topics and their connections.
-

Findings

The results from the bibliometric analysis of the articles retrieved from the WoS database platform are presented here under headings that align with the research questions.

Leading Countries and Institutions

To identify the countries targeted in the examined studies, country information was retrieved from the WoS database and depicted as bubble charts using an online tool in venngage.com. Figure 2 displays the distribution of these articles by country.

Figure 3

Bubble Graph of the Distribution of Selected Articles by Country

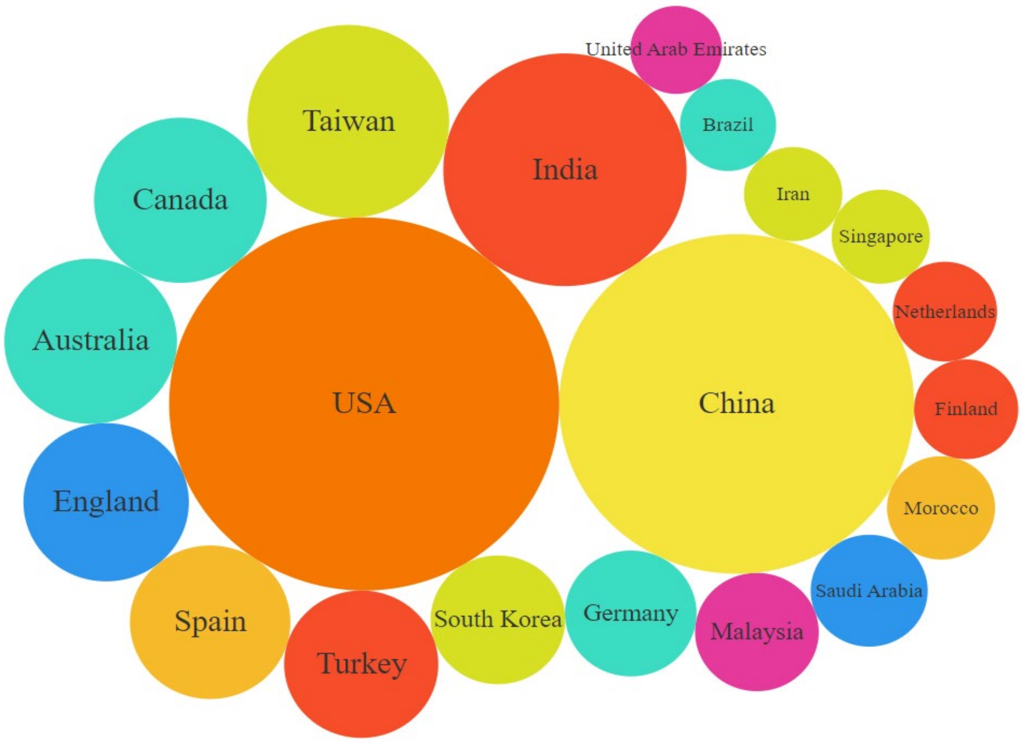
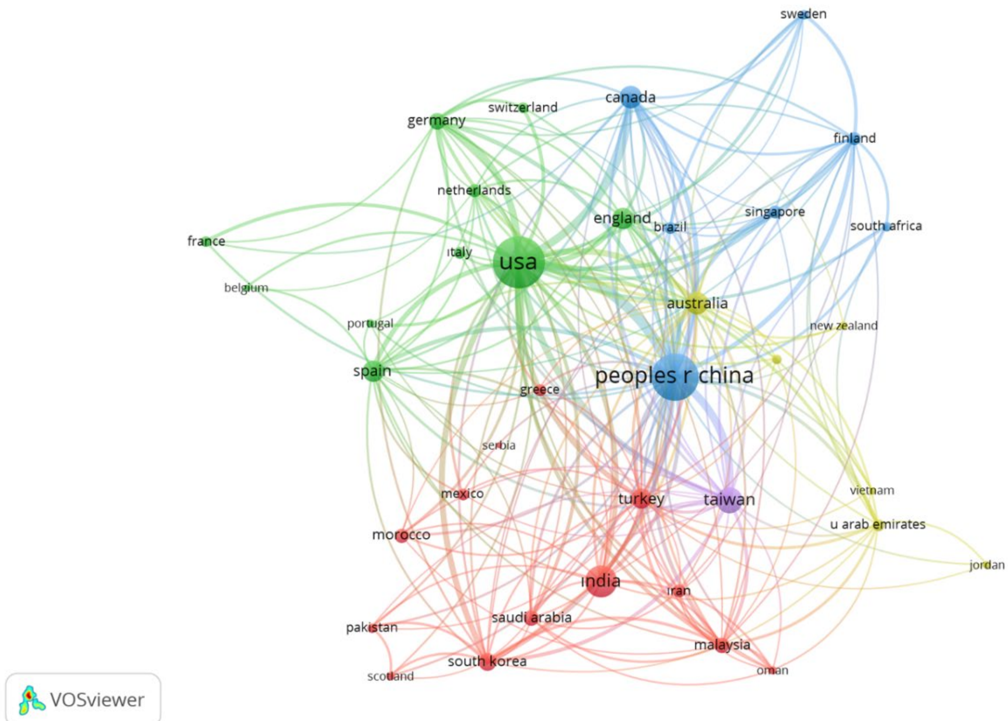


Figure 3 shows 20 countries that have published a minimum of 16 articles. As per the illustration, the U.S. led with 378 articles, followed by China with 313, and India with 147. In total, researchers from 95 distinct countries contributed to 1,726 articles.

Figure 4 displays the bibliographic coupling of the countries with network visualization, offering a comprehensive view of the interconnections among citing publications, which helped to trace the thematic evolution and current advancements in AIEd. As the condition we set, a country must have had a minimum of two documents and 100 citations to be included. Out of 95 countries, 37 met this criterion. For all of the countries, the number of publications, the number of citations, and total link strength (TLS), which represents the number of cited references that two countries share, were calculated.

Figure 4

Bibliographic Coupling of Countries

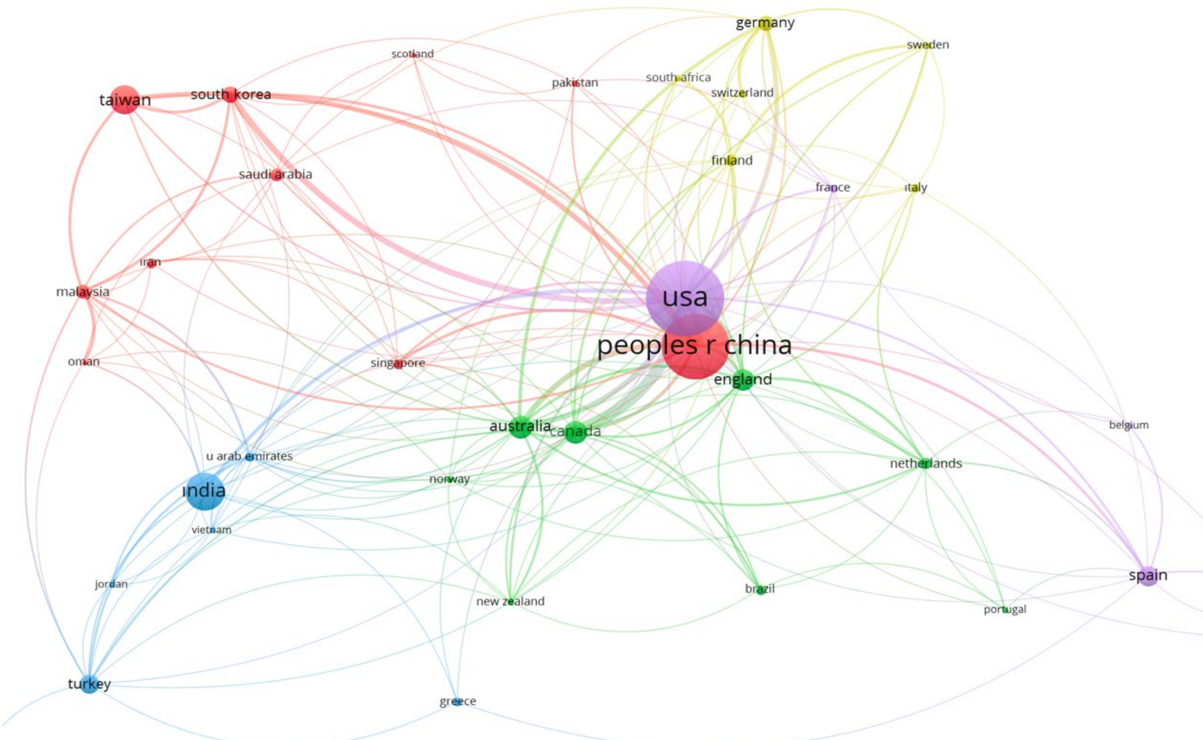


In our analysis of the bibliographic coupling ranked by the number of citations from each country, the U.S. topped the list with 4,533 citations, China came second with 2,513 citations, and Taiwan was third with 1,574 citations. Regarding the highest TLSs, the U.S. dominated with a link strength of 31,570, China followed with 24,166, Taiwan had 12,145, and Turkey had 7,984. Distinct colors represent various clusters that were more commonly interconnected. The line between any two circles indicates that papers from those two countries had similar citations in their reference list. The thickness of the lines shows a greater bibliographic coupling between the countries (Van Eck & Waltman, 2014). Large circles show the dominance of the countries in terms of citations. The green cluster, one of the big clusters, included China and Taiwan. The other big cluster comprised the U.S., England, Germany, Netherlands, Portugal, Spain, Canada, Italy, Switzerland, France, and Belgium. In the third cluster, we found Australia, the People's Republic of China, and Portugal.

Figure 5 represents the co-authorship network of countries. Each circle in the figure represents the country of an author, with the size of the circle indicating the number of their publications. Lines between circles signify the network of collaboration, with thicker lines indicating more intense collaboration. Various clusters are represented by distinct colors to denote similar research areas (van Eck & Waltman, 2018). As the condition we set, the analysis required a minimum of two documents and 100 citations per country, and 37 countries met this criterion.

Figure 5

Co-Authorship Networks of Countries

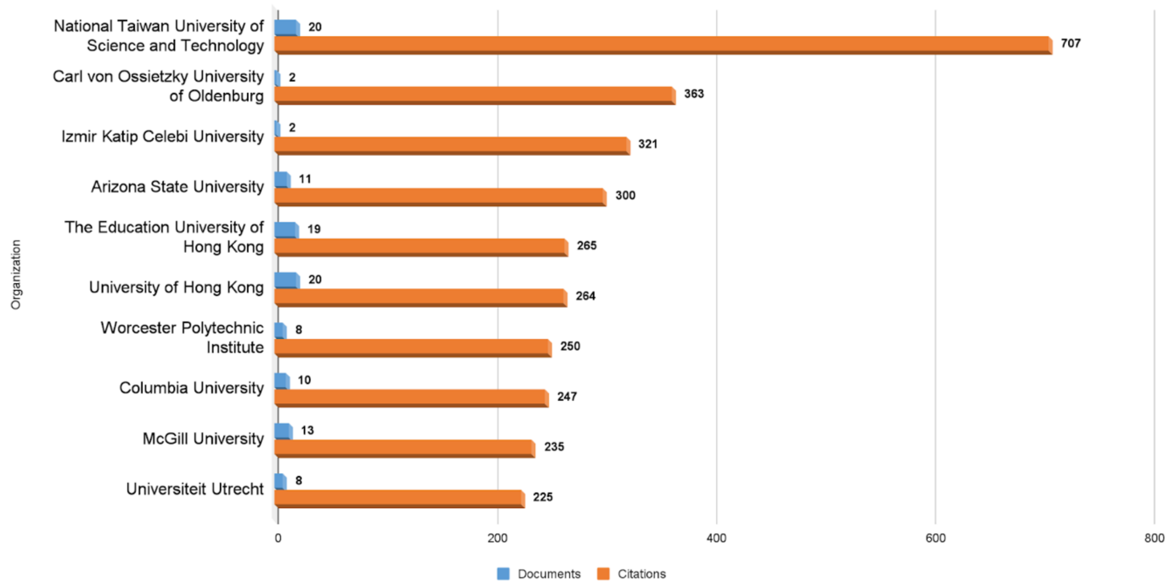


The results revealed 5 clusters through 141 connections. The U.S. emerged as the country with the most frequent collaborations with 113 total link strength, particularly strong with China. Both China and Australia, China and Taiwan, U.S. and South Korea also showed high levels of international collaboration. Specific clusters, such as China, Taiwan, and South Korea (Cluster 1); U.S. and Spain, (Cluster 2); and India and Turkey (Cluster 3) were noted for having similar research focuses. The map provides a detailed view of the collaboration patterns among these and other countries.

In the context of this study, which aimed to map out the key players in the field, Figure 6 plays a crucial role. It displays the leading institutions based on the authors' affiliations, offering insights into which academic and research organizations were most prominently represented in this area of research.

Figure 6

Number of AIEd Publications and Citations of Top 10 Institutions



According to Figure 6, National Taiwan University of Science and Technology was the leading institution with 20 article and 707 citations. Based on the number of citations, it was followed by Carl von Ossietzky University of Oldenburg, İzmir Katip Çelebi University, Arizona State University, and so on. When the number of citations per article was evaluated, Carl von University was prominent. While Figure 6 highlights the leading institutions based on article count and citations, Figure 7 shifts the focus to the interconnectedness of these institutions, showcasing the bibliographic coupling based on the authors' affiliations, which includes only those institutions with at least two articles and 100 citations.

Figure 7

Bibliographic Coupling of Institutions

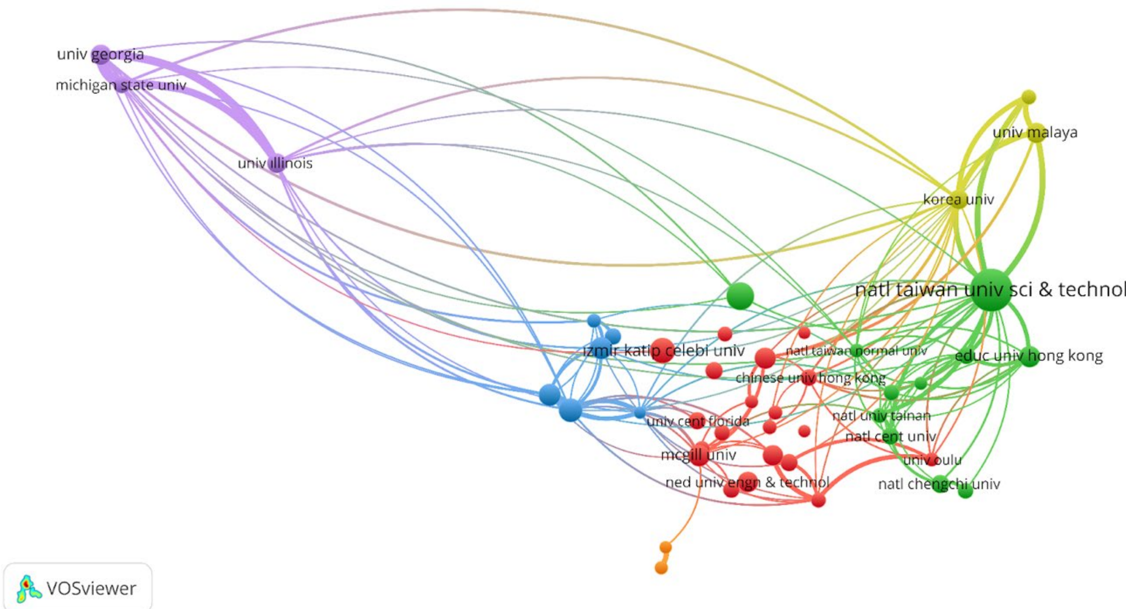


Figure 7 identifies six distinct clusters, each represented by different colors. These clusters highlight key terms such as the number of publications (NP), the number of citations (NC), and the TLS. Out of 2,003 institutions, only 54 had bibliographic coupling ties. The institutions were mapped based on the authors' affiliations and ranked by their total citation count. Considering total link strength values apart from total citation values, the leading institutions can be listed as follows:

- University of Georgia (NP = 11, NC = 224, TLS = 2,580) part of the purple cluster.
- Michigan State University (NP = 11, NC = 184, TLS = 2,282) part of the purple cluster.
- National Taiwan University of Science and Technology (NP = 20, NC = 707, TLS = 2,007) part of the green cluster.
- University of Illinois (NP = 12, NC = 217, TLS = 1,633) part of the purple cluster.
- University of Hong Kong (NP = 20, NC = 264, TLS = 1,460) part of the red cluster.

Leading Journals and Authors

The 1,726 articles examined appeared in 291 unique journals. When these journals were ranked by their publication count, 147 of them had just one article each. A list of the top 10 journals can be found in Table 2.

Table 2

The 10 Most-Cited Journals Regarding AIEd Research (2013–2023)

Journal	NP	NC	TLS
Computers & Education	43	2,131	117
Education and Information Technology	154	1,380	159
International Journal of Emerging Technologies in Learning	116	884	65
Educational Technology & Society	39	907	69
Computer Applications in Engineering Education	43	699	34
International Journal of Educational Technology	21	596	9
Interactive Learning Environments	63	629	69
British Journal of Educational Technology	30	507	33
IEEE Transactions on Learning Technologies	27	362	13
Technology Knowledge and Learning	12	315	31

Table 2 reveals that *Educational Technology & Society* led with 154 articles, followed by *International Journal of Emerging Technologies in Learning* with 116 articles, and *Interactive Learning Environments* with 63 articles. When considering citation counts, *Computers & Education* topped the list with 2,131 citations from 43 articles, *Education and Information Technology* has 1,380 from 154 articles, and *Educational Technology & Society* received 907 citations from 39 articles.

The research examined the citation network map of leading journals as illustrated in Figure 8, focusing on those with at least 2 articles and 100 citations to ensure the inclusion of publications with significant scholarly impact. This citation analysis was employed to evaluate the influence and prestige of these journals, reflected by the frequency of citations they received within the academic community, as a measure of their contribution to the field.

Figure 8

Citation Network Map of Leading Journals

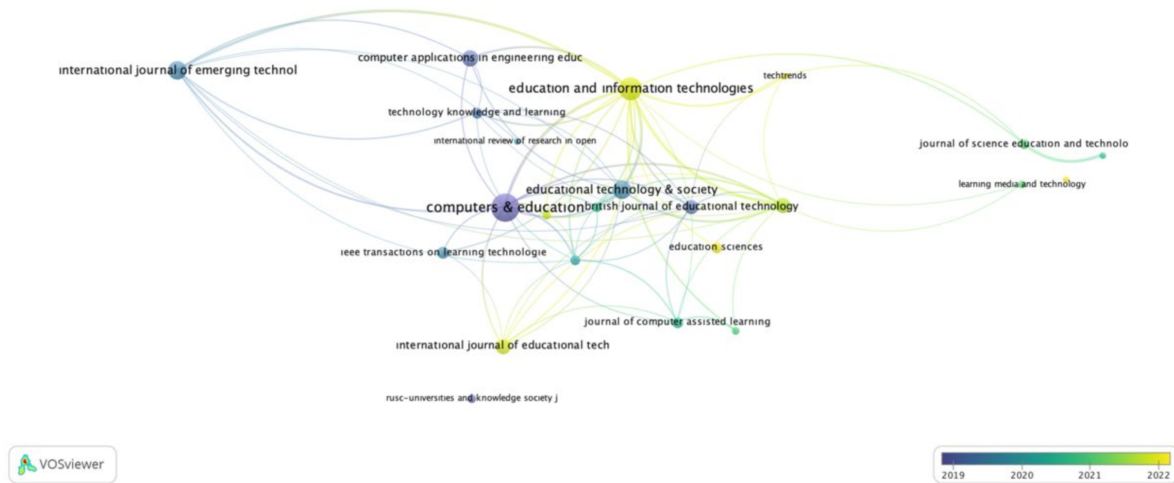


Figure 8 identifies 10 distinct clusters, each represented by a different color. These clusters highlight key terms such as the number of documents, the number of citations, and the total link strength. Out of 348 journals, only 193 had citation ties. *Computers & Education, Education and Information Technology, Interactive Learning Environments*, and *Educational Technology & Society* also stood out in the citation rankings, maintaining robust citation connections with numerous other journals. When evaluated in terms of popularity (shown in yellow), we observed that journals such as *Education and Information Technologies, Interactive Learning Environments,* and *International Journal of Educational Technology in Higher Education* were prominent.

Following the analysis of academic journals, the study presented a ranking of the top 10 authors based on citation counts, and subsequently examined the network of collaborations among authors. Table 3 below highlights the 10 most prominent authors based on their citation numbers.

Table 3

Top 10 Authors by the Number of Citations

Author	NP	NC
Hwang, G.-J.	14	542
Papamitsiou, Z.	2	346
Onan, A.	3	345
Baker, R. S.	3	224
Xie, H.	4	224
Zhai, X.	8	178
Hew, K. F.	3	177
Qiao, C.	2	175
Tang, Y.	2	175
Chu, H.-C.	2	171

Table 3 shows that Hwang, G.-J. was notable with 14 articles and 542 citations. Out of the 77 authors who surpassed the criteria of at least 2 articles and 100 citations, 29 were part of an affiliated network.

Building on the identified leading scholars, the subsequent phase of the study employed co-authorship analysis to delve into the broader landscape of intellectual collaboration among researchers. Co-authorship analysis, an essential tool for understanding intellectual partnerships among researchers, is employed to reveal how scholars interact and contribute collectively. Figure 9, a co-authorship network map, visually interprets these relationships, highlighting significant connections among 828 authors based on specific inclusion criteria. Inclusion in the map required authors to have authored at least two documents and received 100 citations, a criterion met by only 29 authors.

Figure 9

Co-Authorship Analysis

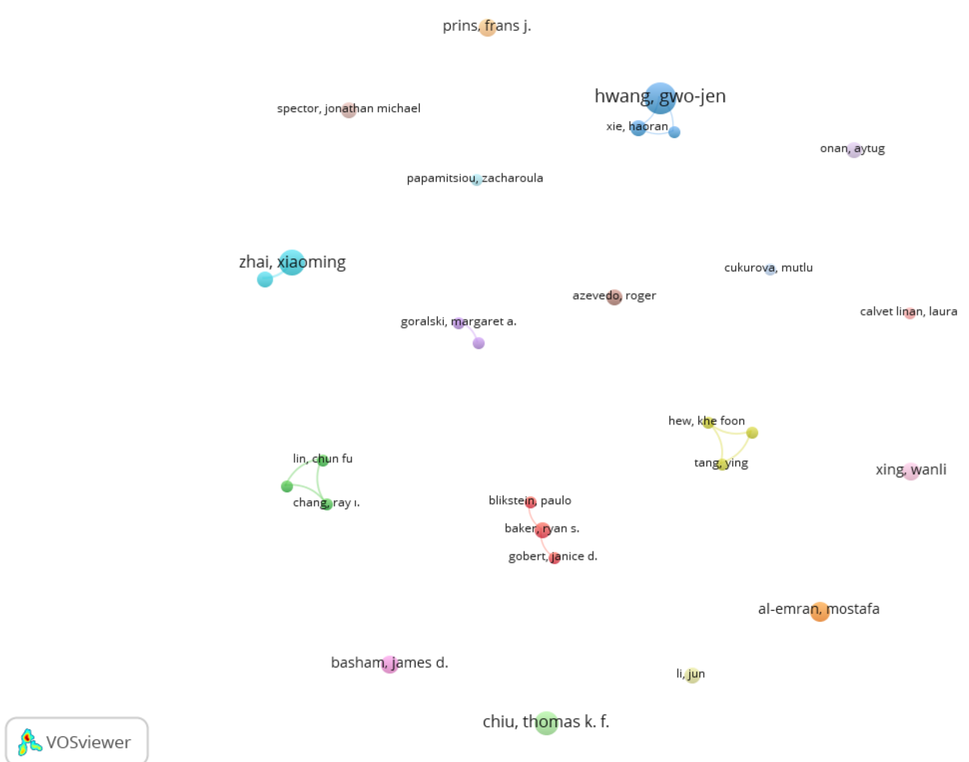
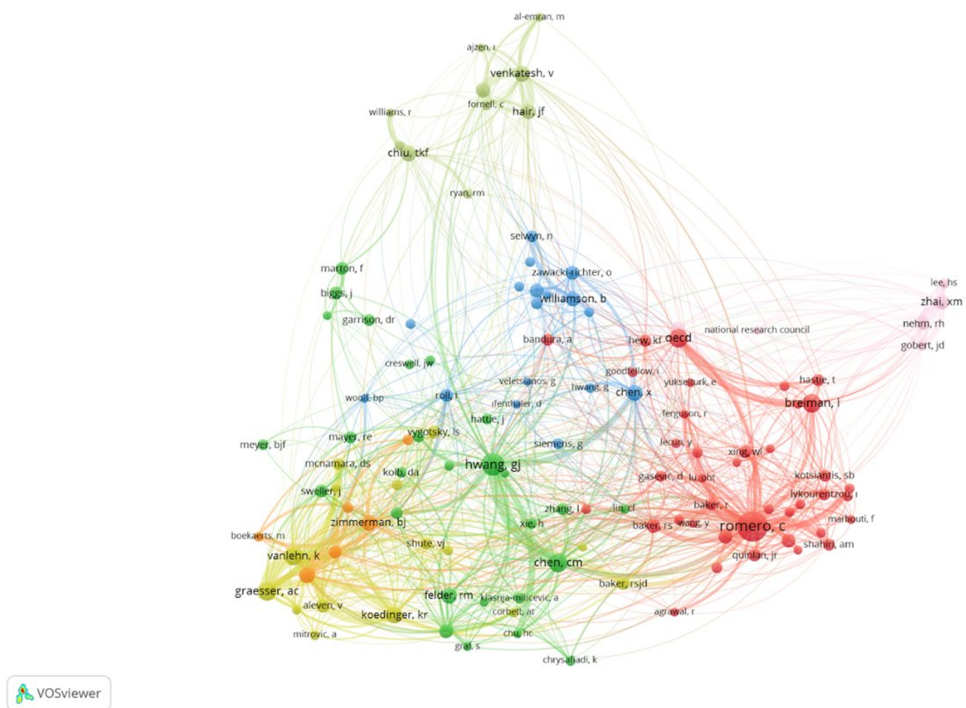


Figure 9 illustrates that 14 clusters were formed in the co-authorship network map of 29 linked authors. Baker, R. S., Blikstein, P., and Gobert, J. D. stood out in terms of centrality and inter-cluster linking. The connections between Lin, C. F. and Chang, R. I. suggested a partnership, likely indicating they had co-authored works. Similarly, Hwang, G.-J. appeared to be another significant contributor, with ties to Xie, H. and Chu, H.-C. which could indicate a shared research interest or a history of collaboration. The overall structure of the network, with its various clusters and connections, indicated a dynamic community of scholars who often work together, sharing ideas and contributing to the collective knowledge of their discipline.

Transitioning from the detailed co-authorship network, the analysis now turned to co-citation patterns to further explore the impact and interrelations of scholarly work within this academic community. This approach not only highlighted how authors were interlinked through shared references, but also shed light on the influential works and ideas that have shaped the discourse and development within the academic community. Figure 10 presents the co-citation analysis of authors, illustrating the patterns of how their works were cross-referenced and interconnected within the scholarly network.

Figure 10

Most-Cited Authors (Co-Citation Analysis)



When the common citation network was examined, seven different colored clusters were seen. Authors who received many citations together were gathered in the same cluster. Publications in the center showed that they were often cited from different fields and had more detailed connections with many clusters. When Figure 10 was examined in its entirety, authors such as Romero, C., Hwang G.-J., Graesser, A. C., Chiu, T. K. F., Chen, X., and Zhai, X. M. were represented by larger clusters, which suggested that these authors were central to their respective clusters. This prominence implied that their work was highly regarded and frequently referenced together with other researchers in their area. Each cluster may have represented a different subfield or a specific area of research focus. For instance, researchers like Chen, C. M. and Hwang, G.-J. appeared to be in the same cluster, which could indicate that they worked on similar topics or within the same discipline.

When evaluated in terms of total link strength, high values for an author suggested widespread recognition and influence in the academic community, indicating their work's diversity across various topics or disciplines and their central role in research networks. It was observed that the authors with the highest TLS values were Romero, C., Hwang, G.-J., Vanlehn, K., and Graesser, A. C. The quantity of lines originating from an author and the thickness of these lines in a bibliometric network map signified the extent and frequency of citations, with thicker lines indicating stronger co-citation connections, all contributing directly to the author's TLS value.

Keyword Analysis

Co-word analysis was essential in this study for mapping the intellectual structure and thematic interrelations within AIEd research, revealing how various concepts within this field were interconnected and how they have evolved over time. It also provided strategic insights into prominent research trends and potential future directions by analyzing the co-occurrence of keywords in the literature. Figure 11 illustrates the network created by incorporating author keywords from WoS dataset that have appeared at least twice in distinct publications. Larger circles represented subjects that were more commonly discussed, and those in yellow showed the most popular subjects.

Figure 11

Analysis of Keyword Co-Occurrence Between 2013 and 2023

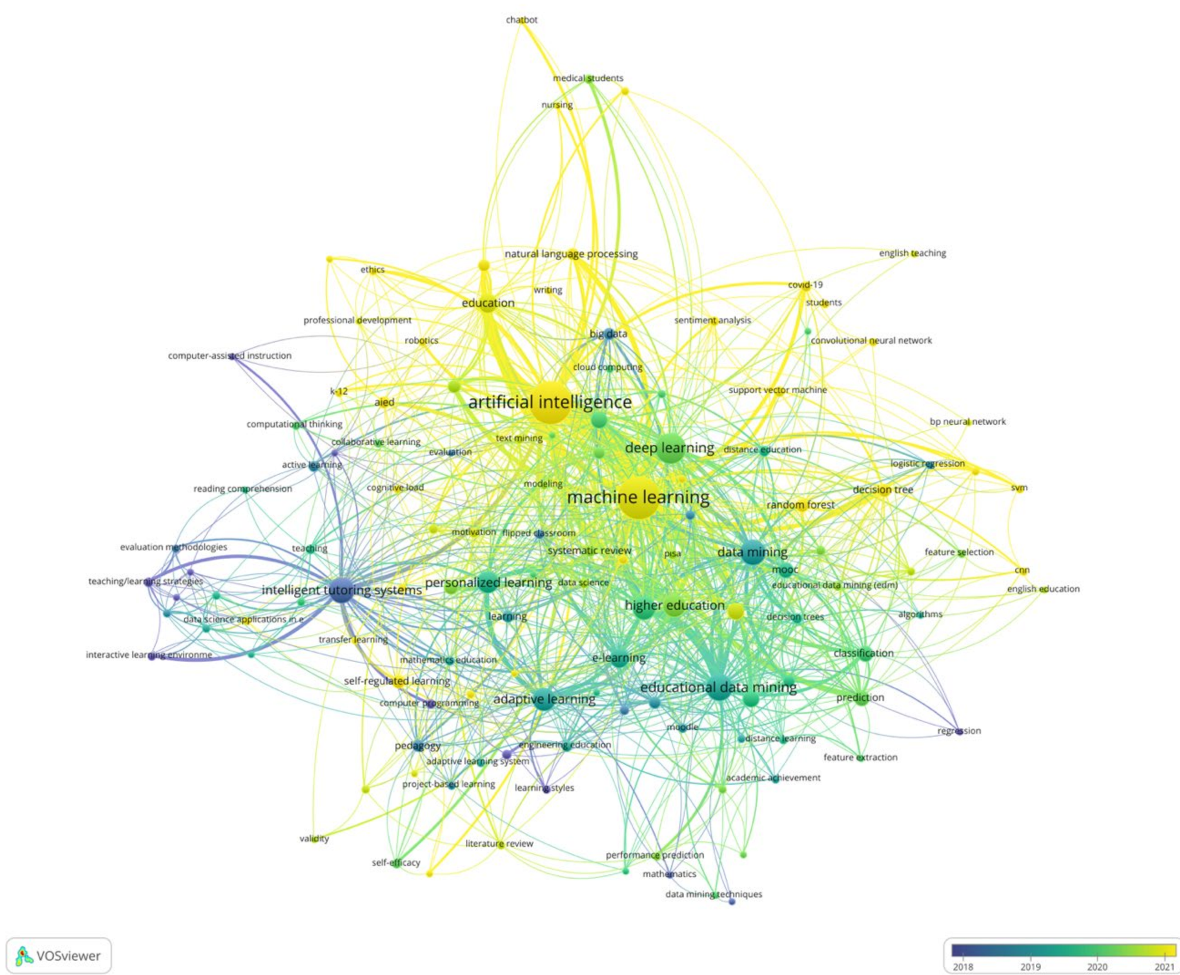


Figure 11 revealed terms such as artificial intelligence, machine learning, deep learning, data mining, and educational data mining situated at the core of the map. These terms stood out as the keywords used most, indicative of concepts frequently researched in conjunction with other thematic clusters. The figure also showed that terms such as intelligent tutoring systems, personalized learning, and adaptive learning were frequently used. When evaluated in terms of popularity (shown in yellow), it was observed that concepts such as artificial intelligence, machine learning, artificial neural network, and decision tree were prominent.

Conclusion and Discussion

The bibliometric analysis of AIEd literature over the past decade revealed the field's evolution, highlighting key countries, institutions, journals, authors, and trends. The study showed a stable publication rate from 2013 to 2017, followed by a sharp increase after 2017, reflecting growing interest in AIEd. This surge was likely due to advancements in AI and its potential in education, as well as a global need for AI educational solutions. Grassini (2023) suggested that AI's role in shaping future educational paradigms and its growing interest among educators will continue to rise.

The analysis showed the U.S., China, and India as leaders in AIEd research, consistent with previous studies (Baek & Doleck, 2020; Chen et al., 2020; Hinojo-Lucena et al., 2019; Jia et al., 2023; Liang et al., 2023; Mohamed et al., 2022; Moreno-Guerrero et al., 2020; Song & Wang, 2020; Talan, 2021; Tang et al., 2023; Zawacki-Richter et al., 2019). Contrary to Baek and Doleck (2020), this research found significant international cooperation, especially in bibliometric coupling and co-authorship networks, with the U.S. and China being particularly collaborative. Both China and Australia, China and Taiwan, U.S. and South Korea also showed high levels of international collaboration. The U.S. dominance in AIEd publications and collaborations has been attributed to its high research and development budgets, prestigious universities, innovation culture, and diverse academic community. These factors, also noted by Hebebci (2021) and Talan (2021), have contributed to the country's pioneering role in AIEd.

The National Taiwan University of Science and Technology led in AIEd publications and citations, with other institutions excelling in either publications or citations. Both metrics are crucial for assessing scientific impact. In bibliometric coupling, US universities like the University of Georgia, Michigan State University, and the University of Illinois, along with National Taiwan University of Science and Technology and Korea University, were notable, aligning with Talan's (2021) findings. Bibliometric coupling measures research integration and collaboration, indicating institutional impact and relationships in specific fields. The prominence of three US universities underscored the U.S. leadership in AIEd across publications, collaborations, and citations, highlighting its global influence and scientific leadership.

In assessing AIEd journals, article count, citation numbers, and total link strength scores were analyzed for academic impact. High article count suggests a journal's activity and content diversity, while high citations and link strength indicate influence and authority. These metrics, important for evaluating a journal's scientific contribution and prestige, should be considered together for a comprehensive understanding of a journal's impact. Additionally, a citation network analysis highlighted *Computers & Education, Education and Information Technologies*, and *Educational Technology & Society* as prominent, with strong citation connections and network popularity. *Computers & Education*, despite fewer publications, had high citation numbers, while *Education and Information Technology* scored highly across all metrics, indicating their significance in AIEd. These findings aligned with studies like Hwang and Tu (2021) and Liang et al. (2023), who attributed the results to the journals' long-standing publication, high impact factors, prestigious academic standing, and attraction of leading AIEd researchers. Their role in disseminating new ideas and accelerating scientific knowledge, reaching wide audiences, and promoting interdisciplinary studies has also contributed to their prominence.

The analysis of AIEd authors focused on their publication count and citation numbers, identifying influential researchers like Hwang G.-J. and Zhai, X. in terms of the number of publications, with Hwang leading in citations, followed by Papamitsiou, Z. Despite fewer publications, some authors' work received high citations, indicating the field's popularity and the impact of these publications. Our study did not compare these findings with the literature due to the dynamic nature of publication and citation data. Additionally, co-authorship and co-citation analyses were conducted. Co-authorship analysis, requiring at least two publications and 100 citations, revealed limited collaborations, suggesting either a lack of collaboration in AIEd or high criteria for analysis. Co-citation analysis helped us understand how researchers' ideas and trends interact and spread within the field.

In bibliometric analysis, keywords are considered the basic elements of representing knowledge concepts. They have been frequently used to uncover the knowledge structure of research domains (Su & Lee, 2010). As expected, the terms artificial intelligence, machine learning, deep learning, data mining, and educational data mining have been used quite extensively, with other keywords typically clustered around them. Similarly, when assessed in terms of popularity, it has been concluded that in recent years, these terms have been the most frequently used. The keyword analysis results of the bibliometric analysis studies applied directly in AI in education or in specific fields of AI were generally on basic topics such as artificial intelligence, deep learning, and machine learning (Baek & Doleck, 2020; Chen et al., 2020; Hwang & Tu, 2021; Kaban, 2023; Liang et al., 2023; Pua et al., 2021) as well as concepts such as mathematics education (Hwang & Tu, 2021), and engineering education (Pua et al., 2021), depending on the specific field of the study. Partial differences in these results can be explained by the period in which the bibliometric analysis was performed and the fact that it was conducted in a specific field.

The following are some recommendations for future research directions in the field of AI applications in education, based on the findings and scope of the current study:

- This study was conducted on the WoS database platform, considered to be one with the most influential publications in the literature. Again, the scope of the related concept can be increased by searching the Scopus database, one of the largest databases in the world, and other field indexes.

- Considering that the concept of AIEd is very popular and of increasing importance, studies comparing some of our findings of the study on a country-by-country basis can be conducted.

- Thematic analysis of the most cited studies in the related field may be important to express the importance of the related studies.

- Detailed analysis of AIEd studies at more specific educational levels (e.g., higher education or high school level) can help reveal and articulate the specific needs and trends in this field.

Acknowledgements

This study was supported by Balıkesir University Scientific Research Projects Coordination Unit (BAUN, BAP, Project Number: 2022/046).

References

Argente, J., Martínez-Rico, G., González-García, R. J., & Cañadas, M. (2023). Bibliometric analysis on the implementation of evidence-based practices through building effective systems. *Children, 10*(5), 813. https://doi.org/10.3390/children10050813

Baek, C., & Doleck, T. (2020). A bibliometric analysis of the papers published in the Journal of Artificial Intelligence in Education from 2015–2019. *International Journal of Learning Analytics and Artificial Intelligence for Education*, *2*(1), 67. https://doi.org/10.3991/ijai.v2i1.14481

Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2021). *An introduction to ethics in robotics and AI*. Springer International Publishing. https://doi.org/10.1007/978-3-030-51110-4

Bozkurt, A. (2023). Generative artificial intelligence (AI) powered conversational educational agents: The inevitable paradigm shift. *Asian Journal of Distance Education, 18*(1). Retrieved from https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/718

Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends, 66(*4), 616–630. https://doi.org/10.1007/s11528-022-00715-y

Chaudhry, M. A., & Kazim, E. (2022). Artificial intelligence in education (AIEd): A high-level academic and industry note 2021. *AI Ethics 2*, 157–165. https://doi.org/10.1007/s43681-021-00074-z

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Chiu, W.-K. (2021). Pedagogy of emerging technologies in chemical education during the era of digitalization and artificial intelligence: A systematic review. *Education Sciences, 11*(11), 709. https://doi.org/10.3390/educsci11110709

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133*, 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Durso, S. D. O., & Arruda, E. P. (2022). Artificial intelligence in distance education: A systematic literature review of Brazilian studies. *Problems of Education in the 21st Century, 80*(5), 679–692. https://doi.org/10.33225/pec/22.80.679

Ellegaard, O., & Wallin, J. A. (2015). The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics, 105*, 1809–1831. https://doi.org/10.1007/s11192-015-1645-z

García-Martínez, I., Fernández-Batanero, J. M., Fernández-Cerero, J., & León, S. P. (2023). Analysing the impact of artificial intelligence and computational sciences on student performance: Systematic review and meta-analysis. *Journal of New Approaches in Educational Research, 12*(1), 171. https://doi.org/10.7821/naer.2023.1.1240

Gaviria-Marin, M., Merigo, J. M., & Popa, S. (2018). Twenty years of the Journal of Knowledge Management: A bibliometric analysis. *Journal of Knowledge Management, 22*(8), 1655–1687. https://doi.org/10.1108/JKM-10-2017-0497

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences, 13*(7), 692. https://doi.org/10.3390/educsci13070692

Hebebci, M. T. (2021). The bibliometric analysis of studies on distance education. *International Journal of Technology in Education, 4*(4), 796–817. https://doi.org/10.46328/ijte.199

Hinojo-Lucena, F.-J., Aznar-Díaz, I., Cáceres-Reche, M.-P., & Romero-Rodríguez, J.-M. (2019). Artificial intelligence in higher education: A bibliometric study on its impact in the scientific literature. *Education Sciences*, *9*(1), Article 1. https://doi.org/10.3390/educsci9010051

Ho, Y. S. (2008). Bibliometric analysis of biosorption technology in water treatment research from 1991 to 2004. *International Journal of Environment and Pollution, 34*(1–4), 1–13. http://dx.doi.org/10.1504/IJEP.2008.020778

Hwang, G.-J., & Tu, Y.-F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics*, *9*(6), 584. https://doi.org/10.3390/math9060584

Jia, F., Sun, D., & Looi, C. (2023). Artificial intelligence in science education (2013–2023): Research trends in ten years. *Journal of Science Education and Technology, 33*. https://doi.org/10.1007/s10956-023-10077-6

Joiner, I. A. (2018). Artificial intelligence: AI is nearby. In I. A. Joiner (Ed), *Emerging library technologies* (pp. 1–22). Chandos Publishing. https://doi.org/10.1016/B978-0-08-102253-5.00002-2

Kaban, A. (2023). Artificial intelligence in education: A science mapping approach. *International Journal of Education in Mathematics, Science and Technology, 11*(4), 844–861. https://doi.org/10.46328/ijemst.3368

Khanam, S., Tanweer, S., & Khalid, S. (2021). Artificial intelligence surpassing human intelligence: Factual or hoax. *The Computer Journal, 64*(12), 1832–1839. https://doi.org/10.1093/comjnl/bxz156

Liang, J.-C., Hwang, G.-J., Chen, M.-R. A., & Darmawansah, D. (2023). Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments, 31*(7), 4270–4296. https://doi.org/10.1080/10494820.2021.1958348

Martins, T., Braga, A., Ferreira, M. R., & Braga, V. (2022). Diving into social innovation: A bibliometric analysis. *Administrative Sciences, 12*(2), 56. https://doi.org/10.3390/admsci12020056

Mohamed, M. Z. B., Hidayat, R., Suhaizi, N. N. B., Sabri, N. B. M., Mahmud, M. K. H. B., & Baharuddin, S. N. B. (2022). Artificial intelligence in mathematics education: A systematic literature review. *International Electronic Journal of Mathematics Education, 17*(3), em0694. https://doi.org/10.29333/iejme/12132.

Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., & Cobo, M. J. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. *Profesional de la Información, 29*(1). https://doi.org/10.3145/epi.2020.ene.03

Moreno-Guerrero, A.-J., López-Belmonte, J., Marín-Marín, J.-A., & Soler-Costa, R. (2020). Scientific development of educational artificial intelligence in Web of Science. *Future Internet*, *12*(8), Article 8. https://doi.org/10.3390/fi12080124

Pelletier, K., Robert, J., Muscanell, N., McCormack, M., Reeves, J., Arbino, N., Grajek, S., Birdwell, T., Liu, D., Mandernach, J., Moore, A., Porcaro, A., Rutledge, R., & Zimmern, J. (2023). *EDUCAUSE horizon report, teaching and learning edition*. EDUCAUSE.

Prahani, B. K., Rizki, I. A., Jatmiko, B., Suprapto, N., & Tan, A. (2022). Artificial intelligence in education research during the last ten years: A review and bibliometric study. *International Journal of Emerging Technologies in Learning, 17*(8), 169–188. https://doi.org/10.3991/ijet.v17i08.29833

Pua, S., Ahmad, N. A., Khambari, M. N. Md., & Yap, N. K. (2021). Identification and analysis of core topics in educational artificial intelligence research: A bibliometric analysis. *Cypriot Journal of Educational Sciences*, *16*(3), 995–1009. https://doi.org/10.18844/cjes.v16i3.5782.

Salas-Pilco, S. Z., Xiao, K., & Hu, X. (2022). Artificial intelligence and learning analytics in teacher education: A systematic review. *Education Sciences, 12*(8), Article 8. https://doi.org/10.3390/educsci12080569

Salas-Pilco, S.Z., Yang, Y. (2022). Artificial intelligence applications in Latin American higher education: a systematic review. *International Journal of Educational Technology in Higher Education, 19*, 21. https://doi.org/10.1186/s41239-022-00326-w

Song, P., & Wang, X. (2020). A bibliometric analysis of worldwide educational artificial intelligence research development in recent twenty years. *Asia Pacific Education Review, 21*(3), 473–486. https://doi.org/10.1007/s12564-020-09640-2

Sapci, A. H., & Sapci, H. A. (2020). Artificial intelligence education and tools for medical and health informatics students: Systematic review. *JMIR Medical Education*, *6*(1), e19285. https://doi.org/10.2196/19285

Su, H. N., & Lee, P. C. (2010). Mapping knowledge structure by keyword co-occurrence: A first look at journal papers in technology foresight. *Scientometrics, 85*(1), 65–79. https://doi.org/10.1007/s11192-010-0259-8

Talan, T. (2021). Artificial intelligence in education: A bibliometric study. *International Journal of Research in Education and Science, 7*(3), 822–837. https://doi.org/10.46328/ijres.2409

Tang, K.-Y., Chang, C.-Y., & Hwang, G.-J. (2023). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments, 31*(4), 2134–2152. https://doi.org/10.1080/10494820.2021.1875001

Van Eck, N. J., & Waltman, L. (2014). Visualizing bibliometric networks. In *Measuring scholarly impact: Methods and practice* (pp. 285-320). Cham: Springer International Publishing.

Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: a systematic review from 2011 to 2021. *International Journal of STEM Education, 9*(1), 59. https://doi.org/10.1186/s40594-022-00377-5

Yin, M. S. (2013). Fifteen years of grey system theory research: A historical review and bibliometric analysis. *Expert Systems with Applications, 40*(7), 2767–2775. https://doi.org/10.1016/j.eswa.2012.11.002

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

Keep the Ball Rolling in AI-Assisted Language Teaching: Illuminating the Links Between Productive Immunity, Work Passion, Job Satisfaction, Occupational Success, and Psychological Well-Being Among EFL Teachers

Tahereh Heydarnejad[1,*] and Fidel Çakmak[2]
*[1]Department of English Language, Faculty of Literature and Humanities, University of Gonabad, Gonabad, Iran; [2]Department of Foreign Language Education, Alanya Alaaddin Keykubat University, Turkey; *Corresponding author*

Abstract

Artificial intelligence (AI) revolutionizes education by fundamentally altering the methods of teaching and processes of learning. Given such circumstances, it is essential to take into account the mental and psychological well-being of teachers as the architects of education. This research investigated the links between teacher immunity (TI), work passion (WP), job satisfaction (JS), occupational well-being (OW-B) and psychological well-being (PW-B) in the context of AI-assisted language learning. In order to achieve this objective, 392 Iranian teachers of English as a foreign language (EFL) were given the Language Teacher Immunity Instrument, the Work Passion Scale, the Job Satisfaction Questionnaire, the Occupational Well-Being Scale, and the Psychological Well-Being at Work Scale. By using confirmatory factor analysis and structural equation modeling, the study identified and quantified the impacts of TI, WP, JS, OW-B, and PW-B via data screening. The findings emphasize the crucial role that TI and WP play in providing a balance in teachers' JS, OW-B, and PW-B while applying AI in their language instruction. The broad ramifications of this research are explored.

Keywords: AI-assisted language teaching, teacher productive immunity, work passion, job satisfaction, occupational success, psychological well-being, EFL teacher

Introduction

In a broad sense, the success of any nation is contingent upon the educational system administered in that country. According to Simmons et al. (2019), education has the potential to be beneficial and efficient provided instructors fulfill the critical role they are expected to play in ensuring that students attain the educational goals imposed by the education system. As stated by Wessels and Wood (2019), the teaching profession is considered to be among the most successful professions in a given community. Due to the fact they are the foundation of any educational system, teachers are regarded as the architects of a country since they are responsible for the care and teaching of future generations. It is a well-known truth that teachers are considered to be the backbone of a country that is both healthy and happy. This is due to the fact that they are the only instructors capable of devotedly managing the challenging process of nation-building (Ryff & Keyes, 1995; Saaranen et al., 2013). In order to accomplish this important goal, it is imperative that educators fulfill their professional responsibilities effectively, especially in AI-assisted language learning. If the objective is to establish a setting where teachers are able to perform their jobs effectively, then some aspects, such as the level of work satisfaction, should get sufficient attention from those in charge of education.

The swiftly expanding discipline of computer science identified as AI is developing intelligent computers capable of simulating human intelligence and performing tasks that are typically executed by humans. Transportation, finance, and healthcare are among the industries that are increasingly adopting AI technology. AI's capacity to make decisions with remarkable speed and precision has the potential to radically alter many different types of enterprises. Using machine learning techniques, AI learns new things and becomes better at what it does. With the help of algorithms, robots can analyze massive databases, spot trends, and find insights that humans cannot fathom (Licardo et al., 2024). Despite AI being extensively used in several industries, its full potential in the context of EFL instructors' psychological well-being remains unfulfilled. The purpose of the current study was to assess the connections between TI, WP, JS, OW-P, and PW-B in the context of EFL education.

Literature Review

The concept of AI was first proposed by McCarthy, who defined it as a scientific and technical concept parallel to the development of intelligent machines (dos Santos & Rosinhas, 2023). AI is a fast-progressing discipline in computer science that focuses on creating robots and software capable of doing activities that typically need human intellect (Terra et al., 2023). These activities manifest the cognitive processes of humans, which include learning, thinking, problem-solving, recognizing patterns, and making predictions (Siemens et al., 2023). AI implementations may emerge in several modalities, using either physical or virtual components, and can function within self-governing or decentralized frameworks. Furthermore, the implementations have the ability to materialize as astute, autonomous entities with the capacity to engage with their surroundings and exercise judgment (Luxton, 2016). Artificial intelligence may be classified into two categories: narrow or weak AI, which focuses on specialized activities, and general or strong AI, which has the capability to do intellectual tasks at a level equivalent to humans (Kay, 2012). Recent inquiries indicate that the use of AI in instruction has a beneficial effect on language learning (Kohnke et al., 2023), providing interactive learning affordances (Chiu et al., 2023). Applying AI in language learning might have both positive and negative effects on teachers, who are at the core of their students' education and who guide their learning step by step.

Immunity is a biochemical defense mechanism that activates the body's naturally inherent defenses and refuses infections, according to Hiver (2015). Its purpose is to shield the inside from outside forces that might cause harm or distress (Hiver, 2017). Teacher immunity, as described by Hiver and Dörnyei (2017), is an approach that effectively addresses many conflicts and difficulties encountered in the field of education. As stated by Haseli Songhori et al. (2018), one end of the teacher immunity spectrum represents teachers' levels of passion for teaching, mental wellness, and openness to change, while the other end represents teachers' levels of educational expectations, weariness, and dropout.

An offshoot of complexity theory, self-organization theory lies at the heart of the teacher immunity establishment (de Boer, 2005). The process of self-organization involves the transformation of a dynamic system's overall functioning through the interaction of its components. This transformation occurs in four distinct phases: activation, integration, adjustment, and equilibrium. (Randi, 2004). When confronted with challenges, language teachers may exhibit their immunity in two primary ways: productive or maladaptive responses (Hiver, 2015; Hiver & Dörnyei, 2017). Productive immunity includes emotions such as optimism, devotion, enthusiasm, resilience, and motivation. Apathy, conservatism, cynicism, and resistance to change may be attributed to maladaptive immunity. Additionally, maladaptive immunity is characterized by a biological counterpart that functions in a similar manner. This distinction emphasizes how adaptive and maladaptive immunity influence individual behavior and broader system dynamics.

Another teacher-associated concept is WP. It is an incentivizing procedure that enables people to efficiently tackle diverse activities. This enthusiasm is evident in workers' willingness to undertake important tasks that demand their energy, ultimately incorporating these behaviours as fundamental to their identity (Vallerand et al., 2003). Vallerand et al. (2003) proposed a dichotomous paradigm for passion, distinguishing two distinct types: harmonious and obsessive passion. Harmonious passion emerges when someone willingly engages in an activity and integrates it into their sense of self. It refers to purposefully engaging in meaningful and essential things, which helps create a sense of harmony with one's complete being. Obsessive passion is distinguished by the integration of control into an individual's psyche as they internalize the action. This fixation is driven by internal cumpulsion and/or external influences such as self-esteem or societal approval, or by overwhelming enthusiasm (Vallerand et al., 2003).

The impact of harmonious and obsessive passion on people results in diverse interactions between passion and work needs. The latter refers to occupations that need exertion and are linked to particular expenses (Vallerand et al., 2007). Consequently, these activities have the capacity to exert control over workers, leading to feelings of discomfort, unease, and fatigue (Vallerand et al., 2003). Overwhelming expectations may fuel workers' motivations, leading to an obsessive zeal that compels them to approach their job responsibilities in inflexible and insufficient ways. This ultimately results in reduced levels of wellness for staff members (Cabrita & Duarte, 2023).

Therefore, it is reasonable to assume that WP mediates the relationship between job demands and emotional health in the workplace. Enthusiasm for one's work is highly related to intrinsic motivation.

According to the self-determination theory, both internal and extrinsic motivations drive and motivate human action (Ryan & Deci, 2017). A driving force that complements motivation, passion enhances motivation, promotes wellness, and infuses daily activities with significance (Cabrita & Duarte, 2023). Due to the joy and fulfillment experienced when performing, people tend to favor some pursuits over others. Moreover, participating in activities that ignite our passions and shape our identities, therefore offering a consistent sense of satisfaction, can significantly influence an individual's psychological well-being (Vallerand et al., 2007). In a nutshell, the desire to reach a goal is dictated by teachers' level of passion, and the process of being motivated is what gets them there.

Within educational settings, the level of JS experienced by teachers may be seen as an indicator of their likelihood to stay in their profession, a factor that influences their level of dedication, and ultimately, a factor that contributes to the overall efficacy of the school (Shan, 1998). In this regard, Buitendach and de Witte (2005) contended that JS has a significant effect in influencing teachers' viewpoints and evaluations of their work. This perception, in return, may greatly impact their objectives and accomplishments inside the school system. JS refers to an individual's emotional reactions to certain characteristics, environment, and conditions related to their work (Werang et al., 2017). Regarding teachers, the term pertains to their affective reactions towards their occupation and professional circumstances (Zhang, 2021). JS may manifest either in a broad or a particular manner. The former refers to a general sense of contentment with one's work, while the latter is more specific and pertains to certain parts of the profession (Lopes & Oliveira, 2020). JS is determined by the extent to which one's wants and wishes are fulfilled in comparison to the actual practices in the workplace (Baluyos et al., 2019).

In order to evaluate their careers, educators look at what makes their jobs special. JS may be defined as a teacher's subjective perception and attitude towards their profession. Similar to other attitudes, it encompasses an intricate combination of awareness, sentiments, and behavioral inclinations (Werang et al., 2017). A teacher who experiences a high degree of JS has favorable views towards their workplace. Conversely, a teacher who is unsatisfied with their employment harbors negative attitudes towards the working atmosphere. Hence, this favorable or adverse attitude might influence the conduct of instructors in the school setting. Employment satisfaction pertains to an individual's comprehensive attitude towards their employment. JS is the emotional response, either favorable or bad, that arises from evaluating one's experience in a job. Professional factors such as subject knowledge, teaching effectiveness, competence, and academic credentials play a role in teachers' JS (Michaelowa, 2002).

OW-B refers to the state of well-being in the context of the workplace. Working on Ryff's (1989) and Warr's (1994) generic definition of well-being, van Horn et al. (2004) related it to a complementary and multi-dimensional phenomenon. Van Horn et al. (2004) focused on emotional state as a critical affective aspect of work well-being. They approached it to measure it through the emotional state of employees, such as JS, emotional exhaustion, and organization commitment. They proposed two further dimensions, psychosomatic well-being (e.g. pain or aches dues to work stress or long working hours) and cognitive well-being (e.g., attention and engagement), to the existing concept of OW-B. OW-B has received interest in education and positive psychology. It is commonly observed that the state of well-being is naturally reflected in teachers' classroom practices (Chan et al., 2023). Therefore, investigating occupational well-being recognizes the teachers' presence in their work lives, giving ideas about improving well-being initiatives at work for the practical outcomes of teaching. The PW-B of teachers is a critical factor that has a substantial effect on their performance and, therefore, affects the achievements of their students. It could be due to the fact that students are generally impacted by the caliber of the instructors. Furthermore, there is a considerable emphasis on teachers' well-being as a possible means to alleviate work stress and discontent. (Parker, 2012). Teacher well-being includes aspects such as managing stress, mental health, overall life satisfaction, and a sense of fulfillment. Well-being among students and teachers is associated with both a more favorable emotional state and improved academic achievement (Paterson & Grantham, 2016). Teacher PW-B refers to a broad spectrum of favorable emotions and states of being at the workplace, together with general contentment with life and one's professional path (McInerney et al., 2015). Moreover, there are several distinct perspectives on the concept of well-being. Several academics have identified self-acceptance, a sense of purpose, personal growth, supportive connections, empowerment, and appreciation for nature as essential factors for achieving a state of thriving (Mercer, 2021).

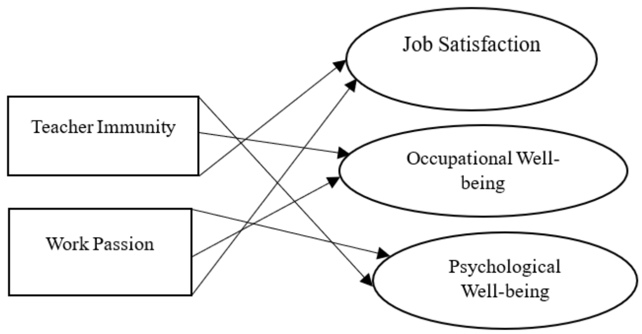
Objectives of the Current Research

Due to the insufficient amount of research that has been carried out in the field of education and the relevance of the components that have been identified in terms of improving language instruction, the aim of this research was to investigate the potential interplay between TI, WP, JS, OW-B, and PW-B of EFL teachers involved in AI-integrated language teaching. A conceptual framework was developed to demonstrate the interaction among TI, WP, JS, OW-B, and PW-B as illustrated in Figure 1. The conceptual framework was set by recent research and theories in the field. The assessment was conducted using confirmatory factor analysis (CFA) and structural equation modeling (SEM), and the findings were subsequently presented. The following research questions were formed:

1. Does the EFL instructors' TI and WP indicate their JS in AI-integrated instruction?
2. Does the EFL instructors' TI and WP indicate their OW-B in AI-integrated instruction?
3. Does the EFL instructors' TI and WP indicate their PW-B in AI-integrated instruction?

Figure 1

Research Model for Exploring Factors of EFL Teacher Well-Being in the AI-Assisted Classroom



Method

Context and Participants

All 392 survey participants were EFL instructors; 105 men and 287 women participated. They were teaching in Iranian private language institutions which are equipped with AI as part of their language teaching. The participants had been teachers from 1 to 25 years, and their ages ranged from 22 to 48. They completed training courses offered by the institutions they attended in order to incorporate AI into their lessons regarding the scalability of teaching resources and materials and provide suggested teaching strategies for specific subjects in the curriculum.

The data was collected via online forms, most especially Google Forms, in 2023. Scales were employed in the target language (English) to maintain the authenticity of the instruments. Data loss was very unlikely because of the meticulous planning that went into the computerized survey. The distribution of the data was initially analyzed using the Kolmogorov-Smirnov test. The data's normality was validated by data screening, proving that parametric procedures would be reliable. With the data assumed to follow a normal distribution, the software LISREL 8.80 (https://ssicentral.com/index.php/products/lisrel/) was used to perform CFA and SEM.

Instruments

The Language Teacher Immunity Instrument (LTII) developed by Hiver (2017) was applied to determine the level of immunity possessed by participants. The 39 questions that make up this instrument are organized into seven subscales, and each of these subscales has a six-point response scale (1 = *strongly disagree*; 6 = *strongly agree*). The subscales include seven items on teaching self-efficacy, five items each on burnout, resilience, and attitudes toward teaching, six items on openness to change and classroom affectivity, and five items on coping. Cronbach's alphas for these items were satisfactory, ranging from 0.72 to 0.83.

Work passion was assessed using the Work Passion Scale (WPS) created by Vallerand and Houlfort (2003). The scale is comprised of fourteen components. Seven questions are used to evaluate harmonious passion (α = 0.76), while the other seven are used to examine obsessive passion (α = 0.77). The anchors of the scale range from 1 (*strongly disagree*) to 7 (*strongly agree*). The WPS's reliability as determined by Cronbach's alpha was satisfactory for the present inquiry, with values ranging from 0.77 to 0.76.

Furthermore, the Job Satisfaction Questionnaire (JSQ) developed by Spector (1985) was used to assess the degree of instructors' job satisfaction. It includes 36 statements that pertain to different aspects of job satisfaction. The questionnaire uses a Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The research results approved the good reliability of the JSQ as evidenced by Cronbach's alpha values varying from 0.71 to 0.83.

A total of twelve questions are included in Warr's (1990) Occupational Well-Being Scale (OWS), which is used to assess the level of happiness experienced by educators. The purpose of this measure is to assess the well-being of educators across two dimensions: physical and emotional. Anxiety, contentment, depression, and enthusiasm were the job-related emotions that participants were asked to rate in the past few weeks. Participants were asked to indicate the degree to which they experienced these emotions. There were six answer alternatives, ranging from 1 (*never*) to 6 (*always*). The Cronbach's alpha estimated reliability of the OWS was satisfactory (α = 0.75).

This research used the Psychological Well-Being at Work (PWBW) scale developed by Dagenais-Desmarais and Savoie (2012) to assess the psychological well-being of educators. PWBW has 25 assertions, each rated on a 5-point scale ranging from 0 (disagree) to 5 (completely agree). PWBW comprises five distinct components: interpersonal fit at work, thriving at work, feeling of competence at work, perceived recognition at work, and desire for involvement at work. The Cronbach's alpha value for PWBW yielded an acceptable result (ranging from 0.95 to 0.71).

Results

This section provides a summary of the data analyzed. In Table 1, the descriptive data from the instruments and scales administered in this study are shown.

Table 1

Descriptive Statistics of Psychological and Professional (P&P) Measures in Teaching

P&P Measures	Minimum	Maximum	*M*	*SD*
Teaching self-efficacy	12	49	31.02	6.05
Burnout	5	35	22.69	4.44
Resilience	5	33	22.80	4.00
Attitudes toward teaching	5	34	22.95	4.60
Openness to change	6	39	27.12	4.88
Classroom affectivity	6	40	28.40	4.35
Coping	5	34	23.24	4.94
Teacher immunity	120	233	178.25	20.11
Harmonious passion	7	46	31.39	6.45
Obsessive passion	18	46	31.93	5.74
Work passion	25	89	63.33	10.20
Supervision	3	15	10.33	2.90
Colleagues and communication	4	20	13.20	3.93
Working conditions and operating procedures	3	15	10.29	2.94
Pay and benefits	4	20	13.73	3.70
Rewards and profits	6	20	13.71	2.95
Work itself	7	20	15.31	2.92
Advancement/Promotion	6	20	14.81	3.34
Operating procedures	6	20	14.73	3.00
Communication	6	15	10.76	2.57
Contingent rewards	12	30	21.67	5.09
Job satisfaction	80	167	127.80	17.47
Occupational well-being	17	72	47.28	7.16
Interpersonal fit at work	10	25	19.71	3.73
Thriving at work	6	25	15.94	4.28
Feeling of competence at work	10	25	17.90	3.98
Perceived recognition at work	5	25	17.82	5.41
Desire for involvement at work	5	25	18.47	4.98
Teacher psychological well-being	43	120	89.84	19.72

Note. N=392

Table 1 shows that the among LTII subcomponents, the teaching self-efficacy score was 31.02 (*SD* = 6.05). On the WPS, the second instrument used, obsessive passion was shown to be more significant (*M* = 31.93, *SD* = 5.74). Contingent rewards had the highest mean score in the JSQ (*M* = 21.67, *SD* = 5.09). The mean score relevant to the OWS was occupational well-being (*M* = 47.28, *SD* = 7.16). Furthermore, interpersonal fit at work had the highest mean score in the PWBW (*M* = 19.71, *SD* = 3.73).

Subsequently, the Kolmogorov-Smirnov test was conducted to find any patterns. As shown in Table 2, the p values of all the instruments and their components exceeded 0.05, indicating that the findings normally distributed, which provided justification for using parametric approaches over the data analysis stage.

Table 2

Results of Kolmogorov-Smirnov Test on the Distribution of Factors Related to EFL Teacher Well-Being

Factors related to EFL Teacher Well-being	Kolmogorov-Smirnov Z	Asymp. Sig. (2-tailed)
Teaching self-efficacy	0.823	0.508
Burnout	1.036	0.233
Resilience	0.943	0.336
Attitudes toward teaching	0.801	0.543
Openness to change	0.799	0.546
Classroom affectivity	0.997	0.273
Coping	1.345	0.054
Teacher immunity	0.446	0.989
Harmonious passion	1.083	0.192
Obsessive passion	0.868	0.439
Work passion	0.892	0.404
Supervision	0.936	0.346
Colleagues and communication	1.088	0.187
Working conditions and operating procedures	0.978	0.294
Pay and benefits	0.567	0.905
Rewards and profits	1.092	0.184
Work itself	0.911	0.377
Advancement/Promotion	0.747	0.633
Operating procedures	1.595	0.052
Communication	1.016	0.253
Contingent rewards	0.904	0.388
Job satisfaction	0.688	0.731
Occupational well-being	0.696	0.719
Interpersonal fit at work	1.137	0.150
Thriving at work	1.134	0.153
Feeling of competence at work	0.997	0.274
Perceived recognition at work	1.054	0.216
Desire for involvement at work	1.295	0.070
Teacher psychological well-being	1.004	0.266

Since all instruments and their subscales had statistically significant values, greater than 0.05, parametric methods were deemed appropriate to evaluate the data since it followed a normal distribution. This study employed a Pearson product-moment correlation to investigate the relation between TI, WP, JS, OW-B, and PW-B. Results are displayed in Table 3.

Table 3

Correlations for TI, WP, JS, OW-B, and PW-B

Factors in teacher well-being	1	2	3	4	5	6	7	8	9	10	11	12
1. Teaching self-efficacy	-*											
2. Burnout	0.678**	-										
3. Resilience	0.589**	0.678	-									
4. Attitudes toward teaching	0.612**	0.627**	0.489**	-								
5. Openness to change	0.633**	0.682**	0.514**	0.477**	-							
6. Classroom affectivity	0.505**	0.589**	0.631**	0.615**	0.578**	-						
7. Coping	0.604**	0.623**	0.477**	0.646**	0.458**	0.615**	-					
8. Harmonious passion	0.689**	0.712**	0.703**	0.705**	0.613**	0.543**	0.469**	-				
9. Obsessive passion	0.703**	0.577**	0.664**	0.531**	0.664**	0.618**	0.441**	0.485**	-			
10. Job satisfaction	0.742**	0.641**	0.751**	0.722**	0.703**	0.677**	0.651**	0.504**	0.488**	-		
11. Occupational well-being	0.579**	0.525**	0.598**	0.641**	0.584**	0.621**	0.543**	0.423**	0.431**	0.612**	-	
12. Psychological well-being	0.887**	0.776**	0.846**	0.894**	0.832**	0.824**	0.801**	0.472**	0.458**	0.631**	0.572**	-

Note. TI = teacher immunity; WP = work passion; JS = job satisfaction; OW-B = occupational well-being; PW-B = psychological well-being.
* Dash is used to report that data was not available.
**Correlation is significant at the 0.01 level (2-tailed).

As displayed in Table 3, significant associations were found across various subcomponents such as job satisfaction, occupational well-being, and psychological well-being, with particularly strong associations noted in areas like teaching self-efficacy and attitudes toward teaching. After this computation, the statistical program LISREL 8.80 was used in combination with CFA and SEM to examine the structural relationships between TI, WP, JS, OW-B, and PW-B. The fit model was assessed using indicators: the chi-square magnitude, the root-mean-square error of approximation (RMSEA), the normed fit index (NFI), the good fit index (GFI), and the comparative fit index (CFI). These metrics evaluate how well the model and data match. The relationship for this study is displayed in Table 4. Table 4 presents the fit criteria for two models assessed in the study.

Table 4

Comparison of Fit Indices in Models Exploring Factors in EFL Teachers' Well-Being

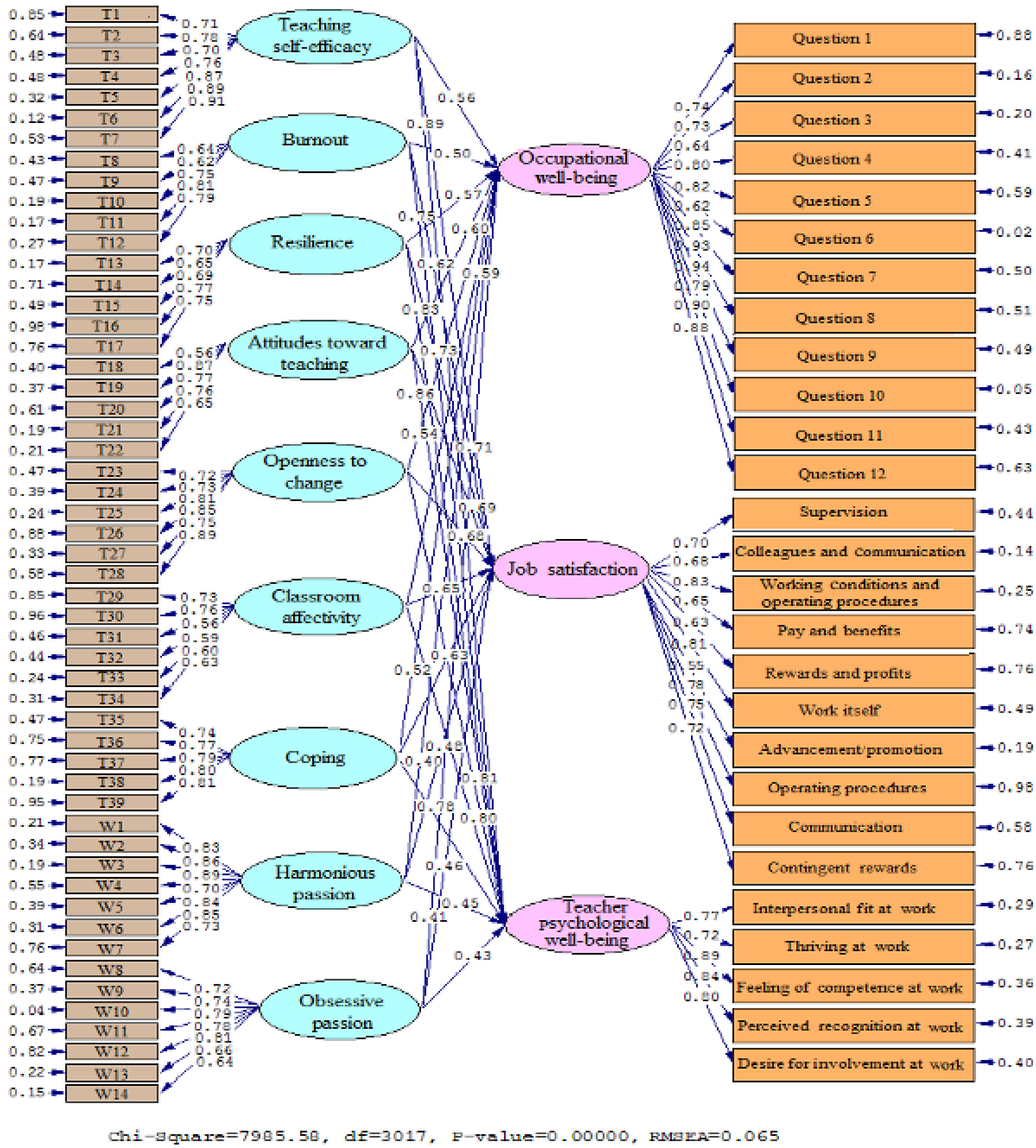
Fit index	$\chi 2$	df	$\chi 2/df$	RMSEA	GFI	NFI	CFI
Cut value			< 3	< 0.1	> 0.9	> 0.9	> 0.9
Model 1	1669.81	587	2.845	0.069	0.931	0.920	0.942
Model 2	7985.58	3017	2.647	0.065	0.931	0.935	0.956

Note. RMSEA = root-mean-square error of approximation; GFI = goodness-of-fit index; NFI = normed fit index; CFI = comparative fit index.

The model 1 fit criteria are met by the chi-square/degrees of freedom ratio of 2.845, the RMSEA of 0.069, the GFI of 0.931, the NFI of 0.920, and the CFI of 0.942. Additionally, this table shows that every fit index associated with model 2 is suitable. The chi-square/degrees of freedom ratio of 2.647, the RMSEA of 0.065, the GFI of 0.931, the NFI of 0.935, and the CFI of 0.956 show that the fit criteria have been met.

Figure 2

Path Coefficients for the Interaction Between TI, WP, JS, OW-B, and PW-B Subcomponents (Model 1)



Note. TI = teacher immunity; WP = work passion; JS = job satisfaction; OW-B = occupational well-being; PW-B = psychological well-being.

Figure 3

t-*Values Indicating the Relevance of Route Coefficients in Model 1*

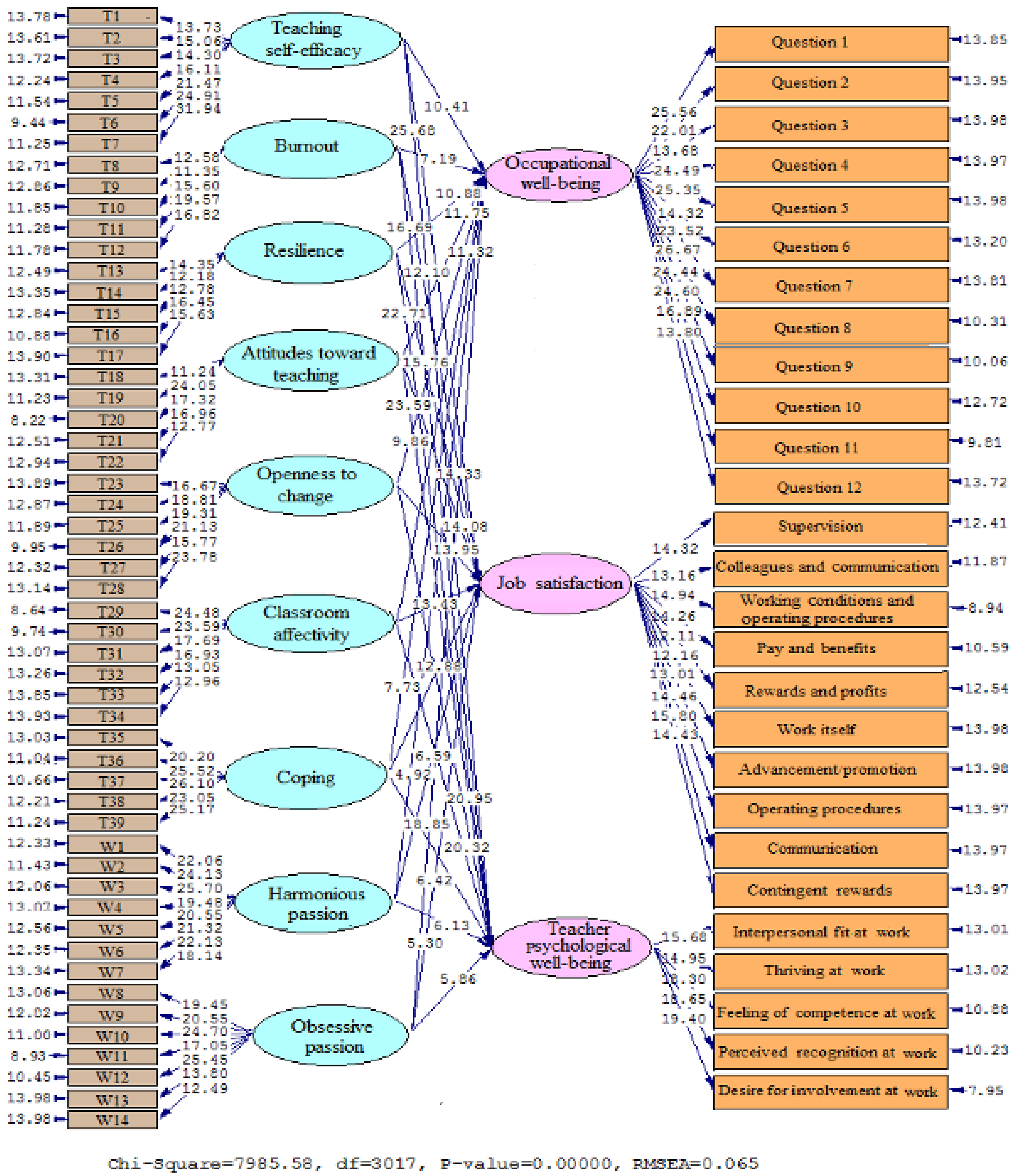


Table 5

Review of Model 1's Outcomes

Path			Path coefficient	*t* statistic	Result
Teacher immunity	→	Job satisfaction	0.67	14.25	Supported
Teacher immunity	→	Occupational well-being	0.55	10.36	Supported
Teacher immunity	→	Teacher psychological well-being	0.82	21.48	Supported
Work passion	→	Job satisfaction	0.47	7.52	Supported
Work passion	→	Occupational well-being	0.41	5.27	Supported
Work passion	→	Teacher psychological well-being	0.44	6.38	Supported

A complete analysis of the robustness of the causal relations among the variables is presented in Figures 2 and 3. These variables are also included in Table 6. During the course of the analysis, it was discovered that TI had a noteworthy and favorable influence on JS (β = 0.67, t = 14.25), OW-B (β = 0.55, t = 10.36), and PW-B (β = 0.82, t = 21.48). Furthermore, WP had a significant and favorable impact on JS (β = 0.47, t = 7.52), OW-B (β = 0.41, t = 5.27), and PW-B (β = 0.44, t = 6.38).

Figure 4

Path Coefficients for the Interaction Between TI, WP, JS, OW-B, and PW-B Subcomponents (Model 2)

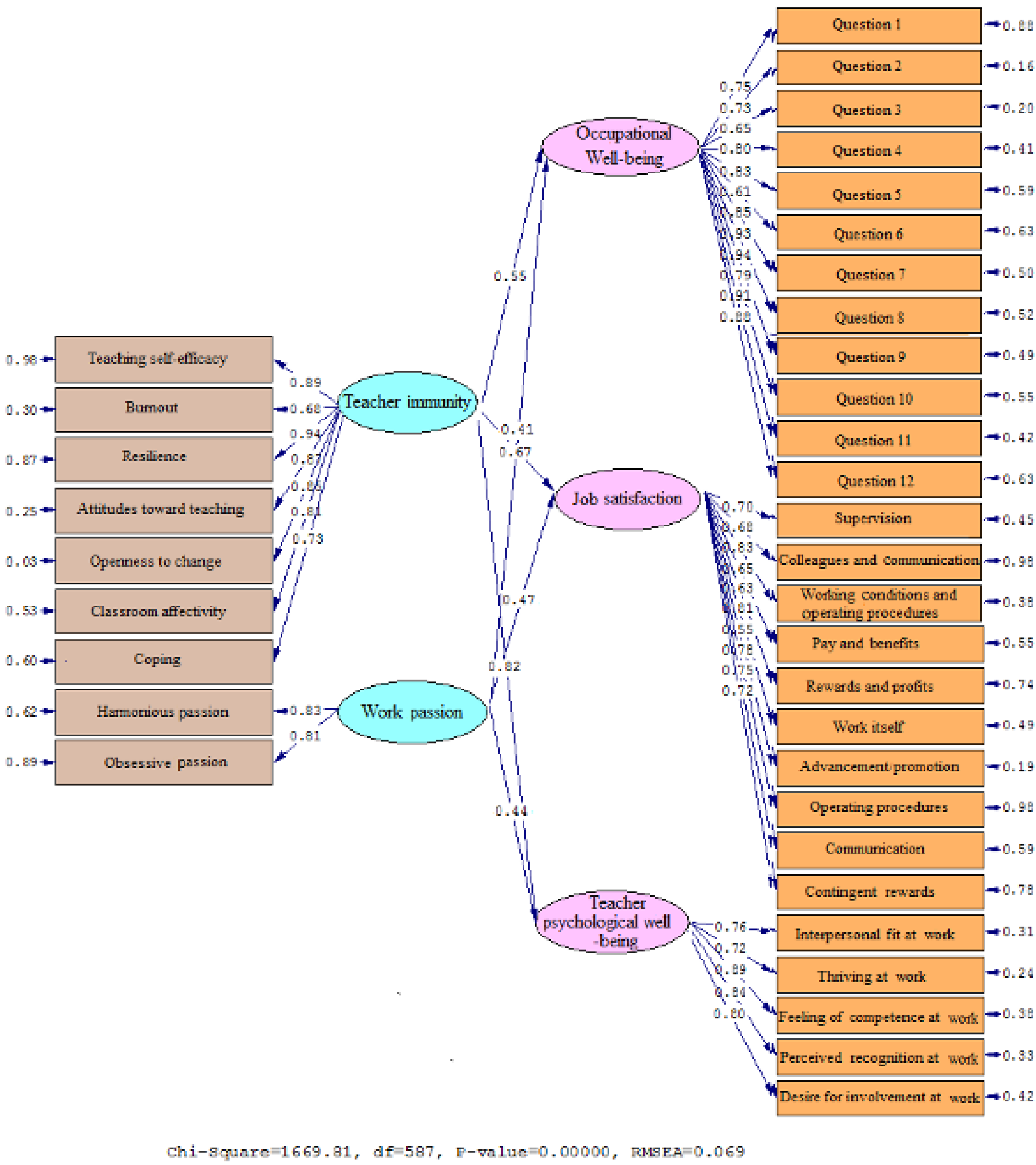
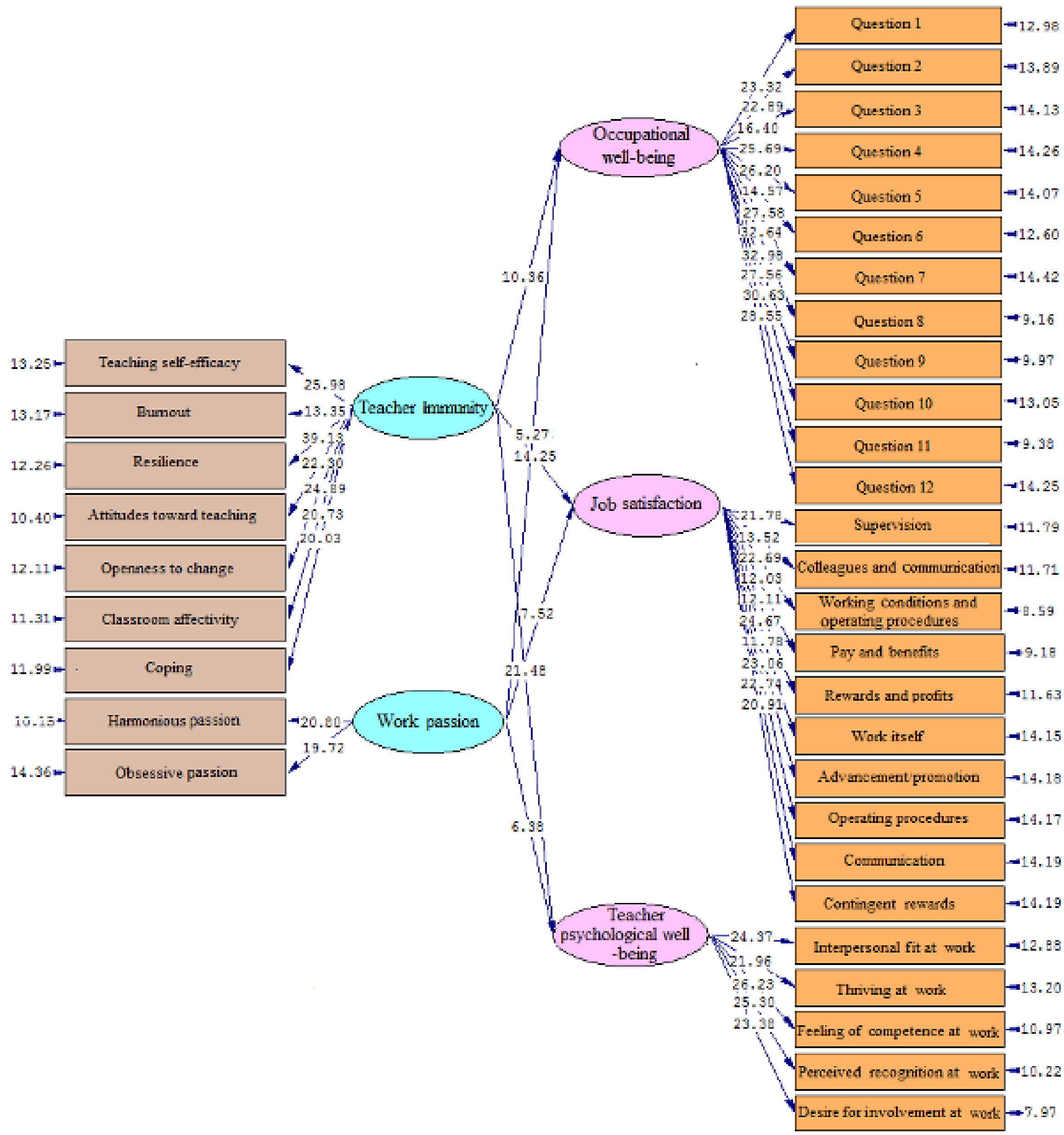


Figure 5

t-*Values Indicating the Relevance of Route Coefficients in Model 2*



Chi-Square=1669.81, df=587, P-value=0.00000, RMSEA=0.069

Table 6

Review of Model 2's Outcomes

		Path	Path coefficient	*t* statistic	Result
Teaching self-efficacy	→	Job satisfaction	0.71	14.33	Supported
Burnout	→	Job satisfaction	0.62	12.10	Supported
Resilience	→	Job satisfaction	0.73	15.76	Supported
Attitudes toward teaching	→	Job satisfaction	0.69	14.08	Supported
Openness to change	→	Job satisfaction	0.68	13.95	Supported
Classroom affectivity	→	Job satisfaction	0.65	13.43	Supported
Coping	→	Job satisfaction	0.63	12.88	Supported
Harmonious passion	→	Job satisfaction	0.48	6.59	Supported
Obsessive passion	→	Job satisfaction	0.46	6.42	Supported
Teaching self-efficacy	→	Occupational well-being	0.56	10.41	Supported
Burnout	→	Occupational well-being	0.50	7.19	Supported
Resilience	→	Occupational well-being	0.57	10.88	Supported
Attitudes toward teaching	→	Occupational well-being	0.60	11.75	Supported
Openness to change	→	Occupational well-being	0.54	9.86	Supported
Classroom affectivity	→	Occupational well-being	0.59	11.32	Supported
Coping	→	Occupational well-being	0.52	7.73	Supported
Harmonious passion	→	Occupational well-being	0.40	4.92	Supported

Obsessive passion	→	Occupational well-being	0.41	5.30	Supported
Teaching self-efficacy	→	Teacher psychological well-being	0.89	25.68	Supported
Burnout	→	Teacher psychological well-being	0.75	16.69	Supported
Resilience	→	Teacher psychological well-being	0.83	22.71	Supported
Attitudes toward teaching	→	Teacher psychological well-being	0.86	23.59	Supported
Openness to change	→	Teacher psychological well-being	0.81	20.95	Supported
Classroom affectivity	→	Teacher psychological well-being	0.80	20.32	Supported
Coping	→	Teacher psychological well-being	0.78	18.85	Supported
Harmonious passion	→	Teacher psychological well-being	0.45	6.13	Supported
Obsessive passion	→	Teacher psychological well-being	0.43	5.86	Supported

Figures 4 and 5 provide a graphical depiction of the route coefficients for the links between the TI, WP, JS, OW-B, and PW-B. These coefficients are also shown in a different format in Table 6. When it comes to the JS, TI, and WP subscales, there is a strong and essential connection between JS and the following subscales: teaching self-efficacy (β = 0.71, t = 14.33), burnout (β = 0.62, t = 12.10), resilience (β = 0.73, t = 15.76), attitudes toward teaching (β = 0.69, t = 14.08), openness to change (β = 0.68, t = 13.43), classroom affectivity (β = 0.78, t = 18.56), coping (β = 0.63, t = 12.88), harmonious passion coping (β = 0.48, t = 8.59), and obsessive passion (β = 0.46, t = 6.42).

In addition, it was observed that there was a noteworthy and favorable correlation between the subfactors of the OW-B, TI, and WP subscales. These subfactors include teaching self-efficacy (β = 0.56, t = 10.41), burnout (β = 0.50, t = 7.19), resilience (β = 0.57, t = 10.88), attitudes toward teaching (β = 0.60, t = 11.57), openness to change (β = 0.54, t = 9.86), classroom affectivity (β = 0.59, t = 11.32), coping (β = 0.52, t = 7.73), harmonious passion coping (β = 0.40, t = 4.92), and obsessive passion (β = 0.41, t = 5.30).

The subsequent inference is the result of an examination of the relationships between the PW-B, TI, and WP subcomponents: teaching self-efficacy (β = 0.89, t = 25.68), burnout (β = 0.75, t = 16.69), resilience (β

= 0.83, t = 22.71), attitudes toward teaching (β = 0.86, t = 23.59), openness to change (β = 0.81, t = 20.95), classroom affectivity (β = 0.80, t = 20.32), coping (β = 0.78, t = 18.85), harmonious passion coping (β = 0.45, t = 6.13), and obsessive passion (β = 0.43, t = 5.86).

Discussion

The aim of this study was to determine the interplay among TI, WP, JS, OW-B, and PW-B. The research revealed a significant and favorable link between TI, WP, JS, OW-B, and PW-B in the EFL setting while the instruction was integrated with AI. The findings of the first part of the research indicate that those who effectively and efficiently fortified their instruction would have been more adept at managing challenging circumstances and conflicts in the workplace. The results of the present study aligned with Rahmati et al.'s (2019) findings, which emphasized the importance of promoting contemplation as a method of enhancing TI.

More precisely, the results indicate that there is a correlation between the level of tenacity in following instructions, enthusiasm and determination in teaching, self-awareness, and attention toward others. Productive immunity, in accordance with the self-organization theory's principles, functions as a means of protection against various obstacles encountered in the workplace (Hiver, 2017). The research revealed a robust association between language teachers' efforts to adapt to changes and their cognitive capacities in this domain. It might be argued that higher level cognitive functions enhance self-awareness and that efficient and productive immunity is a consequence of self-organization. Job satisfaction fosters emotional equilibrium, leading to improved immune function and therefore increasing teachers' dedication to perseverance, purpose in the classroom, excitement, and their awareness of themselves and others.

Another perspective that may be used to comprehend the results of this research is self-organization theory. Language instructors may adapt to the novel circumstances brought about by AI in the classroom by employing productive immunity. In addition, the study's findings suggest that EFL teachers who adopt productive immunity all through their jobs have a better understanding of their instructional environment and the factors that affect their effectiveness. Previous research (e.g., Amirian et al., 2023; Namaziandost & Heydarnejad, 2023; Rahmati et al., 2019) has identified noteworthy correlations between professional achievement, self-efficacy, resilience, and exhaustion (which are the subscales of the LTII). However, the absence of previous research specifically examining the correlation between TI, WP, and JS precludes any ability to draw comparisons between this finding and others. Consequently, this study can inspire further research on the well-being of teachers at the time of AI applications in language education.

The results of the second research question (Does the EFL instructors' TI and WP offer any indication of their OW-B in AI-integrated instruction?) indicate that TI and WP predict the state of OW-B in AI-integrated language instruction. In accordance with positive psychology principles, this result may be supported. Similar to other domains within positive psychology, language education employs self-help principles to enhance the learning experience (Seligman, 2018). So, instructors who exhibit TI and WP are more likely to achieve intrapersonal and interpersonal mindfulness, which can lead to more success. They are less certain and more ardent in the classroom. The study found that positive interactions and peer support enhanced not only the resiliency and determination of EFL instructors in their classrooms but also their sense of purpose and significance. While definitive evidence linking TI, WP, and OW-B is still lacking, the research conducted by Zhang (2021) implies that increased engagement is linked to more persistent behavior as a teacher, which indirectly supports this result.

Additionally, it was demonstrated that TI and WP significantly influence a teacher's level of PW-B. Teachers who are immunized and dedicated are more likely to feel a sense of mission and importance in their teaching. Teachers who have acquired the appropriate immunizations are more likely to develop a sense of professional fulfillment, which in turn improves their overall health and satisfaction. Tolerance, self-efficacy in the classroom, fatigue, perseverance, teaching attitudes, adaptability readiness, and responsiveness in the classroom are all potential contributors to professional engagement. It would appear that educators who possess positive relationships with both students and colleagues, exhibit fruitful immunity, and develop engaging and impactful lesson plans are more inclined to exhibit feelings of competence and self-assurance in their vocation. This sense of competence and reliability may increase job satisfaction and contentment, thereby fostering psychological health as a whole (Noori, 2023). Additionally, it can be inferred that educators are less prone to feel exhausted and emotionally distressed when they have a sense of autonomy in their work and possess the necessary abilities and resources to confront any challenges that may emerge. This can be attributed to their exceptional capacity to manage stress and navigate challenging situations effectively, which eventually brings about improved psychological health. Educators who possess adaptive immunity are more likely to be inclined toward improvement, a trait that positively correlates with their mental and psychological health and assists them in managing stressors (Rahmati et al., 2019). Moreover, the results of the study indicate that educators who have immunity demonstrate an unwavering dedication to accomplishing their scholastic objectives and achieving success.

The implications of these findings for the design and implementation of AI-based programs and initiatives that aim to enhance the well-being of EFL instructors are substantial. As previous research witnessed (e.g., Jamal, 2023; O'Dea & O'Dea, 2023), AI can improve educators' skills by providing them with access to various tools and resources that may help them become more effective teachers. Assessment systems that are driven by AI may also provide instructors with real-time feedback on the performance of their students. This gives teachers the ability to modify their instructional methods to better comply with the requirements of the students. Additionally, AI may assist educators in personalizing learning by enabling them to develop classes that are tailored to meet the requirements to meet the requirements of each of their students.

Conclusion and Pedagogical Implications

The study underscore the significance of TI, WP, JS, OW-B, and PW-B, potentially providing educators with insights to enhance their pre-service and in-service curricula, especially in AI-supported language learning. The potential influence of educators' TI and WP on their responses to reform initiatives suggests that the findings of this research may inspire language instructors to employ strategies for productive immunization and engagement when instructing via AI. The integration of AI in language education only maintains effective teaching methods but also enhances language instruction, ensuring continuous progress in effective teaching practices (keep the ball rolling). Furthermore, it is highly recommended that policymakers consider the findings of the current study so as to develop a holistic comprehension of the elements that contribute to the efficacy of certain programs and instructors while rendering others ineffectual. Policy makers, language educators, and instructors must acknowledge the significance of language instructor immunity, considering the novelty and efficacy of this concept.

Further investigation may be warranted to address certain constraints that were identified in the current study. To begin with, further research is suggested to enhance the applicability of the results acquired across different higher education institutions nationwide, given that the participants were selected via convenience sampling and AI-based applications were used in teaching. Future research may employ mixed-methods designs to examine the correlation between TI, WP, JS, OW-B, and PW-B, as was done in this quantitative investigation, so as to offer a more detailed understanding of the matter. Moreover, due to the cross-sectional design of the current investigation, additional long-term studies are required to examine the relationships between TI, WP, JS, OW-B, and PW-B. Furthermore, additional descriptive variables, such as the demographics of the language instructors, were not examined in this study. Therefore, it is suggested that future research use demographic information regarding language instructors. Last but not least, additional investigation is necessary to determine the degree to which productive immunity, physiological well-being, buoyant inclinations, and learner engagement can serve as predictors of teacher success.

Acknowledgements

The researchers would like to confirm that AI-enhanced applications were used responsibly for editing and proofreading of the manuscript.

References

Amirian, S. M. R., Heydarnejad, T., & Abbasi-Sosfadi, S. (in press). The relationship among language teacher immunity, reflective teaching, and work motivation in EFL context. *Issues in Language Teaching*. https://doi.org/10.22054/ILT.2023.67179.691

Baluyos, G. R., Rivera, H. L., & Baluyos, E. L. (2019). Teachers' job satisfaction and work performance. *Open Journal of Social Sciences, 7*(8), 206–221. https://doi.org/10.4236/jss.2019.78015

Buitendach, J. H., & De Witte, H. (2005). Job insecurity, extrinsic and intrinsic job satisfaction andaffective organisational commitment of maintenance workers in a parastatal. *South African Journal of Business Management*, *36*(2), 27–37. https://doi.org/10.4102/sajbm.v36i2.625

Cabrita, C., & Duarte, A. P. (2023). Passionately demanding: Work passion's role in the relationship between work demands and affective well-being at work. *Frontiers in Psychology*, *14*, Article 1053455. https://doi.org/10.3389/fpsyg.2023.1053455

Chan, S. W., Pöysä, S., Lerkkanen, M. K., & Pakarinen, E. (2023). Teachers' occupational well-being in relation to teacher–student interactions at the lower secondary school level. *Scandinavian Journal of Educational Research*, 1–18. https://doi.org/10.1080/00313831.2023.2204114

Chiu, T. K. F., Moorhouse, L. M., Chai, C. S., & Ismailov, M. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environment, 1–17*. https://doi.org/10.1080/10494820.2023.2172044

Dagenais-Desmarais, V., & Savoie, A. (2012). What is psychological well-being, really? A grassroots approach from the organizational sciences. *Journal of Happiness Studies, 13*, 659–684. https://doi.org/10.1007/s10902-011-9285-3

de Boer, B. (2005). Self-organization in language. In C. Hemelrijk (Ed.), *Self-organization and evolution of social systems* (pp. 123–139). Cambridge Univerity Press.

dos Santos, P. I. G., & Rosinhas, A. C. (2023, September 15). Artificial intelligence and mental health. *Seven Editora*. https://sevenpublicacoes.com.br/index.php/editora/article/view/2618

Greenier, V., Derakhshan, A., and Fathi, J. (2021). Emotion regulation and psychological well-being in teacher work engagement: a case of British and Iranian English language teachers. System, 97, 102446. https://doi.org/10.1016/j.system.2020.102446

Haseli Songhori, M., Ghonsooly, B., & Afraz, S. (2018). Language teacher immunity among Iranian EFL teachers: A self-organization perspective. *Iranian Journal of English for Academic Purposes*. *7*(1), 128–143. https://journalscmu.sinaweb.net/article_87253.html

Hiver, P. (2015). Once burned, twice shy: The dynamic development of system immunity in language teachers. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 214–237). Multilingual Matters.

Hiver, P. (2017). Tracing the signature dynamics of language teacher immunity: A retrodictive qualitative modeling study. *The Modern Language Journal, 101*(4), 669–690. https://doi.org/10.1111/modl.12433

Hiver, P., & Dörnyei, Z. (2017). Language teacher immunity: A double-edged sword. *Applied Linguistics, 38*(3), 405–423. https://doi.org/10.1093/applin/amv034

Licardo, J. T., Domjan, M., & Orehovački, T. (2024). Intelligent robotics—A systematic review of emerging technologies and trends. *Electronics*, *13*(3), 542. https://doi.org/10.3390/electronics13030542

Jamal, A. (2023). The role of artificial intelligence (AI) in teacher education: Opportunities and challenges. *International Journal of Research and Analytical Reviews*, *10*(1), 139–146. http://ijrar.org/viewfull.php?&p_id=IJRAR23A2629

Kay, J. (2012). AI and education: Grand challenges. *IEEE Intelligent Systems*, *27*(5), 66–69. https://doi.org/10.1109/MIS.2012.92

Kohnke, L., Moorhouse, B. L., and Zou, D. (2023). Chat GPT for language teaching and learning. *RELC Journal*, 54, 537–550. https://doi.org/10.1177/00336882231162868

Lopes, J., & Oliveira, C. (2020). Teacher and school determinants of teacher job satisfaction: A multilevel analysis. *School Effectiveness and School Improvement, 31*(4), 641–659. https://doi.org/10.1080/09243453.2020.1764593

Luxton, D. D. (2016). An introduction to artificial intelligence in behavioral and mental health care. In D. D. Luxton (Ed.), *Artificial intelligence in behavioral and mental health care* (pp. 1–26). Elsevier Academic Press.

Michaelowa, K. (2002). *Teacher career satisfaction, student achievement, and the cost of primary education in francophone sub-Saharan Africa*. Hamburg Institute of International Economics. https://www.econstor.eu/bitstream/10419/19349/1/188.pdf

McInerney, D. M., Ganotice, F. A., King, R. B., Morin, A. J. S., & Marsh, H. W. (2015). Teachers' commitment and psychological well-being: Implications of self-beliefs for teaching in Hong Kong. *Educational Psychology, 35*(8), 926–945. https://www.doi.org/10.1080/01443410.2014.895801

Mercer, S. (2021). An agenda for well-being in ELT: An ecological perspective. *ELT Journal, 75*(1), 14–21. https://doi.org/10.1093/elt/ccaa062

Namaziandost, E., & Heydarnejad, T. (2023). Mapping the association between productive immunity, emotion regulation, resilience, and autonomy in higher education. *Asian-Pacific Journal of*

Second and Foreign Language Education, *8*, Article 33. https://doi.org/10.1186/s40862-023-00207-3

Noori, A. Q. (2023). Job satisfaction variance among public and private school teachers: A case study. *Cogent Education*, *10*(1), Article 2189425. https://doi.org/10.1080/2331186X.2023.2189425

O'Dea, X. (C.), & O'Dea, M. (2023). Is artificial intelligence really the next big thing in learning and teaching in higher education? A conceptual paper. *Journal of University Teaching & Learning Practice*, *20*(5). https://doi.org/10.53761/1.20.5.05

Parker, P. D., Martin, A. J., Colmar, S., & Liem, G. A. (2012). Teachers' workplace well-being: Exploring a process model of goal orientation, coping behavior, engagement, and burnout. *Teaching and Teacher Education, 28*(4), 503–513. https://doi.org/10.1016/j.tate.2012.01.001

Paterson, A., & Grantham, R. (2016). How to make teachers happy: An exploration of teacher wellbeing in the primary school context. *Educational and Child Psychology, 33*(2), 90–104. https://doi.org/10.53841/bpsecp.2016.33.2.90

Rahmati, T., Sadeghi, K., & Ghaderi, F. (2019). English as a foreign language teacher immunity: An integrated reflective practice. *Iranian Journal of Language Teaching Research, 7*(3), 91–107. https://doi.org/10.30466/ijltr.2019.120738

Randi, J. (2004). Teachers as self-regulated learners. *Teachers College Record*, *106*(9), 1825–1853. https://doi.org/10.1111/j.1467-9620.2004.00407.x

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. The Guilford Press.

Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, *57*, 1069-1081. https://doi.org/10.1037/0022-3514.57.6.1069

Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, *69*(4), 719–727. https://doi.org/10.1037/0022-3514.69.4.719

Saaranen, T., Tossavainen, K., Ryhanen, E., & Turunen, H. (2013). Promoting the occupational well-being of teachers for the Comenius program. *International Journal of Higher Education, 2*(2), 159–174. https://doi.org/10.5430/ijhe.v2n2p159

Seligman, M. (2018). PERMA and the building blocks of well-being. *The Journal of Positive Psychology*, *13*(4), 333–335. https://doi.org/10.1080/17439760.2018.1437466

Shan, M. H. (1998). Professional commitment and satisfaction among teachers in urban middle schools. *Journal of Educational Research*, *92*, 67–73. http://dx.doi.org/10.1080/00220679809597578

Siemens, G., Marmolejo-Ramos, F., Gabriel, F., Medeiros, K., Marrone, R., Joksimovic, S., & de Laat, M. (2022). Human and artificial cognition. *Computers and Education: Artificial Intelligence*, *3*, 100107. https://doi.org/10.1016/j.caeai.2022.100107

Simmons, M., McDermott, M., Lock, J., Crowder, R., Hickey, E., DeSilva, N., Leong, R., & Wilson, K. (2019). When educators come together to speak about well-being: An invitation to talk. *Canadian Journal of Education/Revue canadienne de l'education, 42*(3), 850–872. https://journals.sfu.ca/cje/index.php/cje-rce/article/view/3959/2773

Spector, P. E. (1985). Measurement of human service staff satisfaction: Development of the Job Satisfaction Survey. *American Journal of Community Psychology, 13*(6), 693–713. https://doi.org/10.1007/BF00929796

Terra, M., Baklola, M., Ali, S., & Karim El-Bastawisy. (2023). Opportunities, applications, challenges and ethical implications of artificial intelligence in psychiatry: A narrative review. *The Egyptian Journal of Neurology, Psychiatry and Neurosurgery*, *59*(1). https://doi.org/10.1186/s41983-023-00681-z

Vallerand, R. J., & Houlfort, N. (2003). Passion at work: Toward a new conceptualization. In S. Gilliland, D. Steiner, & D. Skarlicki (Eds.), *Emerging perspectives on values in organizations* (pp. 175–204). Information Age Publishing.

Vallerand, R. J., Salvy, S.-J., Mageau, G. A., Elliot, A. J., Denis, P. L., Grouzet, F. M. E., & Blanchard, C. (2007). On the role of passion in performance. *Journal of Personality, 75*(3), 505–534. https://doi.org/10.1111/j.1467-6494.007.00447.x

Van Horn, J. E., Taris, T. W., Schaufeli, W. B., & Schreurs, P. J. (2004). The structure of occupational well-being: A study among Dutch teachers. *Journal of Occupational and Organizational Psychology*, *77*(3), 365–375. https://doi.org/10.1348/0963179041752718

Warr, P. (1990). The measurement of well-being and other aspects of mental health. *Journal of Occupational and Organizational Psychology*, *63*(3), 193–210. https://doi.org/10.1111/j.2044-8325.1990.tb00521.x

Warr, P. B. (1994). A conceptual framework for the study of work and mental health. *Work and Stress*, *8(2)*, 84–97. https://doi.org/10.1080/02678379408259982

Werang, B. R., Agung, A. A. G., & Agung, G. (2017). Teachers' job satisfaction, organizational commitment, and performance in Indonesia: A study from Merauke District, Papua. *International Journal of Development and Sustainability, 6*(8), 700–711. https://isdsnet.com/ijds-v6n8-13.pdf

Wessels, E., & Wood, L. (2019). Fostering teachers' experiences of well-being: A participatory action learning and action research approach. *South African Journal of Education*, *39*(1), 1–10. *https://doi.org/10.15700/saje.v39n1a1619*

Zhang F. (2021). Toward the impact of job satisfaction and collective efficacy on EFL teachers' professional commitment. *Frontiers in Psychology*, *13*, 938125. https://doi.org/10.3389/fpsyg.2022.938125

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

The Metaphor of AI in Writing in English: A Reflection on EFL Learners' Motivation to Write, Enjoyment of Writing, Academic Buoyancy, and Academic Success in Writing

Ying He
Department of Humanities, Ningbo University of Finance and Economics, Ningbo 315175 China

Abstract

Several barriers hinder students from producing clear and impactful written work. Writing assignments are often given on an individual basis, similar to homework, and without any assistance. Students in a classroom context have access to both their classmates and the teacher while they are working in groups or pairs as part of their assignments. The majority of students, however, are clueless about how to begin their homework assignments. The introduction of artificial intelligence in education may help solve this problem. The current research intended to demonstrate the effects of employing automated writing evaluation (AWE) in fostering learners' writing skills, motivation to write, enjoyment of writing, and academic buoyancy in open and distributed English as a foreign language (EFL) learning. The participants were 86 intermediate EFL students from China. The participants in the experimental group (n = 44) received instruction and feedback from their teachers only; participants in the control group (n = 42) were exposed to their teachers' instruction as well as AWE. The results of data analysis via one-way multivariate analysis of variance indicated that the participants in the experimental group outperformed their peers in the control group in motivation to write, enjoyment in writing, academic buoyancy, and academic success in writing. Further in-depth discussions proceed regarding the implications of the study.

Keywords: artificial intelligence, motivation to write, enjoyment in writing, academic buoyancy, academic success in writing, automated writing evaluation, open and distributed learning

Introduction

Learning a language and expressing oneself in writing are two of the most important aspects of education (Jin, 2023; Sun & Wang, 2020). Students' intellectual and personal development are both aided by efforts to improve their writing skills. Teaching students to feel positive emotions while they write is an excellent way to improve their writing abilities (Richards, 2022; White & Arndt, 1991). Therefore, more research is needed on students' positive emotions in second- and first-language writing courses, along with the development of innovative methods to assess students' positive emotions, such as enthusiasm, toward writing.

In the field of artificial intelligence (AI)-integrated writing, AI generates automated writing using a variety of techniques, such as machine learning, deep learning, and neural networks. AI is specifically engineered to use these identified habits and frameworks to produce new content that adheres to a particular approach or tone. Regretfully, not much research has been done on AI in language instruction, and studies that have been done have only just started to pinpoint important areas that need more focus. Consequently, it is critical to investigate the potential applications of AI further to determine how it may be used in the teaching of practical language skills. It is also critical to consider how students may benefit from the real-time feedback provided by AI algorithms in achieving their language learning goals, as well as how various forms of feedback might enhance students' computer and mobile self-study. The advantages of using AI to improve writing skills in language learning environments are examined in this study.

Automated writing evaluation (AWE) uses a Web-based platform to offer rapid evaluation and constructive feedback on student-submitted written assignments (Tang & Wu, 2017). When teachers include AWE in their writing instruction, they have the potential to speed up the evaluation process without increasing their workload. An AWE system allows students to submit their written work and receive comments, anytime and anywhere, without human assistance (Liao, 2016). However, the successful implementation of any technological system depends on a thorough understanding of the user approval processes (Davis, 1989). An analysis of the literature revealed that previous studies on AWE primarily focused on its inter-rater reliability as evaluated by human graders (Cao, 2020; Mohsen & Abdulaziz, 2019; Zhai & Ma, 2022) and validation frameworks (Chapelle et al., 2015; Lang et al., 2019). To our knowledge, only one empirical study has specifically examined the individual traits that influence learners' inclination to use AWE (R. Li et al., 2019; Zhai & Ma, 2022). Using AWE may impact the way students think and feel.

Motivation is essential for a student's psychological health and is acknowledged as a key aspect that impacts their success in learning a second or foreign language. This concept focuses on how learners are drawn to academic subjects, influencing their behavior, attitudes toward learning, and responses to challenges. In this regard, Mo (2012) discovered that English writing issues may be remedied by increasing students' interest in and participation in writing courses that emphasize writing exercises. It is worth highlighting that the major hypothesis in understanding motivation is self-determination theory (SDT), developed by Ryan and Deci (2000). SDT posited three fundamental elements of motivation: (a) inherent motivation, (b) external motivation, and (c) lack of motivation.

Within the field of motivation, it was shown that students' ability to write improved when they were allowed to pursue their own intrinsic desires in the context of academic writing classes (LaSalle, 2015). Lo and

Hyland (2007) looked at how students' writing improved when they were motivated to write. Students' interest in the subjects being written about increases motivation, which in turn enhances writing success, particularly for students who were previously performing inadequately. Other studies examined the influence of motivation on the usage of writing methods (e.g., Lee & Wong, 2014; Nasihah & Cahyono, 2017). However, research has not specifically focused on motivation in writing. Therefore, the goal of this research was to look at the link between EFL learners' motivation to write and their writing skills, particularly when they use AI in their writing.

Additionally, this study focuses on academic buoyancy, which is another student-related issue. Academic buoyancy originates from the field of positive psychology (Xu & Wang, 2022). According to positive psychology, it is important to emphasize positive and self-help qualities in the field of language instruction and learning to speed up learners' progress (Jin & Dewaele, 2018). MacIntyre and Mercer (2014) asserted that positive psychology provides significance and reinforcement for the process of language acquisition and teaching. More recently, Jahedizadeh et al. (2019) developed and verified a specialized tool to evaluate academic buoyancy in the unique context of English as a second language/foreign language.

According to Xu and Wang (2022), academic buoyancy is being linked to thoughts, feelings, and actions that are good for learning and academic success. Academic buoyancy has been found to be constructively correlated with involvement, proficiency, dedication, self-efficacy, organizing, tenacity, and pleasant achievement emotions (Ding & He, 2022; Jahedizadeh et al., 2019; Xu & Wang, 2022). On the other hand, academic buoyancy is unfavorably correlated with nervousness, uncertain control, and undesirable accomplishment (Martin, 2013, 2014). More frequent usage of learning techniques is associated with greater levels of academic buoyancy, according to studies (e.g., Martin et al., 2013). This finding is based on the premise that a student who is self-assured in their ability to handle difficulties may focus less on obstacles and be more likely to use effective learning practices. From an empirical point of view, Martin and Marsh (2006, 2008) discovered that academic buoyancy predicted class engagement, general self-worth, and school fulfillment among high school students in Australia. Additionally, they established that academic buoyancy predicted educational participation and nonacademic repercussions. Heydarnejad et al. (2022) and Nurjamin et al. (2023) found that EFL students' psychological and mental involvement was directly impacted by their academic buoyancy. Yun et al. (2018) discovered that among South Korean college students, academic buoyancy was a strong predictor of second-language proficiency and grade point average. Therefore, our hypothesis was that academic buoyancy would be a good predictor of self-regulated learning writing methods when used in a second-language writing setting.

Purpose of the Current Research

There is a deficiency in the amount of in-depth research that has been conducted on the fundamental factors that influence students' adoption of AWE. According to the best of the researcher's knowledge, no study has previously investigated how AWE affects the psychological well-being of students or their academic progress in open and distributed learning. That is, EFL learners' motivation to write, enjoyment in writing, academic buoyancy, and academic success in writing have not been evaluated in any single study. As a result, the following research question and null hypothesis were developed, taking into account the background of the study as well as the review of previous research that was comparable to it:

Research question: Does applying AWE foster EFL learners' motivation to write, enjoyment in writing, academic buoyancy, and academic success in open and distributed learning?

Null hypothesis: Applying AWE does not foster EFL learners' motivation to write, enjoyment in writing, academic buoyancy, or academic success in open and distributed learning.

Methods

This study employed a quasi-experimental design that includes both a pretest and a posttest. The acts that were performed are elaborated upon in further detail in the subsequent paragraphs.

Participants and Procedures

A random sample of 86 participants was chosen from a larger group of 147 first-year EFL students who were enrolled at the Ningbo University of Finance and Economics, China. Based on the results of the Oxford Quick Placement Test, the participants' level of English-language proficiency was concluded to be intermediate. As an extra point of interest, over the course of the study, participants did not take part in any additional English classes. As a result, participants' English-language competence was assumed to be roughly equivalent. The ages of the participants ranged from 19 to 25, and they hailed from a wide spectrum of socioeconomic and cultural backgrounds. The control group (CG) consisted of 44 students, while the experimental group (EG) consisted of 42 individuals. Throughout the first semester of the academic year, students were obliged to attend 16 sessions of an English writing class.

A preliminary assessment was conducted before administering the treatment. After the pretest, a researcher, who was also the instructor for all the courses attended by both the EG and the CG, was responsible for delivering instruction. The study was carried out throughout a single academic semester in 2022 (16 sessions). The CG students received online instruction via the use of webinar software (Adobe Connect). Conversely, the EG acquired and was exposed to online instruction through Adobe Connect, and their writing skills were reinforced with AWE.

A posttest was conducted at the end of the semester after all instructional sessions were completed. The objective of the examination was to assess the achievement of both CG and EG students, as well as to ascertain the degree to which the program had been effective. Both the pretest and the posttest were evaluated by four EFL teachers to ensure that the findings were accurate. Data collection was concluded by taking into account the pretest scores of each student in addition to their posttest averages. The questionnaire was in English since all respondents fulfilled the standards required to grasp English.

Measures

Oxford Quick Placement Test

The Oxford Quick Placement Test (OQPT) was used to assess the participants' English proficiency. OQPT results range from 0.1 to 0.9, with 0.4–0.6 indicating an intermediate level of English proficiency. A reliability of 0.91 was recorded for the OQPT in this study.

Learners' Writing Skills Assessments

Two assessments requiring students to produce two different kinds of essays were used to gauge the students' writing ability. Students were to produce a written piece on process evaluation for the initial test and a written work on causality and impact for the subsequent exam. The mean results of the two essays were used to calculate the students' writing competence. The ESL Composition Profile scoring rubric, created by Hartfiel et al. (1985), was employed. Content accounts for 30% of the components, followed by structure, vocabulary, language used, and mechanics. The content component concentrated on the students' subject-matter knowledge as well as the coherent development of the thesis statement and its supporting information. The arrangement focuses on the degree to which learners put their thoughts for each sort of writing.

Learners' Motivation to Write Scale

To assess the students' motivation in writing, the questionnaire designed and validated by Cahyono and Rahayu (2020) was applied. This is a 6-point Likert scale including 23 items in 7 subsections. In the present investigation, the reliability of this scale was assessed and the result was considered sufficient ($r = 0.89$).

Enjoyment of Writing Scale

To gauge the learners' enjoyment of writing, the scale developed by Jin (2023) was used. This scale includes nine items rated on a 5-point Likert scale. Participants were expected to share their sentiments on writing. The dependability of this scale was assessed, and the outcome was reasonable ($\alpha = 0.911$).

Academic Buoyancy Scale

The participants' academic buoyancy was measured using the academic buoyancy scale (ABS) created and verified by Jahedizadeh et al. (2019). The 27 questions that made up this test assessed four aspects of second-language buoyancy: sustainability, regulation adaptability, positive personal eligibility, and acceptance of academic life. In addition, the ABS uses a 5-point Likert scale, where 1 indicates strong disagreement and 5 indicates strong agreement. Cronbach's alpha ranged from 0.824 to 0.876: the ABS dependability in this study was satisfactory.

Statistical Analysis

To examine the data, a one-way multivariate analysis of variance (MANOVA) was performed. Before calculating the MANOVA, it was necessary to conduct analyses of the related hypotheses. Several parameters were considered, including data normality, sample size, outlier presence or absence, data linearity, and regression homogeneity.

Findings

A one-way MANOVA was used to compare the pretest and posttest scores of EG and CG learners in terms of their English writing abilities. This statistical test is used when there is a single independent variable (in this instance, the implementation of AI in both the EG and CG) and two or more interconnected dependent variables, which are the subcomponents of writing in this study (motivation to write, enjoyment in writing, academic buoyancy, and academic success in writing). It is standard practice to verify that all MANOVA's assumptions (e.g., normality, sample size, outliers, linearity, regression homogeneity) are true before running the analysis. Tables 1 and 2 compare the English writing pretest scores of EG and CG students.

Table 1

Descriptive Statistics Results Comparing EG and CG on Pretest English Writing Scores

Pretest subsection	Group	*M*	*SD*	*n*
ASW	EG	27.09	7.78	44
	CG	28.02	8.04	42
	Total	27.54	7.87	86
LMW	EG	43.50	26.23	44
	CG	53.69	25.76	42
	Total	48.47	26.35	86
EW	EG	29.65	6.65	44
	CG	28.80	6.47	42
	Total	29.24	6.54	86
AB	EG	34.97	18.27	44
	CG	40.73	19.02	42
	Total	37.79	18.76	86

Note. EG = experimental group; CG = control group; ASW = academic success in writing; LMW = learners' motivation to write; EW = enjoyment in writing; AB = academic buoyancy.

Table 1 displays the mean scores of the EG and CG for the learners' motivation to write, enjoyment in writing, academic buoyancy, and academic success in writing in the pretest. There was some disparity between the two groups' mean scores on each subcomponent of writing in English, but the discrepancies were not significant. To determine whether the differences under consideration were statistically significant, the researcher consulted the MANOVA table (Table 2).

Table 2

MANOVA Results Comparing EG and CG on Pretest English Writing Scores

Effect		Value	*F*	Hypothesis *df*	Error *df*	*p*	Partial eta²
Groups	Pillai's trace	.05	1.23	4.00	81.00	.30	.05
	Wilks's lambda	.94	1.23	4.00	81.00	.30	.05
	Hotelling's trace	.06	1.23	4.00	81.00	.30	.05
	Roy's largest root	.06	1.23	4.00	81.00	.30	.05

Note. EG = experimental group; CG = control group.

Wilk's lambda is the most commonly reported statistic; hence, its value is stated here (.94). The related Wilk's lambda significance value was discovered to be .30, which is more than the significance threshold (.30 >.05). This demonstrates that the two groups, EG and CG, did not vary substantially on their pretest scores in terms of (the four subcomponents of) writing in English. Table 3 displays the findings of a similar data analysis approach used for the EG's and CG's writing in English posttest scores. Any posttest changes might be traced to the EG treatment (i.e., employing the AI).

Table 3

Descriptive Statistics Results Comparing EG and CG on Posttest English Writing Scores

Posttest subsection	Group	*M*	*SD*	*n*
ASW	EG	36.61	5.77	44
	CG	35.42	17.59	42
	Total	36.03	12.90	86
LMW	EG	97.31	21.57	44
	CG	63.14	22.14	42
	Total	80.62	27.69	86
EW	EG	38.11	5.31	44
	CG	37.47	15.94	42
	Total	37.80	11.70	86
AB	EG	88.31	28.87	44
	CG	70.07	60.14	42
	Total	79.40	47.44	86

Note. EG = experimental group; CG = control group; ASW = academic success in writing; LMW = learners' motivation to write; EW = enjoyment in writing; AB = academic buoyancy.

According to Table 3, there was a discrepancy between the EG and CG posttest mean scores for academic success in writing, motivation to write, enjoyment in writing, and academic buoyancy. Nevertheless, the researcher referred to the MANOVA table (Table 4) to determine whether these differences were statistically significant.

Table 4

MANOVA Results Comparing EG and CG on Posttest English Writing Scores

Effect		Value	*F*	Hypothesis *df*	Error *df*	*p*	Partial eta²
Groups	Pillai's trace	.41	14.49	4.00	81.00	.00	.41
	Wilks's lambda	.58	14.49	4.00	81.00	.00	.41
	Hotelling's trace	.71	14.49	4.00	81.00	.00	.41
	Roy's largest root	.71	14.49	4.00	81.00	.00	.41

Note. EG = experimental group; CG = control group.

When compared to the significance level (.00 < .05), the related significance value of Wilks's Lambda was .00, which is higher than the significance threshold. The presence of a *p* value that is either lower than or equal to the significance threshold indicates a meaningful difference between the two groups. As a result, the EG and the CG scored substantially differently on their posttests concerning the composite dependent variable of writing in English. Table 5 indicates which of the four subcomponents of wring in English was responsible for the difference between the two groups.

Table 5

Test of Between-Subjects Effects on Writing in English

Dependent variable	Type III sum of squares	*df*	*M* square	*F*	*p*	Partial eta²
ASW	30.17	1	30.178	0.179	.67	.00
LMW	25,097.40	1	25,097.40	52.55	.00	.38
EW	8.73	1	8.732	0.06	.80	.00
AB	7,154.42	1	7,154.42	3.26	.07	.03

Note. ASW = academic success in writing; LMW = learners' motivation to write; EW = enjoyment in writing; AB = academic buoyancy.

A more rigorous significance threshold was recommended to prevent type I errors, given that we examined many independent studies in this case. Bonferroni's adjustment, which involves dividing the number of analyses by the significance level (i.e., .05), is the most used method for this. The original significance level of .052 was divided by 4 since there were four dependent variables in this example. If the significance level (*p*) was less than .012, the findings would be considered significant. Learners' desire to write had a *p* value of .00 (< .012; Table 5). On the other hand, every *p* value examined was higher than the threshold of significance. This indicates that the treatment given to the EG learners had a substantial impact on their desire to write when compared with the CG learners. That is, the disparity between the EG and the CG on the academic writing posttest was caused by the learners' motivation to write.

Discussion

The primary goal of this study was to demonstrate the utility of appropriate AI applications in the context of English-language acquisition, focusing specifically on the development of writing skills through possibilities derived from AWE. The results obtained from this research provide insight into the potential benefits of incorporating AI feed-based tasks into writing training. More precisely, the outcomes underscore the encouraging prospects of AWE as a platform for English writing, proof writing, and editing in open and distributed learning (ODL). The EG students' performance was superior to that of their peers in the CG; therefore, the null hypothesis was rejected: Applying AWE does not foster EFL learners' motivation to write, enjoyment in writing, academic buoyancy, and academic success in ODL.

This study aimed to determine whether EFL students' motivation to write improved by applying AI in ODL. The data analysis demonstrated a positive correlation between EFL students' writing ability and motivation to write. This indicates that offering an effective tool for feedback and reinforcement in writing enhances the motivation of learners in writing. Students who want to excel in writing should consistently increase their drive to attend writing classes and complete course tasks. The present study's findings align with prior research by Brown (2001) and Alves-Wold et al. (2023), indicating that intrinsic motivation plays a crucial role in the effectiveness of language learning. According to Ryan and Deci (2000), intrinsic drive, rather than extrinsic rewards, plays a significant role in determining the success of a language learner. In the same line of inquiry, Ryan and Deci (2009) and Dougherty et al. (2015) emphasized the significance of intrinsic motivation in enhancing language learners' success.

Incorporating motivating elements into writing education is crucial according to this study's findings. This is particularly true when it comes to engaging students, making effective use of instructional material, and fostering positive relationships between students and instructors. All students should have enough feedback and interesting exercises so they can take pleasure in writing instruction and ultimately become better writers. Establishing rapport may be achieved via teacher–student conferences and the occasional use of icebreakers during writing time. To engage EFL learners, students should participate in activities centered on topic selection and writing tasks before starting to complete writing projects, as well as having conversations on interesting themes they are already familiar with.

Based on the findings of Zumbrunn et al. (2019) as well as Myhill et al. (2023), instructors may improve their students' enjoyment of and satisfaction in writing by adjusting writing tasks' level of difficulty. Such obstacles were overcome with the use of AI in the current study. Academic buoyancy may have a crucial role in eliciting positive emotional responses, as seen by its impact on second-language writing. This study's findings suggest that students' beliefs about their ability to control their English writing processes and results are important in English writing. This supports and broadens the control-value theory in the context of second-language writing.

Students find AI-based learning tools beneficial for academic studies, drafting, and other aspects of the writing process (Sumakul et al., 2022a, 2022b; Utami et al., 2023). Although these tools may not have all the capabilities writers need, they are easy to use and adaptable. Some students think that AI may be a fun way to spice up their writing lessons. The results suggest that college students are more likely to regard the AWE system as beneficial and use it for writing instruction when they discover it is simple to manage.

Conversely, college students prioritize the practicality of AWE over its user-friendliness. The efficacy of AWE feedback has emerged as a critical determinant in convincing college students to embrace it. Previous studies on AWE have shown similar results (e.g., Li et al., 2015; Li et al., 2023; Zhai & Ma, 2022). The effectiveness of AWE feedback has emerged as a critical determinant in persuading EFL students to adopt it. Previous research on the technology acceptance model has demonstrated similar results (e.g., Z. Li, 2021; Tian & Zhou, 2020).

Conclusion and Implications

Implementing AI is advisable for its ability to create an engaging and intriguing technical setting, particularly in open and dispersed language learning. This will enhance language learners' ability to effectively practice writing skills, resulting in overall enhancement of their proficiency in self-evaluation and of their writing output. The findings of this research indicate that options and feedback aided by AI led to enhancements in EFL learners' writing skills, motivation to write, enjoyment in writing, academic buoyancy, and academic success in writing. AI technology may therefore help learners oversee and regulate their educational processes. These technologies may assist in establishing objectives, monitoring accomplishments, and implementing any necessary adjustments. AI-driven training empowers learners to take control of their educational process and enhance their oral communication skills by offering personalized coaching and adaptive tasks that promote the development of meta-cognitive strategies.

Limitations and Suggestions for Future Research

In addition to the implications of this study, a number of limitations need to be addressed. Chinese students with an intermediate level of English proficiency participated in this research, which focused on ODL. This study may be replicated in other educational environments in the future to provide a comparative analysis of the current results. Research on the potential effects of employing AI to improve writing abilities on motivation, pleasure, academic achievement, and other aspects of academic success in various fields of study is strongly encouraged. Future studies may also consider other language skills, such as listening, reading, and writing, to evaluate the psychological and academic benefits that may be achieved via the use of AI in ODL environments.

The findings of this study are limited in their relevance since the subjects were picked for a quasi-experimental procedure. There is an acute need for extensive study to be conducted on the effects of using a variety of applications on the levels of motivation, pleasure, academic achievement, and academic success in writing. In addition, little consideration was given to the demographics of the participants. The students' demographic data will be important to gather data in similar future studies. Though this study used a quantitative methodology, a mixed-methods examination might provide more accurate findings.

In conclusion, future studies may investigate the connection between motivation, pleasure, academic buoyancy, and other learner-attributed variables, as well as other fundamental abilities.

Acknowledgments

The authors would like to acknowledge that they used AI applications for editing and proofreading some sections for accuracy and fluency.

References

Alves-Wold, A., Walgermo, B. R., McTigue, E., & Per Henning Uppstad, H. P. (2023). Assessing writing motivation: A systematic review of K-5 students' self-reports. *Educational Psychology Review*, *35*, Article 24. https://doi.org/10.1007/s10648-023-09732-6

Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy* (2nd ed.). Pearson Education.

Cahyono, B. Y., & Rahayu, T. (2020). EFL students' motivation in writing, writing proficiency, and gender. *TEFLIN Journal*, *31*(2), 162–180. https://doi.org/10.15639/teflinjournal.v31i2/162-180

Cao, L. L. (2020). Comparison of automatic and expert teachers' rating of computerized English listening–speaking test. *English Language Teaching*, *13*(1), 18–30. https://doi.org/10.5539/elt.v13n1p18

Chapelle, C. A., Cotos, E., & Lee, J. (2015). Validity arguments for diagnostic assessment using automated writing evaluation. *Language Testing*, *32*(3), 385–405. https://doi.org/10.1177/0265532214565386

Davis, F. D. (1989). Perceived usefulness, perceived ease of use and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Ding, J., & He, L. (2022). On the association between Chinese EFL teachers' academic buoyancy, self-efficacy, and burnout. *Frontiers in Psychology, 13*, 1-11. https://doi.org/10.3389/fpsyg.2022.947434

Dougherty, P. (2015). Creative writing in English and EFL students: Finding motivation in self-expression. *East West Journal of Humanities*, *5*, 1–14. http://lib.ewubd.edu:8080/handle/2525/2785

Hartfiel, V. F., Jacobs, H. L., Zinkgraft, S. A., Wormuth, D. R., & Hughey, J. B. (1985). *Learning ESL composition*. Newbury House.

Heydarnejad, T., Abdel Al Ibrahim, K. A., Abdelrasheed, N. S. G., & Rezvani, E. (2022). The effect of academic emotion regulation on EFL learners' core of self-assessment and academic buoyancy: A structural equation modeling. *Language Testing in Asia*, *12*, Article 57. https://doi.org/10.1186/s40468-022-00207-z

Jahedizadeh, S., Ghonsooly, B., & Ghanizadeh, A. (2019). Academic buoyancy in higher education. *Journal of Applied Research in Higher Education*, *11*(2), 162–177. https://doi.org/10.1108/JARHE-04-2018-0067

Jin, Y. (2023). The development and validation of the English writing enjoyment scale. *Perceptual and Motor Skills*, *130*(1), 555–575. https://doi.org/10.1177/00315125221137649

Jin, Y. X., & Dewaele, J. M. (2018). The effect of positive orientation and perceived social support on foreign language classroom anxiety. *System*, *74*, 149–157. https://doi.org/10.1016/j.system.2018.01.002

Lang, F., Li, S., & Zhang, S. (2019). Research on reliability and validity of mobile networks-based automated writing evaluation. *International Journal of Mobile Computing and Multimedia Communications*, *10*(1), 18–31. https://doi.org/10.4018/IJMCMC.2019010102

LaSalle, D. (2015). Intrinsic motivation and the five-paragraph essay: Lessons learned on practitioner research, the role of academic research in the classroom, and assessing changes in student motivation. *Perspectives on Urban Education*, *12*(1). https://urbanedjournal.gse.upenn.edu/archive/volume-12-issue-1-spring-2015/intrinsic-motivation-and-five-paragraph-essay-lessons-learned-

Lee, I., & Wong, K. (2014). Bringing innovation to EFL writing: The case of a primary school in Hong Kong. *The Asia-Pacific Education Researcher*, *23*(1), 159–163. https://doi.org/10.1007/s40299-013-0149-y

Li, W.-Y., Kau, K., & Yi-Jiun Shiung. (2023). Pedagogic exploration into adapting automated writing evaluation and peer review integrated feedback into large-sized university writing classes. *SAGE Open, 13*(4), 1-18. https://doi.org/10.1177/21582440231209087

Li, J., Link, S., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing*, *27*, 1–18. https://doi.org/10.1016/j.jslw.2014.10.004

Li, R., Meng, Z., Tian, M., Zhang, Z., Ni, C., & Xiao, W. (2019). Examining EFL learners' individual antecedents on the adoption of automated writing evaluation in China. *Computer Assisted Language Learning*, *32*(7), 784–804. https://doi.org/10.1080/09588221.2018.1540433

Li, Z. (2021). Teachers in automated writing evaluation (AWE) system-supported ESL writing classes: Perception, implementation, and influence. *System*, *99*, Article 102505. https://doi.org/10.1016/j.system.2021.102505

Liao, H. C. (2016). Using automated writing evaluation to reduce grammar errors in writing. *ELT Journal*, *70*(3), 308–319. https://doi.org/10.1093/elt/ccv058

Lo, J., & Hyland, F. (2007). Enhancing students' engagement and motivation in writing: The case of primary students in Hong Kong. *Journal of Second Language Writing*, *16*(4), 219–237. https://doi.org/10.1016/j.jslw.2007.06.002

MacIntyre, P. D., & Mercer, S. (2014). Introducing positive psychology to SLA. *Studies in Second Language Learning and Teaching*, *4*(2), 153–172. https://doi.org/10.14746/ssllt.2014.4.2.2

Martin, A. J. (2013). Academic buoyancy and academic resilience: Exploring "everyday" and "classic" resilience in the face of academic adversity. *School Psychology International*, *34*(5), 488–500. https://doi.org/10.1177/0143034312472759

Martin, A. J. (2014). Academic buoyancy and academic outcomes: Towards a further understanding of students with attention-deficit/hyperactivity disorder (ADHD), students without ADHD, and academic buoyancy itself. *British Journal of Educational Psychology*, *84*(1), 86–107. https://doi.org/10.1111/bjep.12007

Martin, A. J., Ginns, P., Brackett, M. A., Malmberg, L. E., & Hall, J. (2013). Academic buoyancy and psychological risk: Exploring reciprocal relationships. *Learning and Individual Differences*, *27*, 128–133. https://doi.org/10.1016/j.lindif.2013.06.006

Martin, A. J., & Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools*, *43*(3), 267–281. https://doi.org/10.1002/pits.20149

Martin, A. J., & Marsh, H. W. (2008). Academic buoyancy: Towards an understanding of students' everyday academic resilience. *Journal of School Psychology*, *46*(1), 53–83. https://doi.org/10.1016/j.jsp.2007.01.002

Mo, H. (2012). A study of the teaching of ESL writing in colleges in China. *International Journal of English Linguistics*, *2*(1), 118–127. https://doi.org/10.5539/ijel.v2n1p118

Mohsen, M. A., & Abdulaziz, A. (2019). The effectiveness of using a hybrid mode of automated writing evaluation system on EFL students' writing. *Teaching English With Technology*, *19*(1), 118–131. https://eric.ed.gov/?id=EJ1204544

Myhill, D., Cremin, T., & Oliver, L. (2023). The impact of a changed writing environment on students' motivation to write. *Frontiers in Psychology*, *14*, Article 1212940. https://doi.org/10.3389/fpsyg.2023.1212940

Nasihah, M., & Cahyono, B. (2017). Language learning strategies, motivation, and writing achievement of Indonesian EFL students. *Arab World English Journal*, *8*(1), 250–263. https://doi.org/10.24093/awej/vol8no1.18

Nurjamin, A., Salazar-Espinoza, D. E., Saenko, N., & Bina, E. (2023). Learner-oriented assessment matters: Testing the effects of academic buoyancy, reflective thinking, and learner enjoyment in self-assessment and test-taking anxiety management of the EFL learners. *Language Testing in Asia*, *13*, Article 30. https://doi.org/10.1186/s40468-023-00247-z

Richards, J. C. (2022). Exploring emotions in language teaching. *RELC Journal*, *53*(1), 225–239. https://doi.org/10.1177/0033688220927531

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement: Motivation, learning, and well-being. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 171–195). Routledge. https://doi.org/10.4324/9780203879498

Sumakul, D. T. Y. G., Hamied, F. A., & Sukyadi, D. (2022a). Artificial intelligence in EFL classrooms: Friend or foe? *LEARN Journal: Language Education and Acquisition Research Network*, *15*(1), 232–256. https://so04.tci-thaijo.org/index.php/LEARN/article/view/256723

Sumakul, D. T. Y. G., Hamied, F. A., & Sukyadi, D. (2022b). Students' perceptions of the use of AI in a writing class. In *Proceedings of the 67th TEFLIN International Virtual Conference & the 9th ICOELT 2021* (pp. 52–57). Atlantis Press. https://doi.org/10.2991/assehr.k.220201.009

Sun, T., & Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System*, *90*, Article 102221. https://doi.org/10.1016/j.system.2020.102221

Tang, J., & Wu, Y. (2017). Impact of AWE rubrics and automated assessment on EFL writing instruction. *International Journal of Computer-Assisted Language Learning and Teaching*, *7*(2), 58–74. https://doi.org/10.4018/ijcallt.2017040104

Tian, L., & Zhou, Y. (2020). Learner engagement with automated feedback, peer feedback and teacher feedback in an online EFL writing context. *System*, *91*, Article 102247. https://doi.org/10.1016/j.system.2020.102247

Utami, S. P. T., Andayani, Winarni, R., & Sumarwati. (2023). Utilization of artificial intelligence technology in an academic writing class: How do Indonesian students perceive? *Contemporary Educational Technology*, *15*(4), Article ep450. https://doi.org/10.30935/cedtech/13419

White, R., & Arndt, V. (1991). *Process writing*. Longman.

Xu, X., & Wang, B. (2022). EFL students' academic buoyancy: does academic motivation and interest matter? *Frontiers in Psychology*, *13,* 1-5. https://doi.org/10.3389/fpsyg.2022.858054

Yun, S., Hiver, P., & Al-Hoorie, A. H. (2018). Academic buoyancy: Exploring learners' everyday resilience in the language classroom. *Studies in Second Language Acquisition*, *40*(4), 805–830. https://doi.org/10.1017/S0272263118000037

Zhai, N., & Ma, X. (2022). Automated writing evaluation (AWE) feedback: A systematic investigation of college students' acceptance. *Computer Assisted Language Learning*, *35*(9), 2817–2842. https://doi.org/10.1080/09588221.2021.1897019

Zumbrunn, S., Marrs, S., Broda, M., Ekholm, E., DeBusk-Lane, M., & Jackson, L. (2019). Toward a more complete understanding of writing enjoyment: A mixed methods study of elementary students. *AERA Open*, *5*(2). https://doi.org/10.1177/2332858419850792

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

Does AI Simplification of Authentic Blog Texts Improve Reading Comprehension, Inferencing, and Anxiety? A One-Shot Intervention in Turkish EFL Context

Ferdi Çelik, Ceylan Yangın Ersanlı, and Goshnag Arslanbay
Ondokuz Mayıs University, Samsun, Türkiye

Abstract

This experimental study investigates the impact of ChatGPT-simplified authentic texts on university students' reading comprehension, inferencing, and reading anxiety levels. A within-subjects design was employed, and 105 undergraduate English as a foreign language (EFL) students engaged in both original and ChatGPT-simplified text readings, serving as their own controls. The findings reveal a significant improvement in reading comprehension scores and inferencing scores following ChatGPT intervention. However, no significant change in reading anxiety levels was observed. Results suggest that ChatGPT simplification positively influences reading comprehension and inferencing, but its impact on reading anxiety remains inconclusive. This research contributes to literature on the use of artificial intelligence (AI) in education and sheds light on ChatGPT's potential to influence language learning experiences within higher education contexts. The study highlights the practical application of ChatGPT as a tool for helping students engage in authentic text readings by making text more comprehensible. Based on the findings, several multifaceted implications that extend to various stakeholders in the field of language education are provided.

Keywords: artificial intelligence, ChatGPT, simplification, reading, language teaching

Does AI Simplification of Authentic Blog Texts Improve Reading Comprehension, Inferencing, and Anxiety? A One-Shot Intervention in Turkish EFL Context

The field of higher education is increasingly recognizing technology's potential to elevate language learning experiences (Hong, 2023; Kalla et al., 2023; Kohnke et al., 2023). Artificial intelligence (AI), a cornerstone of modern life, offers significant opportunities for both teachers and students (Y. Chen et al., 2020; Holmes et al., 2023). AI, including AI in education (AIEd), has transformed pedagogy and human audiovisual literacy, influencing language learning practices (Yang et al., 2021). These advancements underscore a shift in educational paradigms, where AI emerges as a key facilitator in creating immersive and effective language learning environments. The integration of AIEd signifies a transformative era, enriching the educational landscape and shaping new possibilities for language learners and educators alike.

In line with this transformative trend, the dynamic nature of AI is evident in its continuous evolution and the introduction of diverse platforms, such as intelligent tutoring systems, teaching robots, and adaptive learning systems (Y. Chen et al., 2020; Yi et al., 2022). This advancement, exemplified by the merging of generative AI and large language models such as Chat Generative Pre-Trained Transformer (ChatGPT), contributes significantly to shaping language learning experiences (Lund & Wang, 2023; Mhlanga, 2023; Pavlik, 2023; Pogla, 2023). ChatGPT, acknowledged as a sophisticated language generation model, facilitates natural language discussions through machine-learning techniques (Brown et al., 2022; Dida et al., 2023; Susnjak & Maddigan, 2023).

Despite some exploration of ChatGPT's educational potential, its in-depth impact on language learning in higher education remains unexplored. This study addresses this research gap, contributing to the existing scholarly literature by investigating AI's broader capabilities and potential benefits. For this reason, the research questions specifically examine the impact of ChatGPT on university students' reading comprehension, inferencing performance, and reading anxiety levels when simplifying an authentic text on a life advice Website. In line with expectations, ChatGPT's role in fostering effective language practices within higher education is explored.

AIEd: Unraveling ChatGPT's Impact

Within the ever-changing realm of AIEd, the influence of advanced language models on language learning experiences is a central point of exploration. Ouyang and Jiao (2021) provide a comprehensive overview of AIEd development, emphasizing three paradigms where AI techniques address educational challenges, including learner agency, personalization, reflective learning, and a learner-centered, data-driven approach. Moreover, AI demonstrates proficiency across natural language tasks, from generating essays to translations and answering questions (Rospigliosi, 2023). The rapid growth of natural language processing technology recognizes large language models as a significant evolution. ChatGPT, a sophisticated generative language model, proves valuable in enhancing critical thinking, academic research, writing, and problem-solving skills (Dwivedi et al., 2023; Sullivan et al., 2023). Its excellence in generating original content and providing students with a comprehensive understanding and analysis of specific subjects (Kasneci et al., 2023; Tlili et al., 2023) underscores its central role in shaping language learning experiences.

Based on initial findings (Bin-Hady et al., 2023; Holmes et al., 2023), ChatGPT emerges as an adaptable simplification tool in language education. It not only contributes to personalized tutoring, automatic grading of writing, deep learning, and adaptive instruction (X. Chen et al., 2020; Chen & Hsu, 2022; Kim et al., 2019; Pang et al., 2021) but also acts as a scaffold for learning. It offers constructive feedback and functions as a collaborative partner in language practice.

As this study progressed, it aimed to explore the intricacies of reading dynamics involving comprehension, inferencing, and anxiety, assessing the unique advantages that ChatGPT offers and its potential contributions to second- and foreign-language learning.

Reading Dynamics

As highlighted by Hu and Nassaji (2014), reading plays a crucial role in vocabulary acquisition for both second-language (L2) and foreign-language (FL) learners, emphasizing the significance of language development. This importance becomes particularly evident during the inferencing process, where readers employ diverse strategies and background knowledge (Hu & Nassaji, 2014). Subsequently, the development of reading skills initiates with decoding and word fluency, evolving over time to encompass the ability to make inferences (Bayat & Çetinkaya, 2020). These inferences, in turn, function as facilitators to ensure a thorough grasp of the text. Moreover, Kispal (2008) identifies inferencing ability as one of the core comprehension skills in the context of reading: it empowers readers to establish meaningful connections between explicit information in the text and implicit ideas. In other studies (Haastrup, 2008; Wesche & Paribakht, 2009), the process of inferencing is delineated as guessing the meaning of an unfamiliar word or, alternatively, as "reading between the lines." At this crucial intersection, ChatGPT significantly contributes to this process. Operating as an adaptive simplification tool, ChatGPT can be applied judiciously to enhance comprehension. Its effectiveness is evident in its ability to break down complex sentences, use simpler vocabulary, and provide additional contextual information (Pogla, 2023). By supporting readers in making inferences and capturing the main ideas of a text, these approaches aim to alleviate potential reading anxiety, providing a more accessible pathway to comprehension.

Furthermore, to understand the interplay between ChatGPT's impact on language learning and reading dynamics, the associations revealed by the Corpus of Contemporary American English (COCA, n.d.) were explored. The term *reading* most frequently collocates with *comprehension* (2,910), *student* (1,852), and *skill* (1,404). Notably, there are no identified collocations involving *anxiety*, which could be attributed to a gap in the literature concerning reading anxiety. Saito et al. (1999) observed that despite reading's substantial role in the L2 curriculum, there has been relatively little discussion of anxiety in second-language reading. This gap in the literature highlights the need for a closer examination of reading anxiety, particularly within the context of AI-enhanced language learning.

Factors in FL reading, such as negotiating unfamiliar scripts and encountering unfamiliar cultural material, may pose challenges and evoke anxiety among FL readers (Saito et al., 1999). Detecting reading anxiety is challenging, particularly for silent reading, as it does not necessitate immediate reactions, unlike oral communication (Chow et al., 2021). Navigating the complexities of determining whether authentic texts are comprehensible for students, we encounter diverse terminology associated with authenticity, including terms including genuine, authentic, real, natural, semi-authentic, simulated, and simulated-authentic (AbdulHussein, 2014). In this intricate landscape, Tomlinson (2004) provides insightful perspectives, defining an *authentic text* as one not created for language teaching purposes. Examples abound, ranging from newspaper articles, rock songs, novels, and radio interviews to traditional fairy stories. Additionally, Tomlinson emphasizes the integral role of authentic tasks, engaging learners in language use reflective of real-world applications beyond the language classroom.

The literature underscores a strong belief in ChatGPT's integration into language learning for addressing challenges and alleviating anxiety, thereby having a positive impact on learners and marking a transformative step forward in language education. In pursuit of a deeper understanding, the research questions guide the investigation, aligning with the broader context of enhancing language learning experiences through ChatGPT, an advanced AI model. To this end, the following are the research questions of this study:

1. Does using ChatGPT to simplify an authentic text on a life advice Website affect university students' reading comprehension compared to reading the original text without ChatGPT simplification?

2. Does using ChatGPT for text simplification influence university students' inferencing scores?

3. How does the use of ChatGPT for text simplification influence university students' reading anxiety levels?

Method

Research Design

As this study focuses on investigating the impact of ChatGPT simplification of authentic texts, a within-subjects design was appropriate (Keren, 2014). This type of experimental design allowed for the examination of within-participant changes by exposing each participant to both conditions: reading the original text and reading the ChatGPT-simplified text (Lottridge et al., 2011). The exposure to both conditions provided a basis for assessing the effectiveness of the intervention. Each participant served as their own control as their reading comprehension performance and anxiety levels were measured before and after the intervention. Moreover, a one-shot (single-session) intervention (DeBacker et al., 2018, p. 712) was used in this study as it can be equally effective as more extended interventions for academic achievement (Walton & Cohen, 2011) and stress response (Crum et al., 2013).

Study Context

The study took place at a public university in Türkiye. The university had a preparatory school where students were taught English as a foreign language (EFL) for academic purposes for one academic year. At the beginning of their university education, students took a proficiency test, which assessed their English language skills at the B1 (equivalent to pre-intermediate/intermediate) level according to the Common European Framework of Reference for Languages (CEFR). The students who failed the exam took preparatory classes to learn English for academic purposes. These classes aimed to get students to reach the B1 level by the end of the year. The students had 26 hours of English weekly, and the classes were divided into speaking and listening (5 hours/week) and reading and writing (5 hours/week). In the skills lessons, the Oxford Q Skills coursebook series was used. There was also a 16-hour main course where students were taught the basics of English, such as the use of English, including vocabulary and grammar. In the main course, the Oxford Headway series was used.

Participants

The participants were purposefully sampled from the pool of students enrolled in an undergraduate program at a university in Türkiye. Potential participants were identified on the criterion of enrollment in an EFL course. They were asked for consent to participate in the study; 112 students expressed informed consent. However, seven participants withdrew from the study due to unforeseen circumstances, which resulted in a final sample size of 105 participants (45 male, 60 female), ranging in age from 18 to 24.

Data Collection Tools

Demographics Survey

The survey collected data on participant demographics, such as age and gender. It also asked whether the participants used any tools for understanding challenging texts or if they read authentic texts, considering that these prior experiences might influence their performance in the study.

Reading Comprehension Test

The reading comprehension test (RCT) was developed by the researchers based on the original text used in the study. The test was reviewed by three English-language teachers, who reached consensus that it would effectively measure students' comprehension. The test consisted of 10 multiple-choice questions with four options each. Each question was worth 10 points, and a perfect score was 100. The participants took the test immediately after both interventions (Table 1). A specific question (inferencing item) was created to investigate the effect of using AI simplification on learners' inferencing scores.

Foreign Language Reading Anxiety Scale

The foreign language reading anxiety scale (FLRAS), developed by Saito et al. (1999), is a 20-item instrument structured as a 5-point Likert-type scale, ranging from strongly disagree to strongly agree. It was used to measure anxiety levels experienced by participants while engaging with foreign-language reading materials (Mikami, 2019). For the present study, the scale had a satisfactory internal validity (Cronbach's alpha) for both the first intervention (.827) and second intervention (.849). The test was used after both interventions (Table 1).

Table 1

Data Collection Tool Delivery Times

Instrument	Pre-intervention	First intervention	Second intervention
Demographics survey	X		
RCT		X	X
FLRAS		X	X

Note. RCT = reading comprehension test; FLRAS = foreign language reading anxiety scale.

Procedure

To select an authentic text, the researchers focused on the theme "life" as it was a general topic that was thought to be interesting for the participants. A ranked list of 100 best life blogs created by FeedSpot (2024), based on criteria such as Website traffic, social media followers, and timeliness of content, was used. Next, Google's true random number generator was used to select a blog, and a life advice blog was selected. The instructor who would deliver the reading class was asked to choose a blog post from the Website, considering students' interests. Finally, an authentic text was acquired to be used in the study. The instructor created a WhatsApp group before the intervention to easily deliver the link for the authentic text.

Control Intervention

The link for the authentic text was shared with the participants. Participants were asked to read the authentic text from the life advice Website carefully. They had 20 minutes to read the text. Following the reading, participants completed the FLRAS and the RCT.

One-Shot Experimental Intervention

The instructor sent the whole text as a WhatsApp message to the participants. ChatGPT 3.5 was used to generate a simplified version of the authentic text. The participants were asked to go to ChatGPT. Next, they were asked to use the following prompt: "Make this text comprehensible for an a2 level learner. *Put the text here*." The participants then read the simplified version generated by ChatGPT. They had 20 minutes to read it. Next, they were asked to read the authentic text again, using the link to the original Website. After reading the simplified version and revisiting the original text in the blog, participants again completed the FLRAS and the RCT.

Data Analysis

The data collected for this study underwent analysis to address the research questions using SPSS v.26 for Windows. Following the within-subjects design, normality tests (Shapiro–Wilk) were employed to assess data distribution. Non-normally distributed data were then analyzed using the Wilcoxon signed-rank test.

Findings

Demographics

To provide insight into how learners perceived the difficulty of the text used in this study, all participants were asked to rate the difficulty of the text on a scale ranging from 0 (not difficult) to 10 (extremely difficult). On average, the participants perceived the text to be moderately difficult (M = 6.06); there was some variability in individual perceptions (SD = 3.159).

In response to the question "Do you read English texts from original resources such as news sites, blogs, magazines, etc. (e.g., PubMed, BBC News)?" 40 participants (38.1%) responded "Yes" and 65 participants (61.9%) chose "No."

Participants were also surveyed regarding the technologies they used to enhance their understanding of original texts. The majority of respondents (63.8%) reported using Google Translate for this purpose. Additionally, 26.7% employed online dictionaries, while 25.7% relied on mobile phone dictionary apps. Notably, 14.3% indicated that they did not employ any additional technologies to improve their comprehension of original texts. Participants were also given the opportunity to specify other technologies they might use; however, no such information was provided. These findings offer valuable insights into the diverse tools participants use to augment their reading comprehension.

Reading Comprehension

To address the first research question regarding the impact of using ChatGPT to simplify an authentic text on university students' reading comprehension, we performed a series of statistical tests. Initially, Shapiro–Wilk normality tests were conducted, which revealed non-normal distributions for both pretest and posttest results (pretest: W = .953, p = .001; posttest: W = .964, p = .006). Consequently, we turned to the Wilcoxon signed-rank test to analyze the data further (Table 2).

Table 2

Wilcoxon Signed-Rank Test for the Reading Comprehension Test

Posttest–pretest	N	M rank	Sum of ranks	Z	p
Negative ranks	2	21.50	43	−8.142	< .001*
Positive ranks	87	45.54	3,962		
Ties	16				
Total	105				

* p < .05.

A Wilcoxon signed-rank test was conducted to assess the differences between pretest and posttest scores (Table 2). Descriptive statistics revealed a pretest mean of 42.67 (SD = 26.101) and a posttest mean of 60.00 (SD = 24.690). The Wilcoxon test indicated a significant difference between posttest and pretest scores, with a Z value of −8.142 (p < .001, two-tailed). The negative Z value suggests a statistically significant improvement in posttest scores, which implies a positive impact of the intervention.

Inferencing

Concerning the second research question, concerning the influence of using ChatGPT for text simplification for university students' inferencing scores, we carried out similar analyses as described above. First, we applied normality tests to the inferencing item pretest and posttest results. The results indicated non-normally distributed scores for both conditions, with a Shapiro–Wilk statistic of .599 (p = .000) for the pre-condition and .629 (p = .000) for the post-condition. Subsequently, a Wilcoxon signed-rank test was employed to evaluate differences between pre- and posttest scores (Table 3).

Table 3

Wilcoxon Signed-Rank Test for Inferencing Item

Posttest–pretest	N	M rank	Sum of ranks	Z	p
Negative ranks	15	27.50	412.50	−3.266	< .001*
Positive ranks	39	27.50	1,072.50		
Ties	51				
Total	105				

* p < .05.

Descriptive statistics revealed a pre-intervention mean of 3.43 (SD = 4.769) and a post-intervention mean of 5.71 (SD = 4.972) (Table 3). The Wilcoxon test indicated a significant difference between post- and pretest scores (Z = −3.266, p = .001, two-tailed). The negative Z value implies a statistically significant improvement in posttest scores that suggests a positive impact of the intervention on participants' inferencing abilities.

Reading Anxiety

To address the third research question, related to the impact of ChatGPT-simplified text on university students' reading anxiety levels, we initially examined the normality of the data with Shapiro–Wilk statistics (pretest = .989, posttest = .973). The test results suggested that neither the pretest nor the posttest results followed a normal distribution. Therefore, we opted for the Wilcoxon signed-rank test to evaluate the differences (Table 4).

Table 4

Wilcoxon Signed-Rank Test for Reading Anxiety

Posttest–pretest	*N*	*M* rank	Sum of ranks	*Z*	*p*
Negative ranks	58	53.79	3,120	−1.265	.206*
Positive ranks	46	50.87	2,340		
Ties	1				
Total	105				

* $p > .05$.

Results of the Wilcoxon signed-rank test did not reveal a statistically significant difference in reading anxiety levels between pretest and posttest conditions ($Z = -1.265$, $p = .206$) (Table 4). This suggests no significant change in reading anxiety levels following the intervention.

Discussion

It is evident that AI is having a significant impact on modern life, especially in the field of education. OpenAI and ChatGPT are remarkable examples of AI technology that can bring about revolutionary advancements in the field of education. This accessibility of these resources encourages the development of AI-powered solutions that are customized to meet various educational needs, which can improve learning by making it more individualized, efficient, and easily accessible (Hwang et al., 2020). ChatGPT promotes learner autonomy and makes language learning easier by answering questions intelligently and without requiring users to wait for assistance (Taecharungroj, 2023). Furthermore, AI-powered chatbots offer language assistance and encourage frequent conversation practice. As evidenced by research findings, these chatbots have proven effective in promoting language learners' overall language development as well as serving as companions for learners to engage in conversational practice (Jeon et al., 2023).

However, ChatGPT must be understood by researchers, educators, and students in a way that sets it apart from classical AI, chatbots, and information systems. First, it is more than just an intelligent system providing learning content, individualized support, or direction. Second, it performs better than a chatbot that can hold students' attention through natural language communication. Finally, it goes beyond a writing assistant.

To this end, the current study aimed to investigate the effectiveness of using ChatGPT as a learning assistant/scaffolder by focusing on whether the use of ChatGPT to simplify authentic texts may affect university students' reading comprehension, inference skills, and anxiety levels when reading authentic texts. A text may undergo simplification in regard to certain grammatical elements, cultural references, and word choice. Essentially, a simplified text has been adjusted from its original version or crafted explicitly for L2 learners. Learners may require simplification to align with the teaching and learning objectives and for better comprehension. The findings gathered from this study indicated a statistically significant difference between students' authentic text comprehension scores and ChatGPT-simplified text comprehension scores. This finding is contrary to those in the existing literature. Soma et al. (2015) investigated the effect of authentic and simplified texts on reading literacy and vocabulary mastery. Their findings suggest that neither text type was superior to the other in terms of comprehensibility for both high- and low-level achievers. Another study (Gashti, 2018) focused on the effect of authentic and simplified literary texts on the reading comprehension of EFL learners. The results indicated that incorporating both simplified and authentic literary texts had a beneficial impact on the reading comprehension abilities of EFL learners. The investigation also revealed no discernible difference between simplified literary texts and the actual literary materials. Stated differently, language learners' reading comprehension was significantly improved by both simplified and authentic literary resources. Yet, there appeared to be a slight difference, with simplified texts being favored, which is in line with our study findings. In a similar vein, Crossley et al. (2007) stated that although no meaningful difference was found between authentic and simplified texts in reading comprehension, simplified texts were favored because they provided students with more common words and less syntactic complexity.

This study's findings also suggest a notable enhancement in posttest scores, indicating a statistically significant improvement in participants' inferencing abilities as a result of the intervention. Inference making is often acknowledged as a vital component of skilled reading (Cain et al., 2001; Graesser et al., 1994; Laufer, 2020; Oakhill & Cain, 2007). It might be difficult for EFL students to understand authentic texts, especially when they include complicated information. Making inferences from authentic texts is a crucial component of reading comprehension. These inferences are usually automatically done as part of the interpretation process, in which readers use what they already know about the text to figure out the meaning. To sustain understanding, students must constantly create new information or rely on what they already know to fill in the details provided by the text. Therefore, in our current study, reading the ChatGPT-simplified texts before reading the authentic texts may have given the students a general idea about the text content, so that they had the necessary background knowledge, which may have helped them infer the meanings of unknown words in the text.

Another finding of the present study is that there was no statistically significant difference in students' reading anxiety while reading authentic text or while reading ChatGPT-simplified text. Anxiety is recognized as a significant factor affecting students' reading comprehension. It can stem from emotional and physical stress. According to the literature, higher levels of reading anxiety may lead to misinterpretations and negative feelings (Dewi & Pramerta, 2021; Jalongo & Hirsh, 2010; Saito et al., 1999). However, it should also be noted that anxiety may have a facilitating effect on students' reading performance. Meymeh et al.'s (2010) study revealed the facilitating effect of anxiety on university students' reading performance of texts that were lexically and grammatically simplified. In the current study, no statistically significant difference was found between the students' reading anxiety levels while reading the authentic texts and then rereading the authentic texts after reading the ChatGPT-simplified versions.

Limitations

This study had several limitations. It was conducted at a single university in Türkiye, which limited its representativeness. The focus on undergraduate students in an EFL course may have restricted generalizability to a more diverse learner population. The exclusion of higher-proficiency students and reliance on a single text may also impact its applicability. The short 20-minute intervention time may not capture long-term impacts. Reading anxiety is a complex psychological construct influenced by various individual and situational factors. The study might not have captured the full complexity of reading anxiety or adequately assessed its long-term effects, as changes in psychological domains typically require sustained interventions over time. The lack of long-term follow-up is also a notable limitation. Future studies addressing these issues would contribute to the literature. Addressing these limitations through larger sample sizes, longer intervention periods, and diverse participant demographics can enhance the applicability of future studies in this area.

Conclusion and Suggestions

The study highlighted the positive impact of using ChatGPT to simplify authentic texts on university students' reading skills. Specifically, the findings revealed improvement in both reading comprehension and reading inferencing abilities among participants who engaged with ChatGPT-simplified texts. Notably, the observed higher posttest scores shed light on the effectiveness of ChatGPT as an educational tool, which suggests that it has the potential to enhance language learning experiences.

The study did not identify any notable change in overall students' reading anxiety levels when participants used ChatGPT-simplified texts. This finding indicates that while ChatGPT may contribute positively to L2 reading, it does not have a substantial effect on reducing anxiety associated with engaging with original, more complex texts in single-shot interventions. The reasons behind this may be that the students may have gathered the main idea from ChatGPT's simplified version of the text with simpler words and less complicated grammatical structures and that it may have had a slight effect on their anxiety levels arising from the unknown when rereading the authentic text. Future studies with long-term interventions are needed.

The study's results present a notable departure from some existing literature that posits no significant difference between authentic and simplified texts. This highlights the need for further exploration into the multifaceted factors influencing language learning outcomes and the impact of AI-driven tools in educational settings. Researchers are encouraged to delve deeper into the relationship between AI-assisted tools like ChatGPT and changes in students' reading anxiety and also to explore other underlying psychological mechanisms.

The implications of this research extend to the integration of educational technology as the positive impact of ChatGPT suggests its potential incorporation into language learning platforms. English-language teachers may face reading materials that are above learners' proficiency levels, and AI simplification of these texts would be helpful for learners' comprehension. Educators may need to adapt their pedagogical approaches using AI technologies while maintaining a balanced instructional strategy. Furthermore, developers of language learning materials could explore the potential to create AI-simplified versions to accompany authentic texts to cater to diverse learners with varying proficiency levels and contribute to a more adaptive learning environment. The positive implications of ChatGPT in language education suggest a broad trend toward the continuous integration of technology in the learning process. Educational institutions should consider investing in and adopting cutting-edge technologies to stay abreast of advancements, promoting innovation in teaching methodologies. Overall, this study contributes valuable insights that inform the ongoing dialogue on the role of AI in language education.

References

AbdulHussein, F. R. (2014). Investigating EFL college teachers' and learners' attitudes toward using authentic reading materials in Misan. *Procedia—Social and Behavioral Sciences*, *136*, 330–343. https://doi.org/10.1016/j.sbspro.2014.05.338

Bayat, N., & Çetinkaya, G. (2020). The relationship between inference skills and reading comprehension. *Education and Science*, *45*(203), 177–190. https://doi.org/10.15390/EB.2020.8782

Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M. (2023). Exploring the dimensions of ChatGPT in English language learning: A global perspective. *Library Hi Tech*. Advance online publication. https://doi.org/10.1108/LHT-05-2023-0200

Brown, H., Lee, K., Mireshghallah, F., Shokri, R., & Tramèr, F. (2022, June 21–24). What does it mean for a language model to preserve privacy? In *FaccT '22: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 2280–2292). Association for Computing Machinery. https://doi.org/10.1145/3531146.3534642

Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference-making ability, and their relation to knowledge. *Memory & Cognition*, *29*(6), 850–859. https://doi.org/10.3758/BF03196414

Chen, H.-R., & Hsu, W.-C. (2022). Do flipped learning and adaptive instruction improve student learning outcome? A case study of a computer programming course in Taiwan. *Frontiers in Psychology*, *12*, Article 768183. https://doi.org/10.3389/fpsyg.2021.768183

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *1*, Article 100002. https://doi.org/10.1016/j.caeai.2020.100002

Chen, Y., Chen, Y., & Heffernan, N. (2020). *Personalized math tutoring with a conversational agent* (arXiv preprint: arXiv:2012.12121).

Chow, B. W.-Y., Mo, J., & Dong, Y. (2021). Roles of reading anxiety and working memory in reading comprehension in English as a second language. *Learning and Individual Differences*, *92*, Article 102092. https://doi.org/10.1016/j.lindif.2021.102092.

Corpus of Contemporary American English (COCA). (n.d.). Retrieved January 27, 2024, from https://www.english-corpora.org/coca/

Crossley, S. A., McCarthy, P. M., Louwerse, M. M., & McNamara, D. S. (2007). A linguistic analysis of simplified and authentic texts. *The Modern Language Journal*, *91*, 15–30. https://doi.org/10.1111/j.1540-4781.2007.00507.x

Crum, A. J., Salovey, P., & Achor, S. (2013). Rethinking stress: The role of mindsets in determining the stress response. *Journal of Personality and Social Psychology*, *104*(4), 716–733. https://doi.org/10.1037/a0031201DeBacker, T. K., Heddy, B. C., Kershen, J. L., Crowson, H. M., Looney, K., & Goldman, J. A. (2018). Effects of a one-shot growth mindset intervention on beliefs about intelligence and achievement goals. *Educational Psychology*, *38*(6), 711–733. https://doi.org/10.1080/01443410.2018.1426833

Dewi, P. H. R., & Pramerta, G. P. A. (2021). Correlation between anxiety and reading comprehension: A study in a secondary school. *Academic Journal on English Studies, 1*(2), 152–161. https://e-journal.unmas.ac.id/index.php/ajoes/article/view/4606

Dida, H. A., Chakravarthy, D. S. K., & Rabbi, F. (2023). ChatGPT and big data: Enhancing text-to-speech conversion. *Mesopotamian Journal of Big Data*, *2023*, 33–37. https://doi.org/10.58496/MJBD/2023/005

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A.M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M.A., Al-Busaidi, A.S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., Carter, L., & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

FeedSpot. (2024, May 20). *100 best life blogs and websites in 2024*. Lifestyle Bloggers Database. https://lifestyle.feedspot.com/life_blogs/

Gashti, Y. B. (2018). The effect of authentic and simplified literary texts on the reading comprehension of Iranian advanced EFL learners. *Iranian Journal of English for Academic Purposes*, *7*(2), 32–44. https://e-journal.unmas.ac.id/index.php/ajoes/article/view/4606

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*(3), 371–395. https://doi.org/10.1037/0033-295X.101.3.371

Haastrup, K. (2008). Lexical inferencing procedures in two languages. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Vocabulary and writing in a first and second language: Processes and development* (pp. 67–111). Palgrave Macmillan. https://doi.org/10.1057/9780230593404_3

Holmes, W., Bialik, M., & Fadel, C. (2023). Artificial intelligence in education. In C. Stückelberger & P. Duggal (Eds.), *Data ethics: Building trust: How digital technologies can serve humanity* (pp. 621–653). Globethics Publications. https://doi.org/10.58863/20.500.12424/4276068

Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, *5*(1), 37–45. https://jeti.thewsu.org/index.php/cieti/article/view/103

Hu, H. C. M., & Nassaji, H. (2014). Lexical inferencing strategies: The case of successful versus less successful inferencers. *System*, *45*, 27–38. https://doi.org/10.1016/j.system.2014.04.004

Hwang, G. J., Xie, H., Wah, B. W., Gasevic, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *1*, Article 100001. https://doi.org/10.1016/j.caeai.2020.100001

Jalongo, M. R., & Hirsh, R. A. (2010). Understanding reading anxiety: New insights from neuroscience. *Early Childhood Education Journal*, *37*(6), 431–435. https://doi.org/10.1007/s10643-010-0381-5

Jeon, J., Lee, S., & Choe, H. (2023). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education*, *206*, Article 104898. https://doi.org/10.1016/j.compedu.2023.104898

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Keren, G. (2014). Between- or within-subjects design: A methodological dilemma. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Vol. 1. Methodological issues* (pp. 257–272). Psychology Press. https://doi.org/10.4324/9781315799582-10

Kim, S., Park, J., & Lee, H. (2019). Automated essay scoring using a deep learning model. *Journal of Educational Technology Development and Exchange*, *2*(1), 1–17.

Kispal, A. (2008). *Effective teaching of inference skills for reading: Literature review* (Research report DCSFRR031). National Foundation for Educational Research. https://eric.ed.gov/?id=ED501868

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Laufer, B. (2020). Lexical coverages, inference unknown words and reading comprehension: How are they related? *TESOL Quarterly*, *54*(4), 1076–1085. https://doi.org/10.1002/tesq.3004

Lottridge, S. M., Nicewander, W. A., & Mitzel, H. C. (2011). A comparison of paper and online tests using a within-subjects design and propensity score matching study. *Multivariate Behavioral Research*, *46*(3), 544–566. https://doi.org/10.1080/00273171.2011.569408

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News*, *40*(3), 26–29. https://doi.org/10.1108/LHTN-01-2023-0009

Meymeh, M. H., Rashtchi, M., & Mohseni, A. (2019). Anxiety effect: A case of text modification and the effect of high and low anxiety levels on medical students' comprehension performance. *Journal of Paramedical Sciences*, *10*(1), 20–26. https://doi.org/10.22037/jps.v10i1.24111

Mhlanga, D. (2023, February 11). *Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning*. SSRN. https://doi.org/10.2139/ssrn.4354422

Mikami, H. (2019). Reading anxiety scales: Do they measure the same construct? *Reading in a Foreign Language*, *31*(2), 249–268. https://eric.ed.gov/?id=EJ1232213

Oakhill, J., & Cain, K. (2007). Issues of causality in children's reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 47–72). Lawrence Erlbaum Associates Publishers.

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, *2*, Article 100020. https://doi.org/10.1016/j.caeai.2021.100020

Pang, G., Shen, C., Cao, L., Hengel, A. V. D. (2021). Deep learning for anomaly detection: A review. *ACM Computing Surveys*, *54*(2), 1–38. https://doi.org/10.1145/3439950

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, *78*(1), 84–93. https://doi.org/10.1177/10776958221149577

Pogla, M. (2023, April 24). *ChatGPT: Optimizing language models for dialogue*. AutoGPT. https://autogpt.net/chatgpt-optimizing-language-models-for-dialogue/

Rospigliosi, P. A. (2023). Artificial intelligence in teaching and learning: What questions should we ask of ChatGPT? *Interactive Learning Environments*, *31*(1), 1–3. https://doi.org/10.1080/10494820.2023.2180191

Saito, Y., Garza, T. J., & Horwitz, E. K. (1999). Foreign language reading anxiety. *Modern Language Journal*, *83*(2), 202–218. https://doi.org/10.1111/0026-7902.00016

Soma, R., Mukminin, A., & Noprival, A. M. (2015). Toward a better preparation of student teachers' reading skill: The SQ3R strategy with authentic and simplified texts on reading literacy and vocabulary mastery. *Journal of Education and Learning*, *9*(2), 125–134. https://doi.org/10.11591/edulearn.v9i2.1527

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, *6*(1), 31–40. https://doi.org/10.37074/jalt.2023.6.1.17

Susnjak, T., & Maddigan, P. (2023). Forecasting patient flows with pandemic induced concept drift using explainable machine learning. *EPJ Data Science*, *12*, Article 11. https://doi.org/10.1140/epjds/s13688-023-00387-5

Taecharungroj, V. (2023). What can ChatGPT do? Analyzing early reactions to the innovative AI chatbot on Twitter. *Big Data and Cognitive Computing*, *7*(1), Article 35. https://doi.org/10.3390/bdcc7010035

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*, Article 15. https://doi.org/10.1186/s40561-023-00237-x

Tomlinson, B. (2004). *Materials development in language teaching* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781139042789

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*, *331*(6023), 1447–1451. https://doi.org/10.1126/science.1198364

Wesche, P. M. B., & Paribakht, P. T. S. (2009). *Lexical inferencing in a first and second language: Cross-linguistic dimensions*. Multilingual Matters. https://doi.org/10.21832/9781847692245

Yang, S. J., Ogata, H., Matsui, T., & Chen, N. S. (2021). Human-centered artificial intelligence in education: Seeing the invisible through the visible. *Computers and Education: Artificial Intelligence*, *2*, Article 100008. https://doi.org/10.1016/j.caeai.2021.100008

Yi, Y., Cho, S., & Jang, J. (2022). Methodological innovations in examining digital literacies in applied linguistics research. *TESOL Quarterly*, *56*(3), 1052–1062. https://doi.org/10.1002/tesq.3140

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

AI and the Future of Teaching: Preservice Teachers' Reflections on the Use of Artificial Intelligence in Open and Distributed Learning

Fatih Karataş[1] and Erkan Yüce[2]
[1]Nevsehir Haci Bektas Veli University, Nevsehir, Türkiye; [2]Aksaray University, Aksaray, Türkiye

Abstract

The rapid advancement of artificial intelligence (AI) in education underscores transformative prospects for open and distributed learning, encompassing distance, hybrid, and blended learning environments. This qualitative study, grounded in narrative inquiry, investigates the experiences and perceptions of 141 preservice teachers engaged with AI, mainly through ChatGPT, over a 3-week implementation on Zoom to understand its influence on their evolving professional identities and instructional methodologies. Employing Strauss and Corbin's methodological approach of open, axial, and selective coding to analyze reflective narratives, the study unveils significant themes that underscore the dual nature of AI in education. Key findings reveal ChatGPT's role in enhancing educational effectiveness and accessibility while raising ethical concerns regarding academic integrity and balanced usage. Specifically, ChatGPT was found to empower personalized learning and streamline procedures, yet challenges involving information accuracy and data security remained. The study significantly contributes to teacher education discourse by revealing AI's complex educational impacts, highlighting an urgent need for comprehensive ethical AI literacy in teacher training curricula. However, critical ethical considerations and practical challenges involving academic integrity, information accuracy, and balanced AI use are also brought to light. The research also spotlights the need for responsible AI implementation in open and distributed learning to optimize educational outcomes while addressing potential risks. The study's insights advocate for future-focused AI literacy frameworks that integrate technological adeptness with ethical considerations, preparing teacher candidates for an intelligent digital educational landscape.

Keywords: artificial intelligence (AI), ChatGPT, AI in education, AIED, AI in teacher education, narrative inquiry

AI and the Future of Teaching: Preservice Teachers' Reflections on the Use of Artificial Intelligence in Open and Distributed Learning

The integration of artificial intelligence (AI) in teacher education marks a significant paradigm shift, fundamentally altering pedagogical methods and the role of educators (Cavalcanti et al., 2021; ElSayary, 2024). AI technologies including deep learning generative AI, intelligent tutoring systems, and automated grading are transforming teaching and learning processes, emphasizing the need for teachers to possess a thorough understanding of these technologies (Celik, 2023; Edwards et al., 2018). However, the rapid progress of AI in education presents challenges, such as the gap in teachers' AI knowledge and capabilities, and the redefinition of their roles in an increasingly AI-supported teaching and learning environments (Guilherme, 2019; Nazaretsky et al., 2022). This evolving landscape necessitates a comprehensive approach in teacher education, focusing on AI literacy and the ethical, practical, and pedagogical implications of AI integration (Tan & Lim, 2018). As AI continues to reshape education, it is crucial to align these technological advancements with educational goals to enhance learning while preserving the essential human elements of teaching.

Literature Review

The integration of digital technologies into education is significantly influenced by teachers' perspectives on digital learning, which are shaped by their professional experiences and external factors (Liu & Wang, 2024). Recent challenges, such as the increase in teacher disengagement during online lessons and questions regarding the efficacy of digital instruction brought on by the pandemic, have highlighted the intricate nature of digital education (Wang, 2023; Wang et al., 2023). Furthermore, the adoption of educational technology is heavily reliant on teachers' emotional intelligence and their beliefs in their abilities, emphasizing the critical role of psychological aspects in the successful implementation of digital tools in teaching practices (Zhi et al., 2023). This underscores the essential need for fostering digital competencies among teachers, encompassing both technical skills and the psychological readiness to navigate and use digital learning environments effectively.

Digital teacher competencies increasingly emphasize the integration of AI in education, marking a transformative era in teacher education through tools like ChatGPT, which offer personalized learning experiences and enhance student engagement (Addo & Sentance, 2023; Cavalcanti et al., 2021). The COVID-19 pandemic has accentuated AI's role in facilitating online learning, with significant implications for educational and psychological well-being (Jiang et al., 2022; Vadivel et al., 2023). The pandemic has further underscored AI's importance in education, particularly in data mining and information retrieval for enriching learning experiences, as evidenced by Araka et al. (2022) and Cheng et al. (2022), and in its impact on curriculum development (Hsu et al., 2022), illustrating the technology's potential to personalize education and predict student performance, which are crucial for reducing dropout rates (Hwang et al., 2022; Rodriguez et al., 2022; Tzeng et al., 2022).

However, the adoption of AI in teaching faces challenges, including technical, ethical, and legal issues, highlighting the need for updated pedagogical frameworks that incorporate AI literacy and ethical considerations (Celik, 2023; García-Peñalvo, 2023). The literature suggests that ensuring ethical use, transparency, and fairness in the integration of AI technologies is vital for equitable learning outcomes and maintaining educational integrity (ElSayary, 2024; Hashem et al., 2024; Keeley, 2023). Addressing these challenges requires educators to be equipped with the skills to navigate AI technologies effectively;

comprehensive professional development will be crucial for acquiring AI competencies (Kim et al., 2022; Lawrence et al., 2024; Ng et al., 2023).

Studies on teachers' use of AI in education highlight its benefits in lesson customization and material discovery, alongside challenges in practical problem-solving (Keeley, 2023). Concerns about bias, accuracy, and the lack of human interaction in AI teaching tools underscore the need for further exploration (ElSayary, 2024). The literature calls for comprehensive ethical AI literacy in teacher training curricula to prepare teachers for a digitally intelligent educational landscape, advocating for a balanced, ethical approach to AI integration in teacher education (Addo & Sentance, 2023; Gentile et al., 2023). However, realizing AI technologies' full potential is contingent on adequately equipping educators with the necessary skills to effectively navigate and use them. Despite the progress in adopting AI in education, a significant gap remains in understanding how future educators perceive and engage with AI applications like ChatGPT. This gap is crucial, as it influences their readiness and enthusiasm to use AI in fostering accessible, flexible, and personalized learning experiences across diverse educational settings. In addressing this gap, this research aims to provide a comprehensive understanding of preservice teachers' encounters with AI tools, shedding light on both the perceived benefits and challenges. Addressing this critical gap, our research probes the following research question:

> RQ: How do preservice teachers perceive the integration of AI applications in teaching, and how does this integration shape their teaching styles and professional identities?

The insights gained are anticipated to inform the design of teacher education curricula and professional development programs, underscoring the need for augmented support and training in proficiently using AI applications in open and distributed learning contexts.

Methodology

Research Design

This study adopted a narrative inquiry approach, as defined by Kutsyuruba and Stasel (2023) and further elaborated by Chase (2005), to explore the perceptions and experiences of preservice teachers regarding AI in educational settings. Narrative inquiry, deeply rooted in understanding the social and cultural dimensions of individual and community narratives (Adama et al., 2016), is particularly apt for examining the evolving professional identities of preservice teachers as they interact with AI technologies such as ChatGPT. This methodology allows for a profound exploration of personal and shared experiences over time, emphasizing the significance of these narratives in shaping identities and practices within educational contexts. By centering on participants' perspectives, narrative inquiry offers a nuanced understanding of preservice teachers' subjective realities and experiences, thereby aligning with the study's objective and research question. Acknowledging the inherent subjectivity and the potential for bias in narrative inquiry, this study incorporates member checking (Bower et al., 2021) to enhance methodological rigor and credibility.

Participants

The selection of 141 preservice teachers from a state university was based on purposeful sampling, aimed at capturing a broad spectrum of experiences with ChatGPT across various departments and levels of AI familiarity (Cohen et al., 2017; Özdil & Kunt, 2023). This diverse group reflects a thoughtfully structured selection process, ensuring a rich and diverse range of perspectives by establishing an ideal relationship between researchers and participants, thereby contributing to a comprehensive understanding of how future educators perceive and interact with AI technologies such as ChatGPT. Accordingly, Table 1 presents the department, age, gender, and prior AI experience of the participants. Since all participants were higher education students from the education faculty of a university in Türkiye, it was considered unnecessary to include citizenship information in the table, as any differences arising from citizenship would not be significant.

Table 1

Participants' Demographic Data

Department[a]	Age	Gender	Prior AI experience
Preschool education (*n* = 53)	17–18 (*n* = 38)	Female (*n* = 23)	Yes (*n* = 9)
			No (*n* = 14)
		Male (*n* = 15)	Yes (*n* = 4)
			No (*n* = 11)
	19–21 (*n* = 15)	Female (*n* = 9)	Yes (*n* = 3)
			No (*n* = 6)
		Male (*n* = 6)	Yes (*n* = 2)
			No (*n* = 4)
Elementary science education (ESE) (*n* = 49)	17–18 (*n* = 34)	Female (*n* = 21)	Yes (*n* = 10)
			No (*n* = 11)
		Male (*n* = 13)	Yes (*n* = 4)
			No (*n* = 9)
	19–21 (*n* = 15)	Female (*n* = 8)	Yes (*n* = 3)
			No (*n* = 5)
		Male (*n* = 7)	Yes (*n* = 2)
			No (*n* = 5)
Physical education and sports teacher department (*n* = 39)	17–18 (*n* = 27)	Female (*n* = 17)	Yes (*n* = 6)
			No (*n* = 11)
		Male (*n* = 10)	Yes (*n* = 3)
			No (*n* = 7)
	19–21 (*n* = 12)	Female (*n* = 7)	Yes (*n* = 2)
			No (*n* = 5)
		Male (*n* = 5)	Yes (*n* = 0)
			No (*n* = 5)

Note. [a] Department indicates the programs the students were studying.

Data Collection Tool and Procedure

Using reflection papers as the data collection tool, this research adopts narrative inquiry principles to collect and analyze reflective narratives from preservice teachers (Kaminski, 2003; Kayima, 2021). At the end of the 3-week implementation using the Zoom online learning platform, which focused on the use of AI in teaching practices, participants used reflective writing as a means to delve into and express their perceptions, experiences, and readiness to integrate AI technologies such as ChatGPT into their future educational practices. Guiding questions based on narrative inquiry principles (Clandinin, 2006; Xu & Connelly, 2010) were provided to aid in participants' expression of perceptions and experiences, encouraging them to explore beyond these questions to fully articulate their insights and learning outcomes. These guiding questions, outlined in Table 2, were instrumental in prompting students to describe their experiences and insights gained throughout the study.

Table 2

Prompts to Guide Reflections

Category	Reflections prompt
Exploring early perceptions and practical encounters with AI	Think about your initial thoughts on AI such as ChatGPT in education and compare them with your actual experiences during the course. Were there surprises or did things match up? (Carvalho et al., 2022; Gentile et al., 2023) Describe how using AI tools in specific tasks changed or confirmed your original beliefs about them. (Cavalcanti et al., 2021; ElSayary, 2024)
Evaluating AI's influence on pedagogical techniques and educator identity	Discuss how AI can support or enhance traditional teaching methods, using real or imagined examples where AI improves education. (Guilherme, 2019; Kim et al., 2022) Reflect on how using AI in teaching affects your view of being an educator. Do these experiences align with or challenge your initial idea of a teacher's role? (Kim & Kwon, 2023)
Preparedness for incorporating AI and ethical reflections	Assess your readiness and willingness to use AI in your teaching. Are you concerned about any specific issues? (Celik, 2023; Ng et al., 2021) Talk about the ethical challenges or dilemmas you foresee or have faced with AI in education and suggest possible solutions. (Edwards et al., 2018; Tan & Lim, 2018)
Impacts on learner engagement and educator's development	How do you think AI tools affect student engagement and learning? Share your observations or thoughts on students' reactions to AI. (Kim et al., 2022) Reflect on how your understanding of AI in education has evolved. Mention key lessons and how they will shape your future as an educator. (Celik, 2023; Lawrence et al., 2024)
Anticipating future trends and concluding insights	Based on your experience, how can AI be smoothly integrated into education? Suggest ways to improve teacher training for AI in education. (Kim et al., 2022; Nazaretsky et al., 2022) Share any final thoughts on your experience with AI in education not previously mentioned. (Lawrence et al., 2024; Ng et al., 2021)

Note. AI = artificial intelligence.

Reflective Learning Activities and Procedure

The 3-week teacher training course, depicted in Table 3, employed ChatGPT to enhance students' understanding and application of various educational theories. The activities stimulated reflective practice and discussion, empowering teachers to integrate AI into their future teaching methodologies. Table 3 outlines preservice teachers' diverse approaches to engaging with ChatGPT across various educational activities.

Table 3

Reflective Learning Activities and Procedures

Week	Content	Activities	Procedures
1 (2 hours)	Behavioral learning principles	• Presentation and discussion on behavioral learning principles • AIED exploration activity • Activity to analyze student behavior in classroom management scenarios using AI tools	Preservice teachers investigate how AI influences learning and teaching processes by examining case studies. Preservice teachers work on classroom management case studies (generated by AI) to analyze student behavior effectively.
2 (2 hours)	Cognitive learning concepts	• Presentation and discussion on cognitive learning theories • AI-assisted cognitive mapping activity	Participants undertake a task where they use AI tools to create cognitive maps of complex topics; process involves identifying key concepts, relationships, and structures within these topics.
3 (2 hours)	Humanist learning perspective	• Presentation and discussion on humanist education philosophy • AI-driven personalized learning quests • Discussion on use of AI tools enhancing teacher autonomy and personalized teaching practice	Preservice teachers develop AI-supported learning quests tailored to individual student needs. Preservice teachers use AI to support learner autonomy in personalized teaching practice, sharing insights in group discussions.

Note. AI = artificial intelligence; AIED = AI in education.

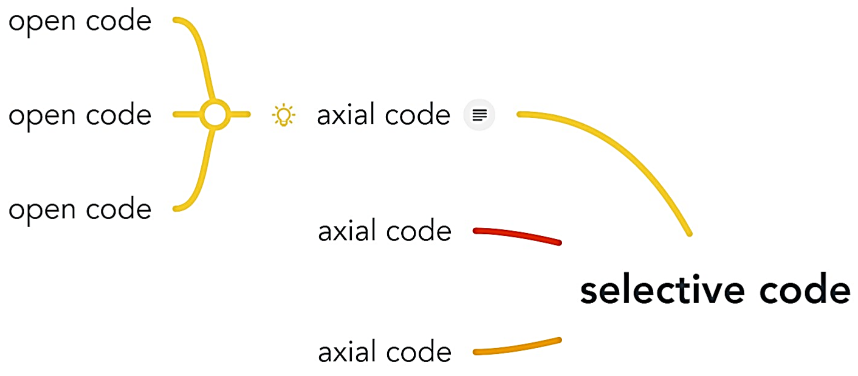
Preservice teachers undertook various activities over the 3 weeks via Zoom. These activities highlighted AI's versatility in educational settings and encouraged critical thinking about teaching practices.

Data Analysis

The narrative data from reflection papers were analyzed using Strauss and Corbin's (2008) open, axial, and selective coding technique, facilitated by NVivo 12 software. The coding technique involves three stages: (a) open coding, which entails the detailed examination of data to identify concepts and explore similarities and differences between events; (b) axial coding, which analyzes the relationships between these concepts to understand them in a broader context; and (c) selective coding, which examines the connections between axial codes to develop an overall theory or framework of understanding. This method focused on narratives to deeply understand preservice teachers' views on AI, offering detailed insights while acknowledging its subjectivity and addressing validity concerns through member checking (Bower et al., 2021). Figure 1 shows the coding process, including open, axial, and selective coding stages.

Figure 1

Open, Axial, and Selective Coding Process



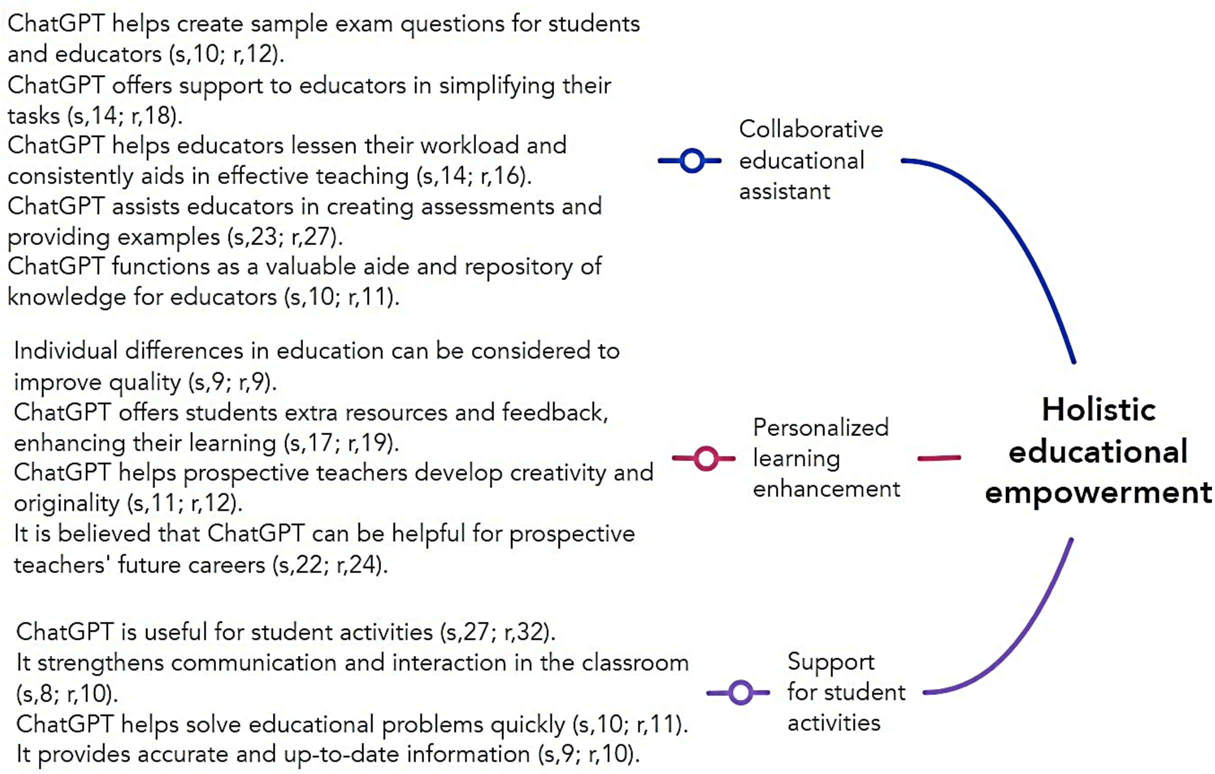
To mitigate the variable effect of students' language proficiency, reflection papers were requested in their native language, Turkish. Analyses were initially conducted in Turkish language, with findings translated into English by the researchers during the manuscript preparation process.

Results

This section presents the findings of our study, exploring how preservice teachers perceive and engage with AI applications in educational settings, mainly focusing on their interactions with ChatGPT. Our narrative inquiry approach enabled us to delve into their experiences, beliefs, and attitudes toward AI's integration into teaching. The research results culminated in six selective codes, each represented by a figure.

Figure 2

Holistic Educational Empowerment



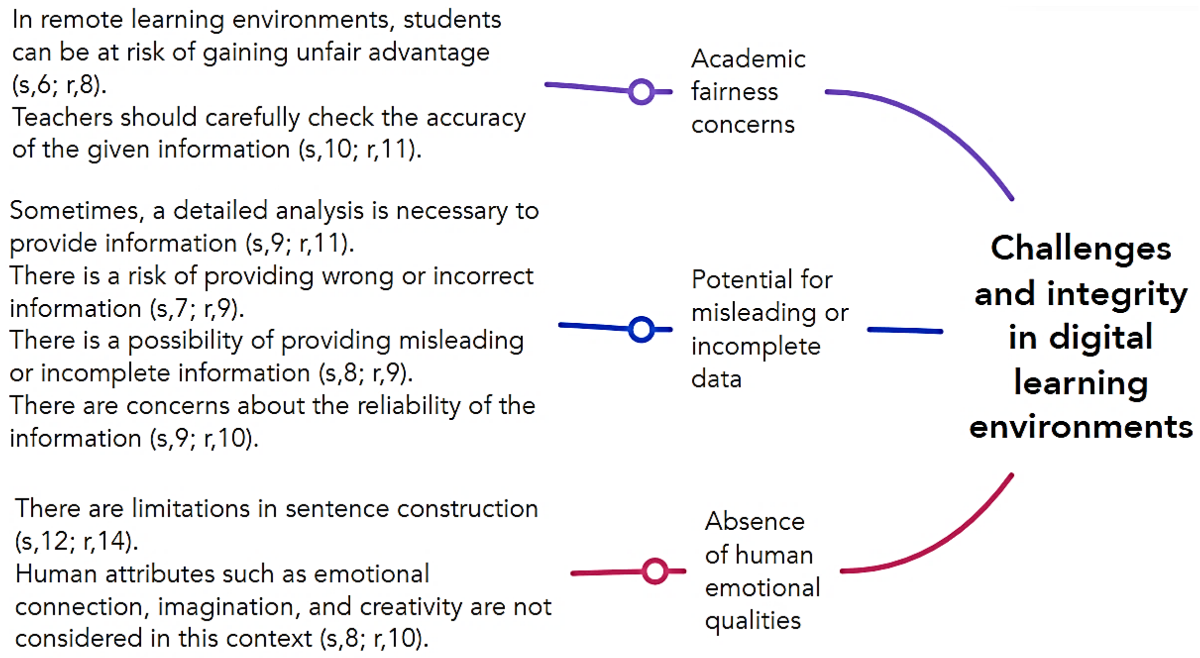
Note. s = source item; shows the number of data sources (number of learners); and r = reference; shows the number of statements reached from the sources.

As illustrated in Figure 2, the findings illuminate ChatGPT's substantial impact on holistic educational empowerment within teaching environments. As a collaborative educational assistant, ChatGPT has significantly reduced educators' workload while enhancing teaching effectiveness; as noted by one student (S45), "ChatGPT's role in reducing educator workload and enabling more personalized teaching is appreciated." The role of ChatGPT goes beyond assisting educators. It is a vital knowledge repository, enriching the teaching–learning process with updated information. The participants stressed its effectiveness in addressing individual learning differences, promoting personalized learning, and fostering inclusivity. One student (S36) noted that "ChatGPT's capability to generate exam questions and offer personalized feedback enhances the learning experience."

Furthermore, the tool facilitates student engagement, strengthens classroom interaction and communication, and fosters creativity among prospective educators. These features simplify educational procedures and cultivate a dynamic, interactive, and personalized learning environment. Nevertheless, one student (S78) advised, "ChatGPT should assist, not substitute, human interaction in education." ChatGPT's comprehensive capabilities significantly contribute to the holistic development of educators and students, aligning with the emergence of an empowered and innovative era in education, as indicated by the qualitative analysis of the student interviews.

Figure 3

Challenges and Integrity in Digital Learning Environments

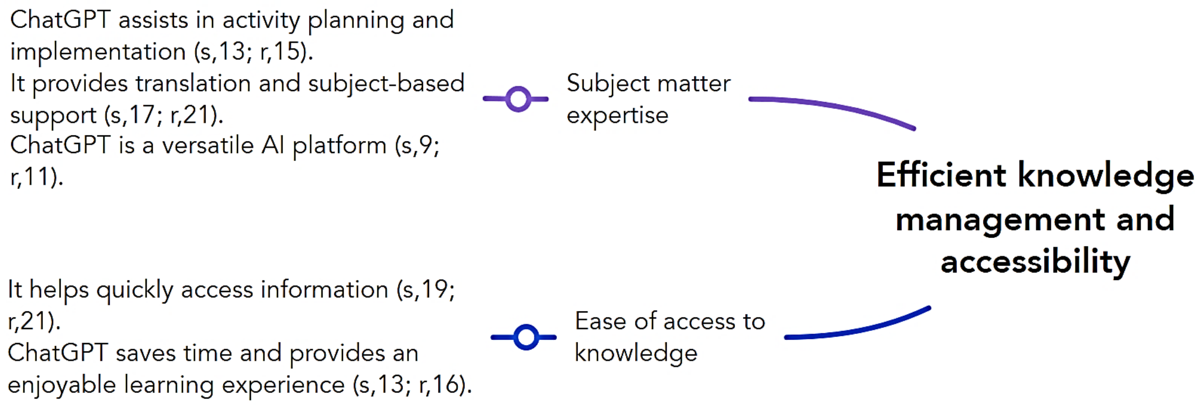


Note. s = source item; shows the number of data sources (number of learners); and r = reference; shows the number of statements reached from the sources.

As outlined in Figure 3, insights from student reflection papers highlighted critical challenges in ensuring integrity and accuracy within digital learning environments. Concerns about the potential for unfair advantages in learning settings were widespread, shedding light on broader academic fairness issues. A common concern among students was that using ChatGPT in the learning environment could result in unfair benefits and potentially diminish critical thinking skills. Students emphasized the responsibility placed on educators to rigorously verify the accuracy of information, highlighting the risks associated with disseminating incorrect or misleading content. It was widely recognized that using ChatGPT in education necessitates extensive fact-checking and explicit guidelines for its use, underlining the critical need for careful scrutiny in selecting and distributing digital educational resources. Additionally, apprehensions regarding the reliability of digitally furnished information were evident, with students noting occasional compromises in their dependability. Concerns were also raised regarding the constraints in sentence construction on digital platforms, which may impede clear communication and comprehension. A notable deficiency identified by the students was the need for more human emotional attributes in digital learning spaces, with core aspects such as emotional connection, imagination, and creativity being marginalized.

Figure 4

Efficient Knowledge Management and Accessibility

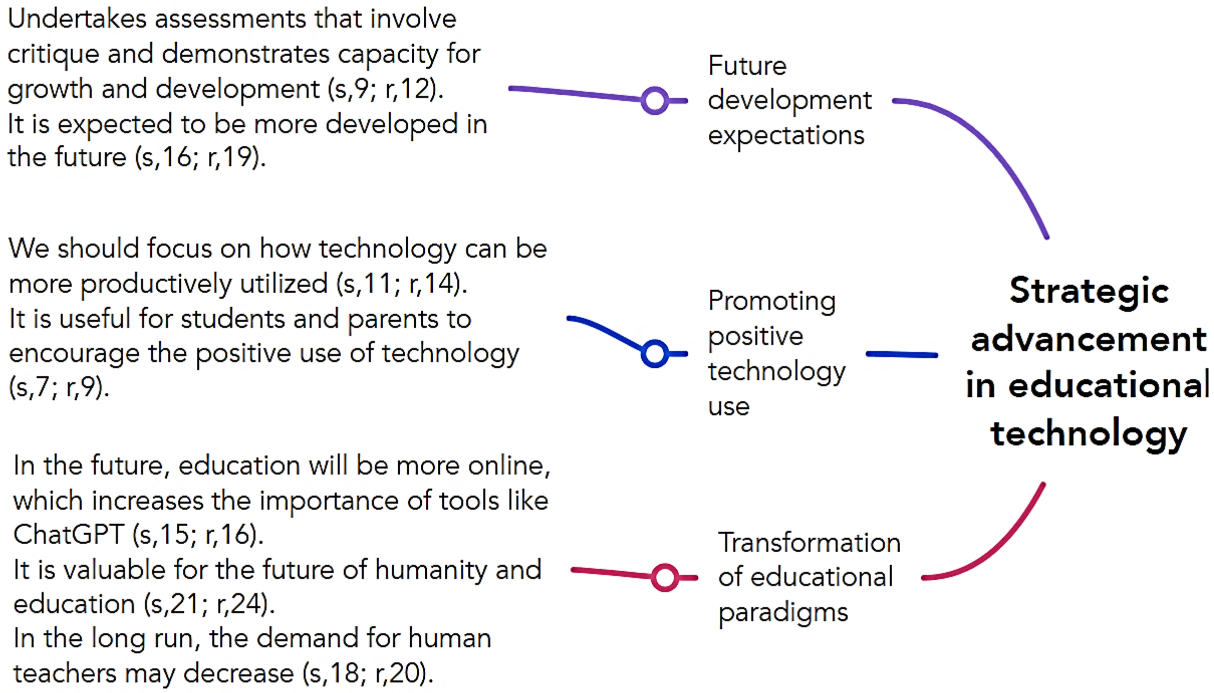


Note. s = source item; shows the number of data sources (number of learners); and r = reference; shows the number of statements reached from the sources.

Figure 4 provides a comprehensive overview of ChatGPT's pivotal role in elevating educational efficiency and broadening accessibility. Specifically, it emerged as a key tool in facilitating activity planning and implementation in educational settings, and students noted its effectiveness in streamlining organizational tasks. This aspect was particularly highlighted in the context of overcoming the logistical challenges in education. As noted by a student (S14), "ChatGPT's customization features and multilingual support are significant for me, aiding in overcoming language obstacles and delivering specialized academic assistance." ChatGPT's capacity to offer translation and subject-specific assistance was greatly appreciated. Students highlighted its essential role in closing language divides and delivering specialized knowledge, thus enriching educational experiences in various learning settings. A prominent benefit of ChatGPT is its rapid access to information, which students credited for significantly improving the efficiency of their learning and research activities. Moreover, the tool was acknowledged to contribute to saved time and more enjoyable learning experiences, which are aspects that students associate with enhanced engagement and satisfaction in the educational process.

Figure 5

Strategic Advancement in Educational Technology

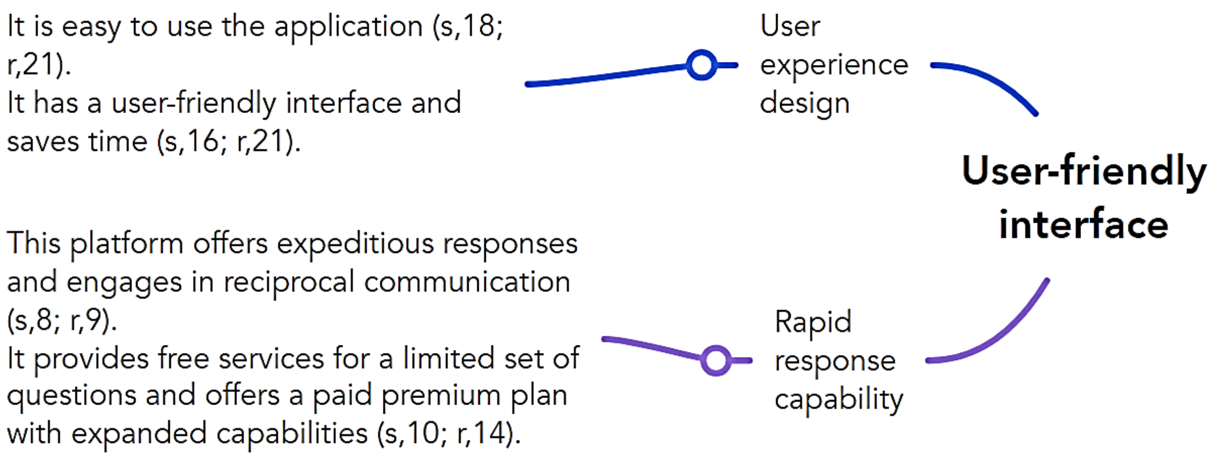


Note. s = source item; shows the number of data sources (number of learners); and r = reference; shows the number of statements reached from the sources.

As outlined in Figure 5, student reflections revealed general agreement on the strategic progression of educational technology. There was clear recognition among students of technology's crucial role in their development and growth, with expectations for more advanced tools to emerge in the future. Students perceived considerable potential in AI tools such as ChatGPT for improving learning experiences, automating tasks, and making education more accessible to people from various socioeconomic backgrounds. Students could foresee the anticipated move toward education, a shift that is likely to heighten the importance of tools like ChatGPT, reflecting a profound change in educational models. In this context, another student remarked that AI presents both opportunities and challenges, improving personalized education while underscoring the need for ethical usage and maintaining human interaction. This shift underscores the importance of technology in shaping the future trajectory of education and its impact on society. Concurrently, the future role of human educators was contemplated, with some students predicting a diminishing demand in the face of advancing digital tools.

Figure 6

User-Friendly Interface

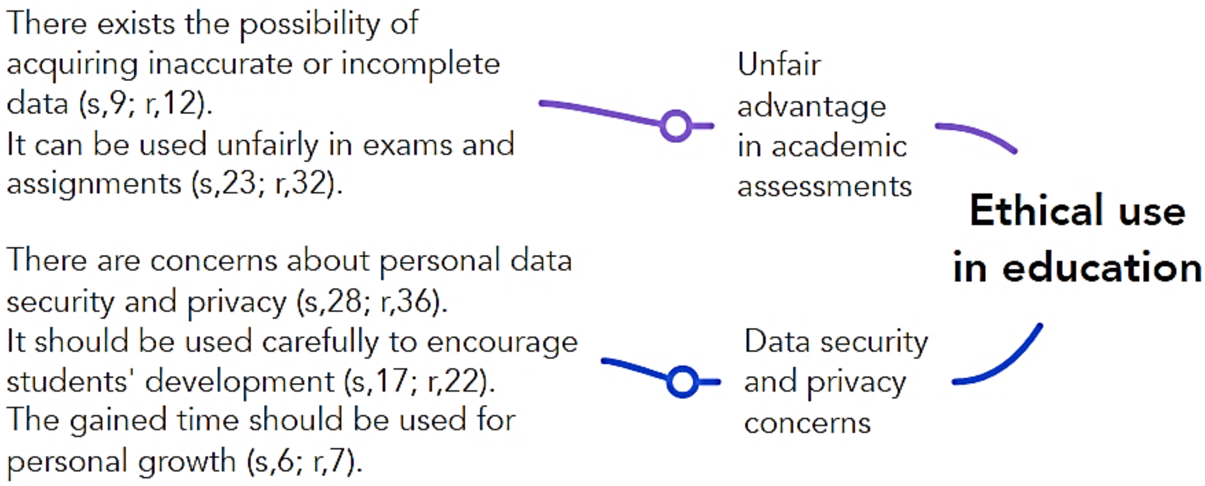


Note. s = source item; shows the number of data sources (number of learners); and r = reference; shows the number of statements reached from the sources.

Figure 6 provides a clear visual summary of how a user-friendly interface is highly valued by students, as reflected in their responses. The findings indicate that students placed high value on a user-friendly interface, which facilitates efficient interaction with the application. Echoing this sentiment, one student (S73) noted, "ChatGPT's user-friendly and responsive interface is accessible to users with different levels of technical skills, which is a major benefit." The importance of intuitive design was underscored by its role in accelerating adoption and saving time. In contrast, a platform's rapid response capability is essential for promoting active communication and enhancing the learning experience. Additionally, free services for basic features and a paid premium plan offering extended capabilities address a broad spectrum of user requirements. This was underscored by a student's (S82) observation: "While I value ChatGPT's user-friendly interface for different applications, the constraints of its free version bring up issues regarding accessibility and inclusivity." Such a strategy reflects a deliberate attempt to balance easy access and offer sophisticated functionalities, a striving to satisfy varied user demands while sustaining a practical financial framework.

Figure 7

Ethical Concerns Regarding AI Use in Education



Note. s = source item; shows the number of data sources (number of learners); and r = reference; shows the number of statements reached from the sources.

As outlined in Figure 7, our data show a strong focus among students on the ethical implications of integrating AI technologies in educational contexts. The data highlight apprehensions about acquiring inaccurate or incomplete data, which may confer an unfair advantage in academic assessments, suggesting the need for rigorous checks to maintain assessment integrity. Students believed that the effectiveness of AI in education largely would depend on prioritizing accuracy (S26, S49, S51), data security (S23, S67, S100), and ethical development to minimize risks and enhance benefits (S16, S88, S109).The discourse extends to the responsible use of technology to foster academic and personal development, with a call to ethically channel time savings from technology use toward personal growth. These perspectives underscore the urgency of incorporating ethical guidelines into the application of educational technologies, ensuring a balance that upholds the integrity of educational practices while maximizing the benefits of technological advancements. Our study advocates for a balanced integration of AI in education. This approach focuses on a framework that not only respects educational integrity but also harnesses technological advancements for empowerment, aiming to minimize inequalities.

Discussion

This research aimed to delve into preservice teachers' perceptions regarding the integration of AI applications, particularly ChatGPT, within educational contexts, with a keen focus on its implications for open and distributed learning environments. It sought to unravel how these emerging digital tools shape future educators' teaching methodologies and professional identities, especially as open and distributed learning becomes increasingly prevalent.

First, the findings on holistic educational empowerment through AI align with broader narratives of AI's potential (e.g., Nazaretsky et al., 2022) and expand on existing research by showcasing AI's ability to reduce workload, personalize teaching, and enhance learning experiences (Hashem et al., 2024). The findings also corroborate with research (Addo & Sentance, 2023; Keeley, 2023), highlighting the significance of ChatGPT in fostering educational effectiveness, accessibility, and teacher–student rapport. Similarly, ElSayary (2024) has provided comparative insights that enrich our understanding of preservice teachers' engagement with ChatGPT. Integrating AI tools like ChatGPT may be more beneficial and transformative than previously anticipated, bridging the gap between theoretical expectations and practical realities of AI in education.

Second, the qualitative analysis revealing ChatGPT's significant role in enhancing educational efficiency and accessibility underscores its value in AI integration within teacher education curricula. Integrating AI into teacher training aligns with calls for its use in education (Chiu et al., 2023; McGovern & Fager, 2007), creating dynamic and flexible learning environments. ChatGPT helps by simplifying tasks, translating languages, offering subject-specific support, addressing language barriers, and diversifying learning methods (ElSayary, 2024; Henry et al., 2021). ChatGPT can be a lifesaver for teachers facing practical challenges (Hashem et al., 2024; Ng et al., 2023), from answering basic questions to solving tough issues. Not only should AI tools be included in teacher training, but they should also be used to manage knowledge efficiently and make education more accessible, ultimately improving teaching practices in the digital age.

Third, the findings on the challenges and integrity in digital learning environments, particularly with AI tools like ChatGPT, resonate deeply with current academic concerns about integrity and the ethical use of technology in open and distributed learning. As highlighted in the literature (Chan, 2023; Chiu et al., 2023), concerns exist that AI tools in education could facilitate cheating. For example, students might use translation apps or automatic grading systems rather than doing the work themselves. This reinforces existing literature highlighting educator guidance as crucial for maintaining academic integrity, especially in online settings (Carvalho et al., 2022; Ng et al., 2021).

Next, the findings illuminate the pivotal role of AI in the strategic advancement of educational technology, particularly resonating with themes of integrating AI in teacher education curricula and the evolution of teacher roles with AI advancements. Reflecting on the implications for teacher education and AI literacy, this study underscores the necessity of integrating AI literacy into teacher education programs, as suggested by Wang and Lu (2023). The comprehensive approach to AI literacy should not only cover the technical use of AI tools but also emphasize ethical considerations and effective strategies for using AI tools (Karataş et al., 2024) in educational settings. This emphasizes the need to prepare future educators for an AI-infused future, where technology augments learning and reshapes teaching. These insights suggest a paradigm shift, highlighting the need for a balanced and ethical approach to AI, where technology and educators work together to enrich the educational experience.

Furthermore, ChatGPT's user-friendly interface, highlighted in this study, aligns with the importance of technology usability for educational adoption (Heintz, 2021). Its intuitive design and broad accessibility support its wide use, echoing calls for user-friendly AI tools in teacher education (Aung et al., 2022). This emphasizes the need for sophisticated yet approachable AI to enhance learning and create a more inclusive educational environment.

Finally, students' ethical concerns about AI in education, including data accuracy and bias, echo existing discussions. Recent studies stress the need for ethical AI use in education, focusing on privacy, bias reduction, and fair access to technology's benefits (Addo & Sentance, 2023; Ng et al., 2021). The link between teachers' understanding of AI and their teaching effectiveness highlights the importance of combining human knowledge with AI technologies (Nazaretsky et al., 2022). To address these issues, research suggests incorporating ethical AI principles into teacher training and curriculum design, aiming for AI integration that respects educational values and promotes inclusivity (Le-Nguyen & Tran, 2023; Wang & Lu, 2023). Participants in our study voiced concerns about AI's accuracy, privacy, and ethical development, emphasizing the need for AI literacy in teacher education to prevent disparities. Echoing this, Keeley (2023) and ElSayary (2024) have called for comprehensive AI literacy programs that include ethical considerations. Additionally, Le-Nguyen and Tran (2023) suggested preparing future educators with strategies to counteract AI's potential negative impacts, advocating for a well-rounded approach to AI literacy that combines technical skills with ethical awareness.

Implications and Significance of the Study in the Context of Teacher Education

This study significantly contributes to the field of teacher education by highlighting the imperative integration of AI literacy, particularly in the context of emerging technologies like ChatGPT, and its relevance for open and distributed learning environments. The findings reveal that AI tools can remarkably enhance holistic educational empowerment, underscoring the need for future educators to be adept in these technologies in order to improve teaching methodologies and address diverse student needs across various learning settings. These insights resonate with current academic discourse that advocates for the inclusion of AI in teacher education curricula (Chiu et al., 2023; Ng et al., 2022). Additionally, the study brings forward critical ethical considerations and practical challenges associated with AI in educational settings, such as concerns over academic integrity and the accuracy of information. The nuanced understanding gained from this research, alongside contributions from ElSayary (2024) and Keeley (2023), further emphasizes the importance of preparing educators for the digital age by incorporating AI literacy into teacher education programs. It is essential to equip preservice teachers with the competencies to use AI tools ethically and effectively, addressing the challenges identified. These revelations underscore the urgency for comprehensive AI literacy programs that balance technological proficiency with ethical use, guiding future research and policy development toward creating a more effective, inclusive, and ethically responsible educational environment. Thus, this research enriches the existing academic dialogue and charts a course for the strategic development of teacher education in an increasingly digital world.

Conclusion

Exploring preservice teachers' perceptions of AI integration in education, particularly within open and distributed learning, uncovers mixed reactions. They see AI as beneficial for enhancing education and tailoring learning to individual needs, yet they express concerns over potential threats to academic integrity and the reliability of information. Reflecting on the narrative inquiry approach, this research highlights preservice teachers' nuanced understanding of AI's role in education, balancing its potential with ethical implications. These insights contribute to the discourse on AI in education and highlight the imperative for balanced, ethically sound, and critical integration of AI technologies in educational settings, including those that extend beyond traditional classroom boundaries. This study contributes new insights into the evolving landscape of teacher education, emphasizing the necessity of a balanced approach to AI integration. It concludes by underscoring AI's transformative yet complex role in education, marking a significant step forward in understanding and preparing for the future of teaching and learning in an increasingly digital and distributed world.

Limitations

This study, investigating preservice teachers' reflections on AI-supported applications in teacher education at a single university, highlights the limited generalizability of findings to broader populations. To extend its applicability, future research should explore diverse educational contexts and include a wider range of participants. The narrative inquiry methodology, while offering deep, personalized insights, introduces subjectivity in data interpretation. Future studies could employ mixed methods to balance subjective narratives with objective measurements, enhancing reliability. Despite these limitations, the insights into AI's integration in teacher education suggest policies should support educators in navigating AI tools, emphasizing training that fosters critical and reflective use. Practices in teacher education must evolve to incorporate AI fluently, preparing educators for AI-augmented teaching environments. This study provides useful insights but is limited by its focus on a specific group. Future research should include participants of diverse nationalities for a more comprehensive understanding, examine the effectiveness of AI tools in diverse educational settings, and create ethical and pedagogical strategies for their integration, addressing the specific advantages and challenges of these technologies.

Recommendations for Future Research and Policy Development

Drawing on the extensive implications of this study, the following recommendations are presented. (a) Future research should focus on evaluating the effectiveness of AI tools such as ChatGPT in varied educational contexts, from open to more structured learning environments, underlining the importance of curriculum development to integrate AI literacy into teacher education curricula, equipping future educators with the skills to use AI tools ethically and effectively (Addo & Sentance, 2023; Wang & Lu, 2023). (b) Particular attention should be paid to developing ethical and pedagogical strategies for integrating AI in ways that enhance learning, alongside professional development opportunities focused on ethical AI use and strategies for leveraging AI to enhance teaching and learning (Nazaretsky et al., 2022; Ng et al., 2023). (c) Simultaneously, policies and regulations need to evolve to foster more AI-inclusive education models, considering the unique benefits and challenges introduced by these technologies across different settings. This includes development and implementation of ethical guidelines for ethical AI use in educational settings to address concerns about academic integrity, information accuracy, and the balance between AI and human interaction in teaching (Le-Nguyen & Tran, 2023). (d) Teachers will likely require enhanced training in both the technical use and ethical implications of AI tools, so investments in educational infrastructure and teacher professional development are warranted. Essential too is (e) further integrating AI literacy, including technical skills as well as considerations of academic integrity and ethical use into teacher training and preparation curricula. Overall, the goal should be developing frameworks for effectively combining AI technologies with more traditional teaching methods and content, aiming to foster a more effective, inclusive, and ethically responsible educational environment.

Acknowledgment

This article builds upon a presentation previously presented in Proceedings of the "Eğitim Araştırmaları Kongresi (EAK2023) [Education Research Congress (ERC2023)]", İstanbul, Türkiye.

Ethical Approval

This research received ethical approval from the Scientific Research and Publication Ethics Committee at Nevsehir Hacı Bektas Veli University in Türkiye, as per decision number 2300062992.

Ethical Use of AI Tools

During the preparation of this research, the authors used Paperpal and Quillbot (AI tools) to paraphrase their writing for more academic enhancement, and ChatGPT and DeepL (AI tools) for language translation and text reduction. After using these AI tools, the authors reviewed and edited the content as necessary, and take full responsibility for the content of the publication. Alongside using AI tools for language tasks, the authors thoroughly reviewed and accurately cited all references in this research. They verified each reference's authenticity, including DOI links. It's crucial to note that all data and findings come from properly cited sources, not AI- generated. The authors fully ensure the research's integrity and accuracy.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

The authors report no funding.

Data Availability

The data and materials used in this study are available upon request from the corresponding author.

References

Adama, E. A., Sundin, D., & Bayes, S. (2016). Exploring the sociocultural aspect of narrative inquiry: A dynamic nursing research methodology. *Clinical Nursing Studies*, *4*(4). https://doi.org/10.5430/cns.v4n4p1

Addo, S. A., & Sentance, S. (2023, December 14–15). Teachers' motivation for teaching AI in K-12 settings. In *HCAIep '23: Proceedings of the 2023 Conference on Human Centered Artificial Intelligence: Education and Practice*. Association for Computing Machinery. https://doi.org/10.1145/3633083.3633192

Araka, E., Oboko, R., Maina, E., & Gitonga, R. (2022). Using educational data mining techniques to identify profiles in elf-regulated learning: An empirical evaluation. *The International Review of Research in Open and Distributed Learning*, *23*(1), 131–162. https://doi.org/10.19173/irrodl.v22i4.5401

Aung, Z. H., Sanium, S., Songsaksuppachok, C., Kusakunniran, W., Precharattana, M., Chuechote, S., Pongsanon, K., & Ritthipravat, P. (2022). Designing a novel teaching platform for AI: A case study in a Thai school context. *Journal of Computer Assisted Learning*, *38*(6), 1714–1729. https://doi.org/10.1111/jcal.12706

Bower, K. L., Lewis, D. C., & Paulus, T. M. (2021). Using ATLAS for Mac to enact narrative analysis: Metaphor of generativity from LGBT older adult life stories. *Qualitative Research*, *22*(6), 933–950. https://doi.org/10.1177/1468794121999008

Carvalho, L., Martinez-Maldonado, R., Tsai, Y.-S., Markauskaite, L., & De Laat, M. (2022). How can we design for learning in an AI world? *Computers and Education: Artificial Intelligence*, *3*, Article 100053. https://doi.org/10.1016/j.caeai.2022.100053

Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y.-S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, *2*, Article 100027. https://doi.org/10.1016/j.caeai.2021.100027

Celik, I. (2023). Towards intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, *138*, Article 107468. https://doi.org/10.1016/j.chb.2022.107468

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, *20*(1), Article 38. https://doi.org/10.1186/s41239-023-00408-3

Chase, S. (2005). Narrative inquiry: Multiple lenses, approaches, voices. In N. Denzin & Y. Lincoln (Eds.), *Sage handbook of qualitative research* (3rd ed., pp. 651–679). Sage Publications.

Cheng, Y.-P., Cheng, S.-C., & Huang, Y.-M. (2022). An Internet articles retrieval agent combined with dynamic associative concept maps to implement online learning in an artificial intelligence course. *The International Review of Research in Open and Distributed Learning*, *23*(1), 63–81. https://doi.org/10.19173/irrodl.v22i4.5437

Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *4*, Article 100118. https://doi.org/10.1016/j.caeai.2022.100118

Clandinin, D. J. (2006). Narrative inquiry: A methodology for studying lived experience. *Research Studies in Music Education*, *27*(1), 44–54. https://doi.org/10.1177/1321103X060270010301

Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education* (8th ed.). Routledge. https://doi.org/10.4324/9781315456539

Edwards, C., Edwards, A., Spence, P. R., & Lin, X. (2018). I, teacher: Using artificial intelligence (AI) and social robots in communication and instruction. *Communication Education*, *67*(4), 473–480. https://doi.org/10.1080/03634523.2018.1502459

ElSayary, A. (2024). An investigation of teachers' perceptions of using ChatGPT as a supporting tool for teaching and learning in the digital era. *Journal of Computer Assisted Learning*, *40*(3), 931–945. https://doi.org/10.1111/jcal.12926

García-Peñalvo, F. J. (2023). La percepción de la Inteligencia Artificial en contextos educativos tras el lanzamiento de ChatGPT: disrupción o pánico [The perception of artificial intelligence in educational contexts after the launch of ChatGPT: Disruption or panic?]. *Education in the Knowledge Society (EKS)*, *24*, Article e31279. https://doi.org/10.14201/eks.31279

Gentile, M., Città, G., Perna, S., & Allegra, M. (2023). Do we still need teachers? Navigating the paradigm shift of the teacher's role in the AI era. *Frontiers in Education*, *8*, Article 1161777. https://doi.org/10.3389/feduc.2023.1161777

Guilherme, A. (2019). AI and education: The importance of teacher and student relations. *AI & Society*, *34*(1), 47–54. https://doi.org/10.1007/s00146-017-0693-8

Hashem, R., Ali, N., Zein, F. E., Fidalgo, P., & Khurma, O. A. (2024). AI to the rescue: Exploring the potential of ChatGPT as a teacher ally for workload relief and burnout prevention. *Research and Practice in Technology Enhanced Learning*, *19*, Article 23. https://doi.org/10.58459/rptel.2024.19023

Heintz, F. (2021). Three interviews about K-12 AI education in America, Europe, and Singapore. *KI—Künstliche Intelligenz*, *35*(2), 233–237. https://doi.org/10.1007/s13218-021-00730-w

Henry, J., Hernalesteen, A., & Collard, A.-S. (2021). Teaching artificial intelligence to K-12 through a role-playing game questioning the intelligence concept. *KI—Künstliche Intelligenz*, *35*(2), 171–179. https://doi.org/10.1007/s13218-021-00733-7

Hsu, T.-C., Abelson, H., & Van Brummelen, J. (2022). The effects on secondary school students of applying experiential learning to the conversational AI learning curriculum. *The International Review of Research in Open and Distributed Learning*, *23*(1), 82–103. https://doi.org/10.19173/irrodl.v22i4.5474

Hwang, G.-J., Tu, Y.-F., & Tang, K.-Y. (2022). AI in online-learning research: Visualizing and interpreting the journal publications from 1997 to 2019. *The International Review of Research in Open and Distributed Learning*, *23*(1), 104–130. https://doi.org/10.19173/irrodl.v23i1.6319

Jiang, P., Namaziandost, E., Azizi, Z., & Razmi, M. H. (2022). Exploring the effects of online learning on EFL learners' motivation, anxiety, and attitudes during the COVID-19 pandemic: A focus on Iran. *Current Psychology*, *42*(3), 2310–2324. https://doi.org/10.1007/s12144-022-04013-x

Kaminski, E. (2003). Promoting preservice teacher education students' reflective practice in mathematics. *Asia-Pacific Journal of Teacher Education*, *31*(1), 21–32. https://doi.org/10.1080/13598660301619

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*. https://doi.org/10.1007/s10639-024-12574-6

Kayima, F. (2021). The role of reflective practice in mediating development of preservice science teachers' professional and classroom knowledge. *Interdisciplinary Journal of Environmental and Science Education*, *18*(1), Article e2262. https://doi.org/10.21601/ijese/11364

Keeley, K. (2023). AI-assisted teacher wellness: Theory and practice. In S. Hai-Jew (Ed.), *Generative AI in teaching and learning* (pp. 201–216). IGI Global. https://doi.org/10.4018/979-8-3693-0074-9.ch008

Kim, J., Lee, H., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies*, *27*(5), 6069–6104. https://doi.org/10.1007/s10639-021-10831-6

Kim, K., & Kwon, K. (2023). Exploring the AI competencies of elementary school teachers in South Korea. *Computers and Education: Artificial Intelligence*, *4*, Article 100137. https://doi.org/10.1016/j.caeai.2023.100137

Kutsyuruba, B., & Stasel, R. S. (2023). Narrative inquiry. In J. M. Okoko, S. Tunison, & K. D. Walker (Eds.), *Varieties of qualitative research methods* (pp. 325–332). Springer. https://doi.org/10.1007/978-3-031-04394-9_51

Lawrence, L., Echeverria, V., Yang, K., Aleven, V., & Rummel, N. (2024). How teachers conceptualise shared control with an AI co-orchestration tool: A multiyear teacher-centred design process. *British Journal of Educational Technology*, *55*(3), 823–844. https://doi.org/10.1111/bjet.13372

Le-Nguyen, H.-T., & Tran, T. T. (2023). Generative AI in terms of its ethical problems for both teachers and learners: Striking a balance. In S. Hai-Jew (Ed.), *Generative AI in teaching and learning* (pp. 144–173). IGI Global. https://doi.org/10.4018/979-8-3693-0074-9.ch006

Liu, G. L., & Wang, Y. (2024). Modeling EFL teachers' intention to integrate informal digital learning of English (IDLE) into the classroom using the theory of planned behavior. *System*, *120*, Article 103193. https://doi.org/10.1016/j.system.2023.103193

McGovern, A., & Fager, J. (2007, March). Creating significant learning experiences in introductory artificial intelligence. *SIGCSE '07: Proceedings of the 38th SIGCSE Technical Symposium on Computer Science Education* (pp. 39–43). Association for Computing Machinery. https://doi.org/10.1145/1227310.1227325

Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology*, *53*(4), 914–931. https://doi.org/10.1111/bjet.13232

Ng, D. T. K., Lee, M., Tan, R. J. Y., Hu, X., Downie, J. S., & Chu, S. K. W. (2022). A review of AI teaching and learning from 2000 to 2020. *Education and Information Technologies*, *28*(7), 8445–8501. https://doi.org/10.1007/s10639-022-11491-w

Ng, D. T K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

Ng, D. T. K., Leung, J. K. L., Su, J., Ng, R. C. W., & Chu, S. K. W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educational Technology Research and Development*, *71*(1), 137–161. https://doi.org/10.1007/s11423-023-10203-6

Özdil, B. M., & Kunt, N. (2023). Do bi/multilingual learners play by the rules of the game? A postmodern approach to L1/L2 use and learner investment. *Journal of Language, Identity & Education*. Advance online publication. https://doi.org/10.1080/15348458.2023.2180372

Rodriguez, M. E., Guerrero-Roldán, A. E., Baneres, D., & Karadeniz, A. (2022). An intelligent nudging system to guide online learners. *The International Review of Research in Open and Distributed Learning*, *23*(1), 41–62. https://doi.org/10.19173/irrodl.v22i4.5407

Strauss, A., & Corbin, J. (2008). *Basics of qualitative research (3rd ed.): Techniques and procedures for developing grounded theory*. Sage Publications. https://doi.org/10.4135/9781452230153

Tan, K.-H., & Lim, B. P. (2018). The artificial intelligence renaissance: Deep learning and the road to human-level machine intelligence. *APSIPA Transactions on Signal and Information Processing*, *7*(1), Article e6. https://doi.org/10.1017/atsip.2018.6

Tzeng, J.-W., Lee, C.-A., Huang, N.-F., Huang, H.-H., & Lai, C.-F. (2022). MOOC evaluation system based on deep learning. *The International Review of Research in Open and Distributed Learning*, *23*(1), 21–40. https://doi.org/10.19173/irrodl.v22i4.5417

Vadivel, B., Namaziandost, E., Rezai, A., & Azizi, Z. (2023). A paradigm shift in teaching and learning due to the COVID-19 pandemic: Areas of potential and challenges of online classes. *English as a Foreign Language International Journal*, *27*(2). https://doi.org/10.56498/5062722023

Wang, Y. (2023). Probing into the boredom of online instruction among Chinese English language teachers during the COVID-19 pandemic. *Current Psychology*, *43*, 12144–12158. https://doi.org/10.1007/s12144-022-04223-3

Wang, Y., & Lu, X. (2023). Research on the evaluation and cultivation design of AI literacy to promote the development of cognitive Intelligence among pre-service teachers. *Proceedings—2023 International Conference on Culture-Oriented Science and Technology, CoST 2023* (pp. 116–120). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/CoST60524.2023.00032

Wang, Y., Pan, Z., & Wang, M. (2023). The moderating effect of participation in online learning activities and perceived importance of online learning on EFL teachers' teaching ability. *Heliyon*, *9*(3), Article e13890. https://doi.org/10.1016/j.heliyon.2023.e13890

Xu, S., & Connelly, M. (2010). Narrative inquiry for school-based research. *Narrative Inquiry*, *20*(2), 349–370. https://doi.org/10.1075/NI.20.2.06XU

Zhi, R., Wang, Y., & Wang, Y. (2023). The role of emotional intelligence and self-efficacy in EFL teachers' technology adoption. *The Asia-Pacific Education Researcher*. Advance online publication. https://doi.org/10.1007/s40299-023-00782-6

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

The Acceptance of AI Tools Among Design Professionals: Exploring the Moderating Role of Job Replacement

Hsi-Hsun Yang
Department of Digital Media Design, National Yunlin University of Science and Technology, Yunlin, Taiwan

Abstract

This study proposes a hypothetical model combining the unified theory of acceptance and use of technology (UTAUT) with self-determination theory (SDT) to explore design professionals' behavioral intentions to use artificial intelligence (AI) tools. Moreover, it incorporates job replacement (JR) as a moderating role. Chinese-speaking design professionals in regions influenced by Confucian culture were surveyed. An analysis of 565 valid cases with AMOS (Analysis of Moment Structures) supported the structural model hypothesis. The model explains 52.1% of the variance in behavioral intention to use (BIU), proving its effectiveness in explaining these variances. The results further validate the importance of performance expectancy (PE) over effort expectancy (EE) in influencing BIU. Additionally, it has been shown that the impact on intrinsic motivation (IM) and extrinsic motivation (EM) can be either amplified or diminished by anxiety about JR. For individuals experiencing higher levels of JR anxiety, there is a marked increase in IM. They may perceive adopting AI tools as an opportunity to enhance their skills and job security. Conversely, this anxiety also significantly boosts EM, as the potential for improved efficiency and productivity with AI use becomes a compelling incentive. These findings suggest new paths for academic researchers to explore the psychological impacts of AI on design professionals' roles. For practitioners, especially in human resources and organizational development, understanding these dynamics can guide the creation of training programs that address job replacement anxiety.

Keywords: unified theory of acceptance and use of technology, UTAUT, self-determination theory, generative artificial intelligence, GenAI, job replacement, performance expectancy

The Acceptance of AI Tools Among Design Professionals: Exploring the Moderating Role of Job Replacement

With the rapid popularization of generative artificial intelligence (AI) technology and the declaration of the year 2023 marking the breakout of generative artificial intelligence (GenAI), profound changes have occurred in various aspects of our daily lives (Aktan et al., 2022). GenAI can create various data, such as images, videos, audio, text, and three-dimensional models, and has significantly impacted fields such as science, education, medicine, technology, and business (Zhang & Aslan, 2021). In just 2 years, AI tools have sprung up rapidly, including text generators like ChatGPT and Bard; image generators such as Midjourney, Stable Diffusion, and DALL-E; and video generators including Runway and Lumen5. Among GenAI's various representative works, Chat Generative Pretrained Transformer (ChatGPT) stands out.

ChatGPT is noted as a premier AI tool in research (Korzyński, 2023) and commercialization (Dwivedi et al., 2023). Frey and Osborne predicted in 2017 that automation would have an impact especially in office and administrative support work. The industry professionals surveyed in this study, such as multimedia artists and animators, have a much lower probability of being affected by automation. However, the pace of AI advancement is often underestimated, and more powerful AI tools are continually emerging. AI tools specifically conceived for painting or design are meant to free designers from monotonous, low-value tasks, allowing them to focus on higher levels of creativity. Additionally, businesses benefit from cost savings and efficiency improvements (Du et al., 2023).

A Pew Research Center survey (Vogels, 2023) suggested AI will significantly impact young people's careers. This was highlighted by the 5-month U.S. Hollywood screenwriters' strike in 2023 over GenAI. Appleby (2023) found that 43% of students had experience using AI tools, and half admitted to relying on these tools for assignments and exams. ChatGPT, known for its capacity to produce responses resembling human language, should be approached cautiously. On the other hand, students who lack trust in technology might reject its use, missing out on learning opportunities. This is the key motivation for our study: understanding how the emerging use of AI tools affects the training and creative processes of students preparing to enter the design profession.

Historical data show that technology innovations trigger complex emotions (Gessl et al., 2019). Challenges include uncertainty in adaptation, trust issues, and AI anxiety, all of which hinders rational engagement (Rahman et al., 2022). Several theories have been proposed to explain and predict the acceptance and use of technology; the most notable are the technology acceptance model (TAM) (Davis, 1989) and the unified theory of acceptance and use of technology (UTAUT) (Venkatesh et al., 2003). Empirical tests have demonstrated that UTAUT 2 explained 74% of the variance in consumers' behavioral intention to use (BIU) and 52% of actual technology use (Venkatesh et al., 2016).

Duong et al. (2023) noted that effort expectancy not only directly affects students' actual use of ChatGPT but also indirectly increases their use through performance expectancy (PE) and the behavioral intention to use ChatGPT. PE and effort expectancy (EE) in UTAUT are similar to the perceived usefulness (PU) and perceived ease of use in the TAM, with its four constructs summarizing personal perceptions of technology (PE and EE) and environmental perceptions (social influence [SI] and facilitating condition [FC]) (Dwivedi et al., 2019). Because UTAUT requires many facets to explain a significant variance and the complexity of

SI and FC, potentially leading to inaccurate measurements (van Raaij & Schepers, 2008), in their 2008 study, van Raaij and Schepers initially overlooked the perception of the environment. In this research, we only extracted PE and EE from UTAUT.

The strength of UTAUT lies in its comprehensiveness in explaining the intrinsic connections among numerous psychological and social factors that may affect technology adoption, demonstrating applicability, validity, and stability in its collected data (Lin & Bhattacherjee, 2008). Therefore, in this study, motivation is considered an external variable within UTAUT, and we posit that motivational factors also affect the model's PE and EE. The motivation theory used in this study is based on the self-determination theory (SDT) proposed by Deci and Ryan (1985). Self-determination behavior includes three types of motivation: intrinsic motivation (IM), extrinsic motivation (EM), and amotivation. IM refers to the drive to act due to internal satisfaction, such as interest, fulfillment, and perceived utility; it becomes a powerful source of motivation when individuals can make autonomous decisions (Deci & Ryan, 1985). In contrast, EM is driven by the desire to achieve valuable external outcomes, such as improved job performance, knowledge acquisition, or a promotion; it can also be driven by a desire to avoid unwanted external outcomes, such as job replacement (JR) (Lawler & Porter, 1967; Wang & Wang, 2022). Amotivation refers to a lack of any intention to act. However, as our study's respondents had already used AI tools, there was no need to discuss amotivation.

Amabile (1993) suggested that there might be an interaction between IM and EM, yet research on such effects within UTAUT, particularly for technology acceptance, remains unexplored. Without a deep understanding of how IM and EM affect technology acceptance among design professionals, fully comprehending their intent to use AI tools amid AI advancements and job security concerns becomes unrealistic. For that reason, this study examines JR and its moderating role in the behavioral intention to use AI tools. JR anxiety refers to the anxiety caused by the concern that AI might replace people's current jobs (Wang & Wang, 2022; Wang et al., 2022).

The current research aims to address four gaps in the field. First, most AI tool use studies have centered on ChatGPT users (Duong et al., 2023; Rahman et al., 2022; Shahsavar & Choudhury, 2023), with relatively little research done on GenAI-related tools used by design professionals. Second, many studies have used UTAUT as their only theoretical framework, not deeply exploring vital motivational factors (Du et al., 2023; Shahsavar & Choudhury, 2023). Third, UTAUT's predictability differs by culture (King & He, 2006; Yoo et al., 2012), requiring more research on its use in various cultural contexts. Fourth, research on JR as a moderator between UTAUT and motivation types is limited and needs more investigation. To address these gaps, this study explores how the UTAUT framework and SDT relate, aiming to better understand design professionals' intent to use AI tools and JR's moderating role. Therefore, the study explores Confucian cultures within Chinese-speaking regions to assess UTAUT applicability in non-Western contexts, aiming to enhance its predictive accuracy.

Therefore, this study had three main objectives:

1. To investigate the factors influencing the behavioral intention to use AI tools among design professionals.

2. To develop an expanded UTAUT model that incorporates IM, EM, and JR in the context of AI tool usage.

3. To empirically validate the proposed model.

Hypotheses Development

Performance Expectancy (PE)

PE refers to an individual's anticipated level of improvement in job performance due to using a specific system (Venkatesh et al., 2003). Engel et al. (1995) and Chou et al. (2018) identified PE as a crucial predictor for mobile commerce. In this study, we operationally define PE as an individual's anticipation of improvement in one's ability to complete tasks, achieve goals, and efficiently alleviate their workload by using AI tools.

Studies have shown a close relationship between PE and BIU in the adoption of various technologies (Nikolopoulou et al., 2021). Therefore, we predict that the intention and action of design professionals to use AI tools will significantly grow with their increased PE. If AI tools can meet the PE of design professionals, they will become more attractive to them, making these professionals more willing to continue using AI tools. Based on the studies previously discussed, PE significantly influences design professionals' BIU AI tools. Hence, the following hypothesis was proposed:

H1: A positive relationship exists between PE and BIU in using AI tools.

Effort Expectancy (EE)

EE refers to the anticipated ease of using an information system (Nikolopoulou et al., 2021; Venkatesh et al., 2003), and it is closely related to the amount of effort required while using the system. Additionally, EE is viewed as a fundamental premise in predicting technology acceptance (Nikolopoulou et al., 2021).

Duong et al. (2023) found that EE directly influenced students' actual use of ChatGPT and indirectly increased their use through PE and the intention to use ChatGPT. Teo and Noyes (2014) discovered that EE affects consumers' behavioral intention to use technology. However, some studies found contrary results. Research on the adoption intention of mobile technology did not show a significant direct relationship between EE and BIU (Morosan & DeFranco, 2016). Thus, this remains an open issue worthy of attention.

In this study, EE is operationally defined as an individual's perception of ease and effortlessness in using AI tools. When design professionals perceive AI tools as seamless and efficient to use, they are more likely to integrate AI tools into their work. A positive perception of EE when using AI tools is a strong indicator significantly influencing design professionals' willingness to accept AI tools to optimize their design work. When the interface is friendly, intuitive, and easy to interact with, the possibility of user–system interaction increases (Duong et al., 2023). Thus, we proposed this hypothesis:

H2: A positive relationship exists between EE and BIU in using AI tools.

Intrinsic Motivation (IM)

According to Ryan and Deci (2000), people differ in their amounts of motivation and the types of motivation they experience. In other words, different individuals possess different motivational orientations, namely IM and EM, and varying levels of motivational intensity. Crucially, there is an interplay between IM and EM. IM refers to the driving force that originates from an individual, such as finding an activity exciting or challenging, instead of being driven by external stimuli, pressures, or rewards. Individuals can gradually develop IM when they can freely express their feelings under certain conditions.

In this study, IM refers explicitly to the psychological satisfaction that design professionals experience when using AI tools. Ryan and Deci (2000) noted that when people work in an environment that supports autonomy, they feel capable, which enhances IM. Thus, to maximize the intrinsic drive of design professionals, it is necessary for them to achieve goals and receive appropriate rewards for their work. However, it is also important to be aware that IM can be weakened by the forces of consistency in the environment, social recognition, and the reduction of expected tangible rewards.

Oliver (1974) found that EM can serve as an indicator for measuring PE. However, Tyagi (1985) argued that the impact of IM on PE is more significant compared to its effect on EE. In the study by Zhao et al. (2018), self-presentation was regarded as a second-order formative indicator of IM, and the authors noted that if Twitch could satisfy the intrinsic and extrinsic needs of the broadcasters, their PE would be strengthened and enhanced. Therefore, this study put forward the following hypotheses:

H3: A positive relationship exists between IM and PE in using AI tools.

H5: A positive relationship exists between IM and EE in using AI tools.

Extrinsic Motivation (EM)

EM guides the behavior taken by individuals to achieve specific outcomes, such as receiving external rewards (Ryan & Deci, 2000). Hars and Ou (2014) considered EM to include direct and indirect economic rewards and social recognition elements. Zhao et al. (2018) categorized anticipated extrinsic reward, self-esteem benefits, social benefits, and feedback as second-order formative indicators of EM, noting that if Twitch satisfied broadcasters' social benefits gained from audience feedback and interactions, it would directly impact their PE.

The theories of motivation by Rotter et al. (1972) and Overmier and Lawry (1979) demonstrated that people act only when they anticipate achieving a certain result, or they will choose actions that are valuable to them. Thus, EM primarily focuses on achieving outcomes or goals that are separate from the behavior itself. Based on this, the current study integrated the concept of EM with that of JR. Wang et al. (2022) showed that when people felt anxious about learning AI, their motivation to learn decreased because they could not perceive the practicality and enjoyment of learning AI. However, when people feared that AI might replace human jobs, it actually motivated them to learn AI. The motivation of design professionals to perceive their job as enhanced rather than replaced by AI contributes to increased EE. This encouragement prompts them to actively learn and engage with AI tools, providing more opportunities and fostering a positive perception of the tools' ease of use. Therefore, based on these findings, we propose the following hypotheses:

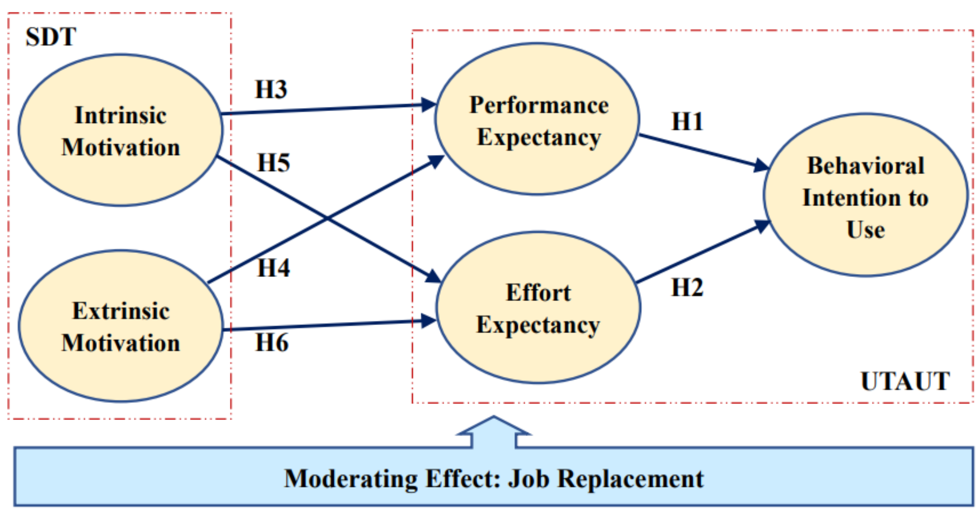
H4: A positive relationship exists between EM and PE in using AI tools.

H6: A positive relationship exists between EM and EE in using AI tools.

This study therefore study suggests that both IM and EM impact PE and EE, directly or indirectly affecting BIU to use AI tools. Additionally, JR is considered a moderating factor. The conceptual framework of this research is shown in Figure 1.

Figure 1

Research Model



Note. SDT = self-determination theory; UTAUT = unified theory of acceptance and use of technology.

Methodology and Research Design

Research Site and Sampling

This cross-sectional study, conducted from January 8 to 15, 2024, employed a self-administered online questionnaire to collect data from samples. The target respondents were users of AI tools in regions influenced by Confucian culture, specifically those with experience using AI for design-related tasks and who were part of the Chinese-speaking community. All respondents provided informed consent, and we guaranteed the confidentiality of their responses. The survey was created using Google Forms and an online platform named Wenjuanxing (https://www.wjx.cn/) and was distributed through social media platforms. The study protocol excluded individuals who (a) were under age 18 and (b) lacked prior experience with the AI tools in question. Among the 568 individuals who initially engaged with the survey, three questionnaires were filled out with only “strongly agree” or “strongly disagree” as responses to all questions. Therefore, we determined that the other 565 responses were considered valid and suitable for further data analysis. The demographic profile of these respondents is illustrated in Table 1.

Table 1

Demographic Profile of Respondents (n = 565)

Variable	Value label		Frequency	Valid %
Gender	1.	Male	232	41.06
	2.	Female	333	58.94
Age	1.	18–20	167	29.56
	2.	21–25	342	60.53
	3.	26–30	29	5.13
	4.	31–40	11	1.95
	5.	41+	16	2.83
Education	1.	Undergraduate design students	429	75.93
	2.	Graduate design students (Master’s/PhD level)	67	11.86
	3.	Alumni with a design major	57	10.09
	4.	Design graduates (Master’s/PhD-level alumni)	12	2.12
Frequency of using AI tools (times per week)	1.	1–3	212	37.52
	2.	4–6	193	34.16
	3.	7–10	106	18.76
	4.	11–15	27	4.78
	5.	16+	27	4.78
Anxiety of job replacement by AI	1.	Yes	268	47.43
	2.	No	297	52.57
		Total	565	100.00

Note. AI = artificial intelligence.

Instrument Development

This research measured the latent variables, as illustrated in Figure 1, using reflective latent constructs slightly adapted from prior studies (Duong et al., 2023; Engel et al., 1995; Hars & Ou, 2014; Morosan & DeFranco, 2016; Nikolopoulou et al., 2021; Overmier & Lawry, 1979; Rotter et al., 1972; Ryan & Deci, 2000; Teo & Noyes, 2014; Zhao et al., 2018). The hypotheses aimed to elucidate the nature of certain relationships or to identify differences among groups or the independence of two or more factors in a given scenario. Reflective constructs were selected because each latent variable was represented by multiple observed variables, which were considered manifestations of the underlying construct. This selection aligns with the capabilities of structural equation modeling (SEM) to rigorously test these complex relationships and to assess the reliability and validity of the constructs.

The questionnaire used in this study comprised structured, closed-ended questions. Respondents provided their answers based on their personal feelings and cognitions. The items were scored a 7-point Likert scale format. First, the questionnaire gathered basic respondent data such as gender, age, education, frequency of AI tool usage, and JR concerns. See Appendix for survey items.

Next, we focused exclusively on PE and EE. This decision was informed by the direct relevance of these constructs to our study's aims, their demonstrated impact on technology acceptance, and considerations of measurement reliability and validity within our specific research context. Each of these constructs comprised three items: PE, EE, and BIU (Duong, et al., 2023; Venkatesh et al., 2003; Venkatesh et al., 2012).

Last, we assessed design professionals' IM and EM for using AI tools based on the motivation structure emanating from SDT. We employed the motivation scale developed by Deci and Ryan (1985) and Deci et al. (2001), which has received global application and validation. Four items gauged IM and five items measured EM. The selection and adaptation of these items were informed by their established reliability and validity across various contexts. In addition, we integrated additional scale items from recent studies by Fan et al. (2012) and Wang et al. (2022). Overall, the measurement instrument incorporated a total of 21 items, and our research model consisted of five constructs.

Analysis Method

In this study, IBM SPSS 28 was employed for deriving descriptive statistics, conducting item analysis, and carrying out reliability and validity assessments. Additionally, SEM was performed using SPSS AMOS 26 to evaluate the fit of the research model. SEM is a robust statistical technique capable of simultaneously analyzing multiple regression equations and is notably prevalent in social work–related literature (Shek & Yu, 2014). This research focused on exploring the structural relationships between SDT and UTAUT, assessing both direct and indirect interactions among exogenous and endogenous variables within a complex structure, as guided by SEM analysis (Barbara, 1998; Kline, 2005).

Empirical Analysis and Results

Sample Profile

The research encompassed a sample size of 565 individuals (Table 1). In terms of gender, female participants accounted for the largest number (333, 58.94%). In regard to age, the 21–25 group was the largest (342, 60.53%). Regarding education, undergraduate design students composed the highest number (429, 75.93%). Regarding use of AI tools, the group using AI one to three times per week had the largest number (212, 37.52%). In response to JR, the group responding "No" accounted for the largest number (297, 52.57%).

Model Reliability and Validity

Reliability and Convergent Validity

Table 2 shows that the absolute value of skewness is less than 2 and the absolute value of kurtosis is less than 7 (Kline, 2005). Thus, the data are normally distributed. The item PE01 had the highest mean (5.742), while EM03 had the lowest (5.067). That is, the respondents agreed the most with PE01 and disagreed the most with EM03.

All constructs exhibited strong composite reliability and average variance extracted (AVE), meeting recommended standards (Fornell & Larcker, 1981; Hair et al., 2019). See Table 2 for details on standard deviations and composite reliabilities in the range of .800 to .851. These results confirm acceptable convergent validity.

Table 2

Statistics for Each Construct

Construct	Item	*M*	*SD*	Skewness	Kurtosis	Std.	CR	AVE
IM	IM01	5.442	1.271	−0.754	0.357	0.719	0.841	0.569
	IM02	5.382	1.322	−0.858	0.759	0.782		
	IM03	5.412	1.279	−0.811	0.431	0.757		
	IM04	5.524	1.178	−0.890	1.436	0.758		
EM	EM01	5.579	1.170	−1.048	1.755	0.757	0.836	0.509
	EM02	5.381	1.308	−0.682	0.136	0.739		
	EM03	5.067	1.338	−0.404	−0.027	0.667		
	EM04	5.465	1.423	−0.972	0.562	0.568		
	EM05	5.736	1.173	−1.197	1.865	0.810		
PE	PE01	5.742	1.121	−1.572	4.200	0.790	0.849	0.585
	PE02	5.476	1.261	−0.824	0.732	0.767		
	PE03	5.674	1.194	−1.201	2.032	0.758		
	PE04	5.657	1.194	−0.971	1.333	0.743		
EE	EE01	5.265	1.292	−0.701	0.542	0.731	0.800	0.501
	EE02	5.287	1.292	−0.654	0.357	0.742		

	EE03	5.127	1.278	−0.583	0.179	0.670		
	EE04	5.280	1.269	−0.756	0.640	0.685		
BIU	BIU01	5.418	1.214	−0.836	0.998	0.790	0.851	0.588
	BIU02	5.297	1.295	−0.753	0.346	0.766		
	BIU03	5.458	1.241	−0.905	0.907	0.716		
	BIU04	5.471	1.276	−0.841	0.639	0.793		

Note. Std. = standardized factor loadings; CR = composite reliability; AVE = average variance extracted; IM = intrinsic motivation; EM = extrinsic motivation; PE = performance expectancy; EE = effort expectancy; BIU = behavioral intention to use.

Discriminant Validity

Discriminant validity, assessed following Fornell and Larcker's (1981) method, confirms that all AVE values exceed correlation coefficients (Table 3). The report found that the correlation between IM and PE is slightly larger than the AVE root value of PE, but the difference is only 0.017 (< 0.1) which can be viewed as a negligible correlation based on random sampling error (Schober et al., 2018). The result still shows great discriminant validity among constructs.

Table 3

Results of Discriminant Validity by Average Variance Extracted

	AVE	IM	EM	EE	PE	BIU
IM	.569	**.754**				
EM	.509	.616	**.713**			
EE	.585	.578	.509	**.765**		
PE	.501	.725	.671	.507	**.708**	
BIU	.588	.723	.622	.528	.630	**.767**

Note. AVE = average variance extracted; IM = intrinsic motivation; EM = extrinsic motivation; EE = effort expectancy; PE = performance expectancy; BIU = behavioral intention to use. The items in bold represent the square roots of the AVE; off-diagonal elements are the correlation estimates.

Model Fit

Whittaker and Schumacker (2022) recommend reporting nine widely accepted fitness metrics to assess model fit. A good model fit typically results in a Chi-square value/degrees of freedom ratio below 3. Additionally, Hu and Bentler (1999) recommend evaluating each fitness metric independently and controlling type I errors with more demanding model fit metrics, such as the comparative fit index (> .90), standardized root mean square residual (< .08), and root mean square error of approximation (< .08) (Table 4).

Table 4

Model Fit

Model fit	Criteria	Model fit of research model
ML χ^2	The smaller the better	487.430
df	The larger the better	182
Normed χ^2 (χ^2/df)	$1 < \chi^2/df < 3$	2.678
RMSEA	< .08	.055
SRMR	< .08	.048
TLI (NNFI)	< .90	.938
CFI	< .90	.946
GFI	< .90	.924
AGFI	< .90	.904

Note. ML = maximum likelihood; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; TLI =Tucker-Lewis index; NNFI =non-normed fit index; CFI = comparative fit index; GFI =goodness of fit index; AGFI = adjusted goodness of fit index.

Path Analysis

In Table 5, the results of path analysis demonstrate significant associations among the constructs. For instance, PE (β = 0.575, p < .001) and EE (β = 0.292, p < .001) significantly affected BIU. The combined influence of these values explained 52.1% of the variance of BIU. IM (β = 0.511, p < .001) and EM (β = 0.367, p < .001) significantly affected PE. The combined influence of these values explained 65.3% of the variance of PE. IM (β = 0.460, p < .001) and EM (β = 0.264, p < .001) significantly affected EE. The combined influence of these values explained 39.5% of the variance of EE (Figure 2).

Table 5

Regression Coefficients

Hypothesis	DV	IV	Unstd.	*SE*	Unstd./*SE*	*p*	Std.	R^2	Result
H1	BIU	PE	0.575	0.057	10.015	.000	.531	.521	Supported
H2		EE	0.292	0.052	5.649	.000	.288		Supported
H3	PE	IM	0.511	0.054	9.453	.000	.528	.653	Supported

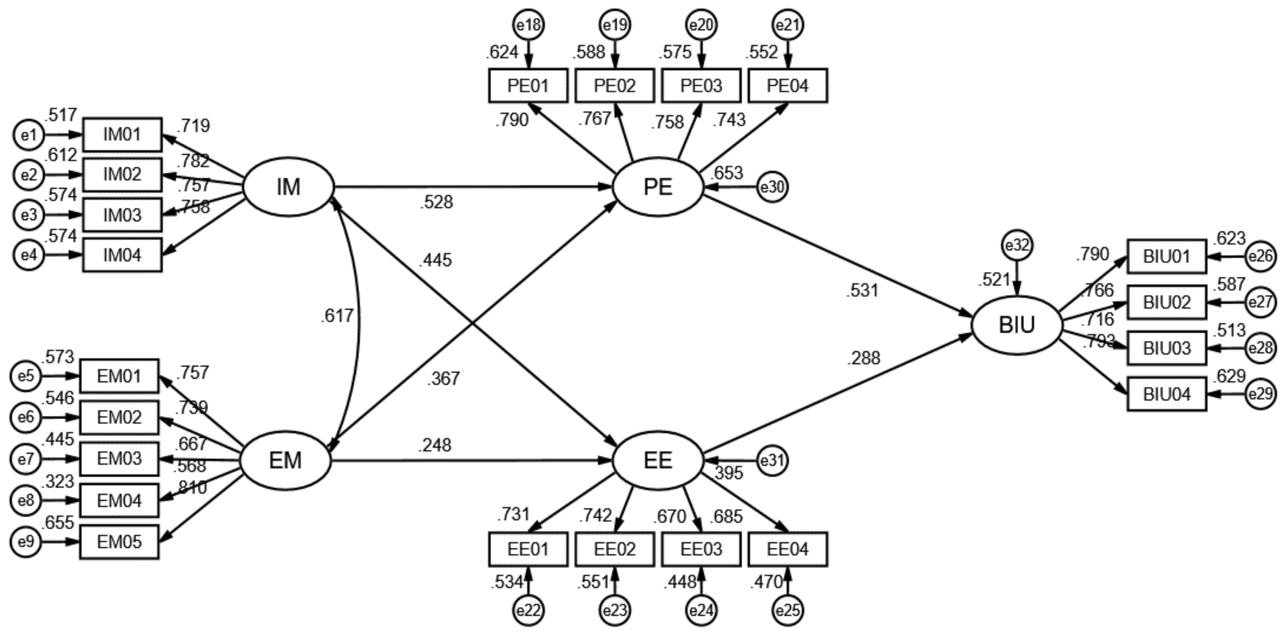
H4		EM	0.367	0.052	7.110	.000	.367		Supported
H5	EE	IM	0.460	0.067	6.838	.000	.445	.395	Supported
H6		EM	0.264	0.065	4.041	.000	.248		Supported

Note. DV = dependent variable; IV = independent variable; Unstd. = unstandardized factor loadings; SE = standard error; Std. = standardized factor loadings; BIU= behavioral intention to use; PE = perceived effort; EE = expected effort; IM = intrinsic motivation; EM = extrinsic motivation.

*** $p < .001$.

Figure 2

Structural Equation Modeling



Note. IM = intrinsic motivation; PE = perceived effort; BIU = behavioral intention to use; EM = extrinsic motivation; EE = expected effort.

Mediation Effects

The bootstrapping method is most commonly used to examine the indirect effect of intermediary variables. It is statistically more powerful than causal path methods and coefficient product methods (Williams & MacKinnon, 2008). Confidence intervals for indirect effects obtained through bootstrapping are statistically stable. However, when 0 is not found within the CIs' lower and upper bounds, bias correction from bootstrapping is suggested (Briggs, 2006; Williams & MacKinnon, 2008).

As shown in Table 6, the total effect IM→BIU, $p < .05$, bias-corrected CI does not include 0. The existence of total effect was supported. The specific indirect effect IM→PE→BIU, $p < .05$, bias-corrected CI does not include 0. Thus, the hypothesis that the existence of a specific indirect effect was supported. The specific indirect effect IM→EE→BIU, $p < .05$, bias-corrected CI does not include 0. Thus, the hypothesis that the existence of a specific indirect effect was supported.

The total effect EM→BIU, $p < .05$, bias-corrected CI does not include 0. The existence of total effect was supported. The specific indirect effect EM→PE→BIU, $p < .05$, bias-corrected CI does not include 0. Thus, the hypothesis that the existence of a specific indirect effect was supported. The specific indirect effect EM→EE→BIU, $p < .05$, bias-corrected CI does not include 0. Thus, the hypothesis that the existence of a specific indirect effect was supported.

Table 6

Mediating Effects

Effect	Point estimate	Product of coefficients			Bootstrap 1,000×	
		SE	*Z*	*p*	Bias-corrected 95% CI	
					Lower bound	Upper bound
Total effect						
IM→BIU	.429	0.082	5.210	.000	0.291	0.603
Specific indirect effect						
IM→PE→BIU	.294	0.085	3.456	.001	0.144	0.479
IM→EE→BIU	.134	0.062	2.165	.030	0.054	0.311
Total effect						
EM→BIU	.289	0.071	4.078	.000	0.156	0.441
Specific indirect effect						
EM→PE→BIU	.211	0.065	3.258	.001	0.101	0.356
EM→EE→BIU	.077	0.038	2.044	.041	0.023	0.182

Note. IM = intrinsic motivation; BIU = behavioral intention to use; PE = perceived effort; EE = expected effort; EM = extrinsic motivation.

Moderator Effects: Job Replacement

To examine the impact of concerns about job replacement, this study categorized responses into "Yes" (indicating anxiety about job replacement) and "No" (indicating no anxiety about job replacement). "Job replacement anxiety" was used as a moderator. Tables 7 and 8 display JR coefficients for these two groups. Out of the six slope comparisons, the IM to PE, IM to EE, and EM to PE comparisons reached statistical significance. Notably, the IM to PE and IM to EE coefficients are higher for "Yes." The EM to PE coefficient is higher for "No."

Table 7

Job Replacement Estimates

IV	DV	Yes (268)				No (297)			
		Estimate	*SE*	*Z*	*p*	Estimate	*SE*	*Z*	*p*
PE	BIU	.568	0.091	6.222	.000	.583	0.074	7.911	.000
EE	BIU	.239	0.079	3.045	.002	.320	0.070	4.599	.000
IM	PE	.582	0.069	8.431	.000	.333	0.088	3.768	.000
IM	EE	.566	0.087	6.497	.000	.271	0.109	2.480	.013
EM	PE	.276	0.058	4.784	.000	.562	0.102	5.514	.000
EM	EE	.218	0.076	2.874	.004	.355	0.120	2.958	.003

Note. IV = independent variable; DV = dependent variable; PE = perceived effort; BIU = behavioral intention to use; EE = expected effort; IM = intrinsic motivation; EM = extrinsic motivation.

Table 8

Job Replacement of Nested Model Differences

Model	Model fit				Nested model differences		
	NPAR	χ^2	*df*	χ^2/df	Δdf	$\Delta\chi^2$	*p*
Default	98	756.072	364	2.077			
PE→BIU	97	756.088	365	2.071	1	0.016	.900
EE→BIU	97	756.609	365	2.073	1	0.537	.464
IM→PE	97	760.830	365	2.084	1	4.757	.029
IM→EE	97	760.473	365	2.083	1	4.401	.036
EM→PE	97	762.544	365	2.089	1	6.472	.011
EM→EE	97	757.015	365	2.074	1	0.943	.332

Note. NPAR = number of parameters; PE = perceived effort; BIU = behavioral intention to use; EE = expected effort; IM = intrinsic motivation; EM = extrinsic motivation.

Discussion and Conclusions

Key Findings

The statistical analysis in Chapter 4 of this dissertation showed that PE and EE significantly and positively impacted the intention to use AI tools. IM and EM also significantly and positively affected PE and EE. The study identified four mediating effects: PE and EE mediate the relationship between IM and BIU, and both also mediate between EM and BIU. Furthermore, three significant moderation effects were found: JR moderates the effects of IM on PE and EE, as well as the effect of EM on PE. This study developed a relationship model covering four major aspects. The overall structural model demonstrated goodness of hit, and hypotheses 1–6 were supported. For the endogenous latent variables of BIU, PE, and EE, the R^2 values reached 52.1%, 65.3%, and 39.5%, respectively. The study's research model can therefore effectively explain these variables' variance.

The findings of this study indicate that PE and EE significantly and positively impacted design professionals' willingness to use AI tools. This aligns with the original hypotheses of the UTAUT model (Venkatesh et al., 2003). PE and EE have been confirmed as key factors influencing behavioral intentions (Davis, 1989; Venkatesh & Davis, 2000). This means that when design professionals believe AI tools can effectively complete tasks and are convenient, their willingness to use them increases. This result is consistent with current research findings on the behavioral intention to use GenAI tools (Du & Gao, 2023; Duong et al., 2023; Shahsavar & Choudhury, 2023).

In addition, this study uncovered a relatively less discussed phenomenon: IM and EM can positively influence design professionals' perceptions of PE when using AI tools. Through SDT, this research showed how, without external pressures and distractions, individuals' needs for internal growth and psychological needs can be met (Deci & Ryan, 2000). Additionally, this study introduces a relatively new hypothesis: IM and EM are hypothesized to positively influence design professionals' EE toward AI tools. Tyagi (1985) suggested that the impact of IM on PE is more significant than on EE; the results of this study show a similar trend. However, design professionals' positive perception of EE significantly increased when AI tools met their IM and EM needs.

The first mediator in the link from IM to BIU involves mediation. The effect of PE mediating between IM and BIU is over twice as strong as that of EE. This shows PE's greater importance compared with EE, as seen in studies by Shahsavar and Choudhury (2023) and Zhao et al. (2018), where only PE is considered. Another mediator is from EM to BIU: PE's mediating effect between EM and BIU is three times stronger than EE's. This further proved PE's substantial impact on BIU.

In the moderation aspect, the results revealed a notable phenomenon: for design professionals who expressed concerns about AI tools potentially replacing human jobs, IM had a stronger effect on PE and EE (Tables 7 and 8). This suggests that when job security is perceived as threatened, IM plays a more crucial role in enhancing PE and EE**.** In contrast**,** for those not worried about AI tools replacing jobs, the increasing influence of EM on PE is also thought-provoking. Design professionals concerned about job security seemed more inclined to boost their self-efficacy by enhancing their needs for relatedness, competence, and autonomy (Ryan & Deci, 2000), leading to a higher acceptance of AI tools. Thus, the study showed that concerns about JR (a moderating variable) amplified the influence of IM on PE and EE.

The participants' anxiety about JR due to AI is considered facilitating anxiety (Alpert & Haber, 1960) that positively affected IM. This study's results are partially contrary to those of Wang et al. (2022). In the study by Wang et al. (2022), EM, but not IM, was found to have a positive effect on the participants. This contrasts with the findings of Donnermann et al. (2021), in which there was no significant correlation between IM and PU, where PU is equivalent to PE. However, our findings indicate that JR anxiety among design professionals has a positive impact on both EM and IM. We also found that those with lower JR anxiety had higher EM than people with higher levels of anxiety. We found that people with higher JR anxiety due to AI had stronger IM, which is consistent with Piniel and Csizér's (2013) results showing that individuals with higher degrees of facilitating anxiety were found to invest more effort and persistence into learning professional knowledge and skills.

Statistics from Tables 7 and 8 indicate that the effect of EM on PE is significant. Consequently, anxiety about JR due to AI had a more significant impact on the pleasure of learning itself than on the rewards of learning AI-related skills, thus relatively weakening the influence on EM. The following are possible reasons for this. First, Confucianism, promoting moderation, seeks a balance between AI technology and the human, stressing that tech progress should boost social harmony and human growth. Work is a livelihood meant to fulfill personal values and duties (Zhu, 2020). Second, design professionals have higher professional confidence than the general public, which we will detail in section of Practical Implications.

JR anxiety has a dual effect on IM and EM, either boosting or lessening it. This resonates with previous studies on technology avoidance attitudes and behaviors (Huang & Haried, 2020; Maduku et al., 2023). This research innovatively reveals the varied impacts within the model, examining how JR influences design professionals' attitudes toward using AI tools. To our knowledge, this topic has yet to be discussed. Furthermore, we discovered that the moderating role of JR in intrinsic and extrinsic motivation differs significantly.

Theoretical Implications

This study has four main theoretical implications. First, the empirical results provide additional evidence that clarifies the relationship between design professionals and AI tools under the integration of UTAUT and SDT, offering a more comprehensive perspective on how design professionals accept and use AI tools. Second, the study confirms the importance of IM and EM for technology acceptance and usage intentions, further revealing how motivational factors affect BIU through PE and EE. This underscores the necessity of considering motivational factors in technology acceptance research. Third, this study is the first to explore the role of JR anxiety as a moderating variable in using AI tools, finding that JR concerns affect the relationship between IM and EM and behavioral intentions. This offers new theoretical insights into psychological factors in technology acceptance. Fourth, by focusing on AI tools, this research provides a deeper understanding of AI technology acceptance and use behavior, specifically in design professionals. This helps us theoretically understand how professionals accept emerging technologies.

Practical Implications

The design industry is a highly specialized and innovative organization where, in addition to professionals needing keen observation and skillful hands, the integration of AI technology is an inevitable trend in the modern era. Through this study's understanding of design professionals' behavioral intentions to use AI tools, developers and marketers can more accurately design and promote AI tools to meet their actual needs and expectations. The primary focus is on improving PE and EE, while design firm managers should focus on intrinsic and extrinsic motivations. Given the positive impact of IM and EM on enhancing PE and EE, design companies should devise effective incentive strategies to encourage employees to learn and use AI tools. For instance, appropriate training could be provided to reduce professionals' learning curve in using these tools, and practical reward systems could be offered to drive the successful implementation of technological innovations.

This research showed that concerns about JR significantly affect the acceptance and use of AI tools. Organizations and managers should recognize this concern and mitigate employees' fear of AI replacing human jobs through education and training, emphasizing the role of AI tools as assistants to enhance work efficiency rather than replacements for humans. More importantly, managers can promote an innovative culture within the organization, always remaining attentive to employees' psychological changes. After all, the design field is also a highly competitive industry, and striving for performance in the market requires good technical and psychological qualities.

Limitations and Future Research

We point out four main limitations of our study. First, although our results support the proposed model on how design professionals use AI tools and the important role of JR in this, we need more research to confirm these findings. Second, our model of AI-induced JR does not consider cultural elements. We gathered data through a network survey from people familiar with AI tools used by design professionals in Confucian cultures and Chinese-speaking areas. It is unclear if this model and the survey questions work well for people in different regions. Third, the current study is cross-sectional, meaning its scope is limited because it only captures the thoughts and intended actions of design professionals at a single time. It is known from research that such perceptions and behavioral intentions can change, particularly with the rapid development of GenAI tools. Fourth, we did not make any conclusions or suggestions about learning. Future studies should investigate how teachers with design backgrounds are dealing with the quick arrival of GenAI tools in higher education and how these AI tools impact their teaching.

Acknowledgements

The research was supported by the National Science and Technology Council of Taiwan, and the Contract Numbers are NSTC 113-2410-H-224-009.

References

Aktan, M. E., Turhan, Z., & Dolu, I. (2022). Attitudes and perspectives towards the preferences for artificial intelligence in psychotherapy. *Computers in Human Behavior*, *133*, Article 107273. https://doi.org/10.1016/j.chb.2022.107273

Alpert, R., & Haber, R. N. (1960). Anxiety in academic achievement situations. *The Journal of Abnormal and Social Psychology*, *61*(2), 207–215. https://doi.org/10.1037/h0045464

Amabile, T. M. (1993). Motivational synergy: Toward new conceptualizations of intrinsic and extrinsic motivation in the workplace. *Human Resource Management Review*, *3*(3), 185–201. https://doi.org/10.1016/1053-4822(93)90012-S

Appleby, C. (2023, March 7). *Will colleges ban ChatGPT?* BestColleges. https://www.bestcolleges.com/news/will-colleges-ban-chatgpt/

Barbara, M. B. (1998). *Structural equation modeling with Lisrel, Prelis, and Simplis: Basic concepts, applications, and programming*. Lawrence Erlbaum Associates Publishers.

Briggs, N. E. (2006). *Estimation of the standard error and confidence interval of the indirect effect in multiple mediator models* [Doctoral dissertation, Ohio State University].

Chou, Y. H. D., Li, T. Y. D., & Ho, C. T. B. (2018). Factors influencing the adoption of mobile commerce in Taiwan. *International Journal of Mobile Communications*, *16*(2), 117–134. https://doi.org/10.1504/IJMC.2018.089754

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, *19*(2), 109–134. https://doi.org/10.1016/0092-6566(85)90023-6

Deci, E. L., Ryan, R. M., Gagné, M., Leone, D. R., Usunov, J., & Kornazheva, B. P. (2001). Need satisfaction, motivation, and well-being in the work organizations of a former Eastern Bloc country: A cross-cultural study of self-determination. *Personality and Social Psychology Bulletin*, *27*(8), 930–942. https://doi.org/10.1177/0146167201278002

Donnermann, M., Lein, M., Messingschlager, T., Riedmann, A., Schaper, P., Steinhaeusser, S., & Lugrin, B. (2021). Social robots and gamification for technology supported learning: An empirical study on engagement and motivation. *Computers in Human Behavior*, *121*, Article 106792. https://doi.org/10.1016/j.chb.2021.106792

Du, Y., Li, T., & Gao, C. (2023). Why do designers in various fields have different attitude and behavioral intention towards AI painting tools? An extended UTAUT model. *Procedia Computer Science*, *221*, 1519–1526. https://doi.org/10.1016/j.procs.2023.08.010

Duong, C. D., Vu, T. N., & Ngo, T. V. N. (2023). Applying a modified technology acceptance model to explain higher education students' usage of ChatGPT: A serial multiple mediation model with knowledge sharing as a moderator. *The International Journal of Management Education*, *21*(3), Article 100883. https://doi.org/10.1016/j.ijme.2023.100883

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., Carter, L., … Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, Article 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2019). Re-examining the unified theory of acceptance and use of technology (UTAUT): Towards a revised theoretical model. *Information Systems Frontiers*, *21*, 719–734. https://doi.org/10.1007/s10796-017-9774-y

Engel, J. F., Blackwell, R. D., & Miniard, P. W. (1995). *Consumer behavior* (8th ed.). Dryden Press.

Fan, W., Williams, C. M., & Wolters, C. A. (2012). Parental involvement in predicting school motivation: Similar and differential effects across ethnic groups. *The Journal of Educational Research*, *105*(1), 21–35. https://doi.org/10.1080/00220671.2010.515625

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Addison-Wesley.

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, *18*(3), 382–388. https://doi.org/10.1177/002224378101800313

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, *114*, 254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Gessl, A. S., Schlögl, S., & Mevenkamp, N. (2019). On the perceptions and acceptance of artificially intelligent robotics and the psychology of the future elderly. *Behaviour & Information Technology*, *38*(11), 1068–1087. https://doi.org/10.1080/0144929X.2019.1566499

Hair, J. F., Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.

Hars, A., & Ou, S. (2014). Working for free? Motivations for participating in open-source projects. *International Journal of Electronic Commerce*, *6*(3), 25–39. https://doi.org/10.1080/10864415.2002.11044241

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, C. L., & Haried, P. (2020). An evaluation of uncertainty and anticipatory anxiety impacts on technology use. *International Journal of Human–Computer Interaction*, *36*(7), 641–649. https://doi.org/10.1080/10447318.2019.1672410

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, *43*(6), 740–755. https://doi.org/10.1016/j.im.2006.05.003

Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*. Sage Publications. https://doi.org/10.4135/9781483385693

Korzyński, P., Mazurek, G., Altmann, A., Ejdys, J., Kazlauskaite, R., Paliszkiewicz, J., Wach, K., & Ziemba, E. (2023). Generative artificial intelligence as a new context for management theories: Analysis of ChatGPT. *Central European Management Journal*, *31*(1), 3–13. https://doi.org/10.1108/CEMJ-02-2023-0091

Lawler, E. E., & Porter, L. W. (1967). The effect of performance on job satisfaction. *Industrial Relations*, *7*(1), 20–28. https://doi.org/10.1111/j.1468-232X.1967.tb01060.x

Lin, C. P., & Bhattacherjee, A. (2008). Elucidating individual intention to use interactive information technologies: The role of network externalities. *International Journal of Electronic Commerce*, *13*(1), 85–108. https://doi.org/10.2753/JEC1086-4415130103

Maduku, D. K., Mpinganjira, M., Rana, N. P., Thusi, P., Ledikwe, A., & Mkhize, N. H. B. (2023). Assessing customer passion, commitment, and word-of-mouth intentions in digital assistant usage: The moderating role of technology anxiety. *Journal of Retailing and Consumer Services*, *71*, Article 103208. https://doi.org/10.1016/j.jretconser.2022.103208

Morosan, C., & DeFranco, A. (2016). It's about time: Revisiting UTAUT2 to examine consumers' intentions to use NFC mobile payments in hotels. *International Journal of Hospitality Management*, *53*, 17–29. https://doi.org/10.1016/j.ijhm.2015.11.003

Nikolopoulou, K., Gialamas, V., & Lavidas, K. (2021). Habit, hedonic motivation, performance expectancy and technological pedagogical knowledge affect teachers' intention to use mobile Internet. *Computers and Education Open*, *2*, Article 100041. https://doi.org/10.1016/j.caeo.2021.100041

Oliver, R. L. (1974). Expectancy theory predictions of salesmen's performance. *Journal of Marketing Research*, *11*(3), 243–253. https://doi.org/10.1177/002224377401100302

Overmier, J. B., & Lawry, J. A. (1979). Pavlovian conditioning and the mediator of behavior. *Psychology of Learning and Motivation*, *13*, 1–55. https://doi.org/10.1016/S0079-7421(08)60080-8

Piniel, K., & Csizér, K. (2013). L2 motivation, anxiety and self-efficacy: The interrelationship of individual variables in the secondary school context. *Studies in Second Language Learning and Teaching*, *3*(4), 523–550. https://doi.org/10.14746/ssllt.2013.3.4.5

Rahman, M. S., Sabbir, M. M., Zhang, J., Moral, I. H., & Hossain, G. M. S. (2022). Examining students' intention to use ChatGPT: Does trust matter? *Australasian Journal of Educational Technology*, *39*(6), 51–71. https://doi.org/10.14742/ajet.8956

Rotter, J. B., Chance, J. E., & Pharses, E. J. (1972). *Applications of a social learning theory of personality*. Holt Rinehart & Winston.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Shahsavar, Y., & Choudhury, A. (2023). User intentions to use ChatGPT for self-diagnosis and health-related purposes: Cross-sectional survey study. *JMIR Human Factors*, *10*, Article e47564. https://doi.org/10.2196/47564

Shek, D. T. L., & Yu, L. (2014). Confirmatory factor analysis using AMOS: A demonstration. *International Journal on Disability and Human Development*, *13*(2), 191–204. https://doi.org/10.1515/ijdhd-2014-0305

Teo, T., & Noyes, J. (2014). Explaining the intention to use technology among pre-service teachers: A multi-group analysis of the unified theory of acceptance and use of technology. *Interactive Learning Environments*, *22*(1), 51–66. https://doi.org/10.1080/10494820.2011.641674

Tyagi, P. K. (1985). Relative importance of key job dimensions and leadership behaviors in motivating salesperson work performance. *Journal of Marketing*, *49*(3), 76–86. https://doi.org/10.1177/002224298504900030

van Raaij, E. M., & Schepers, J. J. (2008). The acceptance and use of a virtual learning environment in China. *Computers & Education*, *50*(3), 838–852. https://doi.org/10.1016/j.compedu.2006.09.001

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157–178. https://doi.org/10.2307/41410412

Venkatesh, V., Thong, J., & Xu, X. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems*, *17*(5), 328–376. https://doi.org/10.17705/1jais.00428

Vogels, E. A. (2023, May 24). *A majority of Americans have heard of ChatGPT, but few have tried it themselves*. Pew Research Center. https://www.pewresearch.org/short-reads/2023/05/24/a-majority-of-americans-have-heard-of-chatgpt-but-few-have-tried-it-themselves/

Wang, Y. Y., & Wang, Y. S. (2022). Development and validation of an artificial intelligence anxiety scale: An initial application in predicting motivated learning behavior. *Interactive Learning Environments*, *30*(4), 619–634. https://doi.org/10.1080/10494820.2019.1674887

Wang, Y. M., Wei, C. L., Lin, H. H., Wang, S. C., & Wang, Y. S. (2022). What drives students' AI learning behavior: A perspective of AI anxiety. *Interactive Learning Environments*. Advance online publication. https://doi.org/10.1080/10494820.2022.2153147

Whittaker, T. A., & Schumacker, R. E. (2022). *A beginner's guide to structural equation modeling* (3rd ed.). Routledge. https://doi.org/10.4324/9781003044017

Williams, J., & MacKinnon, D. P. (2008). Resampling and distribution of the product methods for testing indirect effects in complex models. *Structural Equation Modeling*, *15*(1), 23–51. https://doi.org/10.1080/10705510701758166

Yoo, S. J., Han, S. H., & Huang, W. (2012). The roles of intrinsic motivators and extrinsic motivators in promoting e-learning in the workplace: A case from South Korea. *Computers in Human Behavior*, *28*(3), 942–950. https://doi.org/10.1016/j.chb.2011.12.015

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*, *2*, Article 100025. https://doi.org/10.1016/j.caeai.2021.100025

Zhao, Q., Chen, C. D., Cheng, H. W., & Wang, J. L. (2018). Determinants of live streamers' continuance broadcasting intentions on Twitch: A self-determination theory perspective. *Telematics and Informatics*, *35*(2), 406–420. https://doi.org/10.1016/j.tele.2017.12.018

Zhu, Q. (2020). Ethics, society, and technology: A Confucian role ethics perspective. *Technology in Society*, *63*, Article 101424. https://doi.org/10.1016/j.techsoc.2020.101424

Appendix

Measurement Items and Sources

Construct	Items	Scale reference
Intrinsic motivation (IM)	IM01: I find the actual process of using AI tools to be enjoyable.	Deci & Ryan (1985); Deci et al. (2001); Wang et al. (2022)
	IM02: Using AI tools enhances my personal development.	
	IM03: I find it interesting to use AI tools to solve my design task problems.	
	IM04: I believe using AI tools can be immensely beneficial to me.	
Extrinsic motivation (EM)	EM01: Using AI tools can improve my work performance.	Deci & Ryan (1985); Deci et al. (2001); Fan et al. (2012); Wang et al. (2022)
	EM02: Using AI tools helps enhance my design knowledge.	
	EM03: Using AI tools can assist in achieving higher income in the future.	
	EM05: Overall, I find AI tools to be very useful for my learning.	
Performance expectancy (PE)	PE01: Using AI tools has increased my learning efficiency.	Venkatesh et al. (2003); Venkatesh et al. (2012)
	PE02: AI tools can help me achieve my goals.	
	PE03: Using AI tools gives me more opportunities to gain knowledge and skills.	
	PE04: I find AI tools very useful in my daily life.	
Effort expectancy (EE)	EE01: I can easily become proficient in using AI tools.	Venkatesh et al. (2003); Venkatesh et al. (2012)
	EE02: I find it easy to use AI tools for knowledge management.	
	EE03: The user interface of AI tools is friendly.	
	EE04: Learning how to handle and operate AI tools is easy for me.	
Behavioral intention to use (BIU)	BI01: I am willing to recommend others to use AI tools.	Duong et al. (2023); Venkatesh et al. (2003); Venkatesh et al. (2012)
	BI02: I plan to use AI tools as learning tools.	
	BI03: I would be interested in participating in teaching activities involving AI tools.	

	BI04: I am interested in using AI tools more frequently in the future.	

Note. AI = artificial intelligence.

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

Video Lectures With AI-Generated Instructors: Low Video Engagement, Same Performance as Human Instructors

Selay Arkün-Kocadere[1] and Şeyma Çağlar-Özhan[2,*]
*[1]Department of Computer Education and Instructional Technology, Hacettepe University; [2]Department of Computer Technology and Information Systems, Bartın University; *Corresponding author*

Abstract

Via AI video generators, it is possible to create educational videos with humanistic instructors by simply providing a script. The characteristics of video types and features of instructors in videos impact video engagement and, consequently, performance. This study aimed to compare the impact of human instructors and AI-generated instructors in video lectures on video engagement and academic performance. Additionally, the study aimed to examine students' opinions on both types of videos. Convergent-parallel approach mixed method was used in this study. A total of 108 undergraduate students participated: 48 in the experimental group, 52 in the control group, and eight in the focus group. While the experimental group (AI-generated instructor) and control group (human instructor) watched 10 minutes of two videos each in two weeks, the students in the focus group watched both types of videos with human and AI-generated instructors. Data were collected through the Video Engagement Scale (VES) after the experimental process, and the Academic Performance Test as a pretest and posttest was administered in both groups. The findings of the experimental part revealed that learners' video engagement was higher in the course with the human instructor compared to the course with the AI-generated instructor. However, the instructor type did not have a significant effect on academic performance. The results based on the qualitative part showed that students thought the AI-generated instructor caused distraction, discomfort, and disconnectedness. However, when the video lesson topic was interesting or when students focused on the video with the intention of learning, these feelings could be ignored. In conclusion, even in today's conditions, there is no difference in performance between human and AI-generated instructors. As AI technology continues to develop, the difference in engagement is expected to disappear, and AI-generated instructors could be used effectively in video lectures.

Keywords: generative AI, human instructor, AI-generated instructor, video lecture, video engagement

Introduction

AI developments, whose place in the education system has been discussed and subject to research for some time, have vivified with generative AI. In the *EDUCAUSE Horizon Report* (Pelletier et al., 2023), generative AI was considered one of the top technologies that will shape the learning and teaching process. While AI can be defined as the simulation of human intelligence processes by computer systems, generative AI is defined as “the production of previously unseen synthetic content, in any form and to support any task, through generative modeling” by García-Peñalvo and Vázquez-Ingelmo (2023) as a result of their systematic mapping study. The form of this content can be visual, text, audio, or video. While there are concerns about the application of generative AI in education (Grassini, 2023; Lambert & Stevens, 2023), it has a huge potential waiting to be explored. When considered in the context of education, it is possible to produce instructional content with generative AI, give automated feedback, individualize the learning process, or support courses with conversational educational agents (Bozkurt, 2023; Lo, 2023; Pelletier et al., 2023). The focus of this study is on video lectures produced with generative AI.

Instructional videos, which could be used as part of larger lessons and either in online or face-to-face classrooms, include both visual and verbal material; the verbal part could be in the form of voice and/or text (Fiorella & Mayer, 2018). A real instructor or pedagogical agent often delivers the verbal part in instructional videos. For a while, massive open online courses (MOOCs) have been popular, and they mainly consist of video lectures; on the other hand, with the COVID-19 pandemic and recent learning movements such as hybrid or flipped classrooms, even face-to-face courses include video lectures. In this respect, it can be said that video lectures are frequently used for education, especially for online learning.

There are many challenges in the video lecture producing process with instructors (Crook & Schofield, 2017). AI video generators have started to be used as a solution. While AI video generators can be used to automate video editing, enhance video features such as resolution, perform automatic video summarization, generate subtitles for spoken content, and perform translation, they can also generate video from text/image and manipulate facial expressions, swap faces, and create deepfake videos. Through AI-generated video platforms, individuals with no video editing skills can create highly convincing synthetic video lectures, with AI-generated humanistic avatars and voices, simply by providing a script (Pellas, 2023). As expected, this is less time-consuming and costly than producing a video lecture through traditional means (Pellas, 2023). There are different types of video lecture productions with AI. Having an agent presenting slides with the real instructor’s face and voice is an interesting example of these productions (Dao et al., 2021).

It is known that the effect of video lecture types on video engagement differs (Chen & Thomas, 2020). Especially, video lectures in which the instructor is present have been the subject of research from several perspectives. Even the instructor’s presence in the video lecture is controversial (Alemdag, 2022). In a large-scale study in which 4,466 learners from 10 highly rated MOOCs participated, Hew (2018) found that having the instructor’s face in a video is preferred, but it is not necessary that it be there throughout the video. Recent studies have shown that instructor presence is critical in different video lecture types. For example, in a course that mainly includes declarative knowledge, performance was higher with instructor voice-over handbook videos as compared to videos showing the instructor and a whiteboard or slides, while in a procedural knowledge intensive course, performance was higher with videos that showed the instructor in front of a whiteboard (Urhan & Kocadere, 2024). Horovitz and Mayer (2021) studied how the virtual and human instructor being happy or bored in the video affects the learning process and outcomes. Although students perceived the emotional state of both instructors, the learning outcome did not differ. Learners responded to the emotional state of instructors in a similar way, but learners were found to be better at recognizing the human instructor's emotions.

Whether virtual or human, factors such as how much of the instructor's body is seen, the tone of their voice, displaying enthusiasm, showing empathy, and using non-verbal cues have been shown to have different effects on video engagement and learning (Dai et al., 2022; Verma et al., 2023). Even the context in which video is used has an impact on its effect; students in online courses engage with videos more than students in blended courses (Seo et al., 2021). With the perspective that teachers' physical characteristics are effective on students' learning and engagement, Daniels and Lee (2022) stated that AI can be used to create computer-generated teachers customized by race, gender, age, voice, language, and ethnicity when developing online courses.

In the research, AI-generated pedagogical agents in human form are described using terms such as AI-generated avatar, animated agent, virtual human, digital human, virtual tutor, and synthetic virtual instructor. However, not all of these terms are used in the same sense. For example, a virtual tutor can even be an animal cartoon character. Studies about AI-generated instructors that look and sound like humans are not very common in the literature. One example in this context is the study of Leiker et al. (2023). They examined the impact of using generative AI to create video lectures with synthetic virtual instructors. In the comparison, no difference was found between improvement and learning experience perception. As a result of the study, it was stated that AI-generated learning videos have the potential to replace those produced in the traditional method.

Vallis et al. (2023) used AI-generated avatars for presenting videos and online activities in a course. Through qualitative research, they examined students' perceptions of learning with AI-generated avatars. Although students drew attention to the avatar's less social and personal nature, they found the AI-generated avatar appropriate for lecture delivery.

Daniels and Lee (2022) examined the impact of avatar teachers on student learning and engagement via survey and interview. The research had mixed results; while some students found the avatar teacher engaging, others found it distracting. The importance of using appropriate teacher voice in the videos was expressed. However, some students stated that the teacher's physical characteristics had no effect as long as content was delivered well. In addition, some stated that the presence of a human teacher fosters a feeling of connection, establishing confidence in the information presented during the video lecture.

While the literature has not yet reached saturation even for human avatars (Beege et al., 2023), and it is not clear which agents with which characteristics should be included in videos for which conditions (Henderson & Schroeder, 2021), AI-based agents have come to the surface and are being used. This study aimed to analyse the difference between video lectures which include human and AI-generated instructors from the video engagement and academic performance perspective. It is possible to say that the higher the video engagement, the higher the success (Ozan & Ozarslan, 2016; Soffer & Cohen, 2019). For this reason, it is essential to produce videos with high engagement. Whether this is possible with AI-generated instructors is an issue that needs to be examined. From this point of view, we posed these research questions :

1. Does the type of instructor (AI-generated or human) affect the academic performance of participants?

2. Does the type of instructor (AI-generated or human) affect participants' engagement in video lectures?

3. What are students' views about AI-generated and human instructors in video lectures?

Method

Research Model

A convergent parallel approach was used in the study based on mixed methods. In this approach, quantitative and qualitative data are collected independently and combined in the interpretation phase to support the quantitative data and understand the subject more deeply (Creswell & Clark, 2011). In the quantitative dimension of this study, an experimental design was used to examine the effect of instructor types on the video engagement and academic performances of participants. In the qualitative dimension of the research, a focus group was formed to examine the factors that caused the results. The collected quantitative and qualitative data were analyzed and interpreted.

Study Group

The participants were undergraduate students in the 2nd and 3rd grade at the Department of Computer Technologies and Information Systems at a state university in Bartın, Turkey. Initially, there were 112 participants, 32 female and 80 male, all between the ages of 19 and 22. Participants had similar backgrounds and previously had not taken any courses on gamification, the subject of the video lecture. On the voluntary participation form signed by participants, it was verified whether they had taken a gamification course before. We then assigned 52 participants to a control group and 52 others to an experimental group. Four people in the experimental group were eventually excluded from the study because they could not complete the data collection process. Eight students participated in the focus group.

Procedure

Participants were randomly assigned to control and experimental groups according to whether their student numbers were odd or even. The control group watched video lectures that featured a human instructor while the experimental group had video lectures delivered by an AI-generated instructor. The methodological process of the research is summarized in Table 1.

Table 1

Summary of Research Method Comparing Student Responses to Human and AI-Generated Instructors in Video Lectures

Group	Participants	Pretest	Video lecture type	Posttest
Experimental	48 students randomly assigned	Academic performance	With AI-generated instructor	VES and academic performance
Control	52 students randomly assigned	Academic performance	With human instructor	VES and academic performance
Focus	8 volunteer students		With AI-generated instructor and with human instructor	Interview

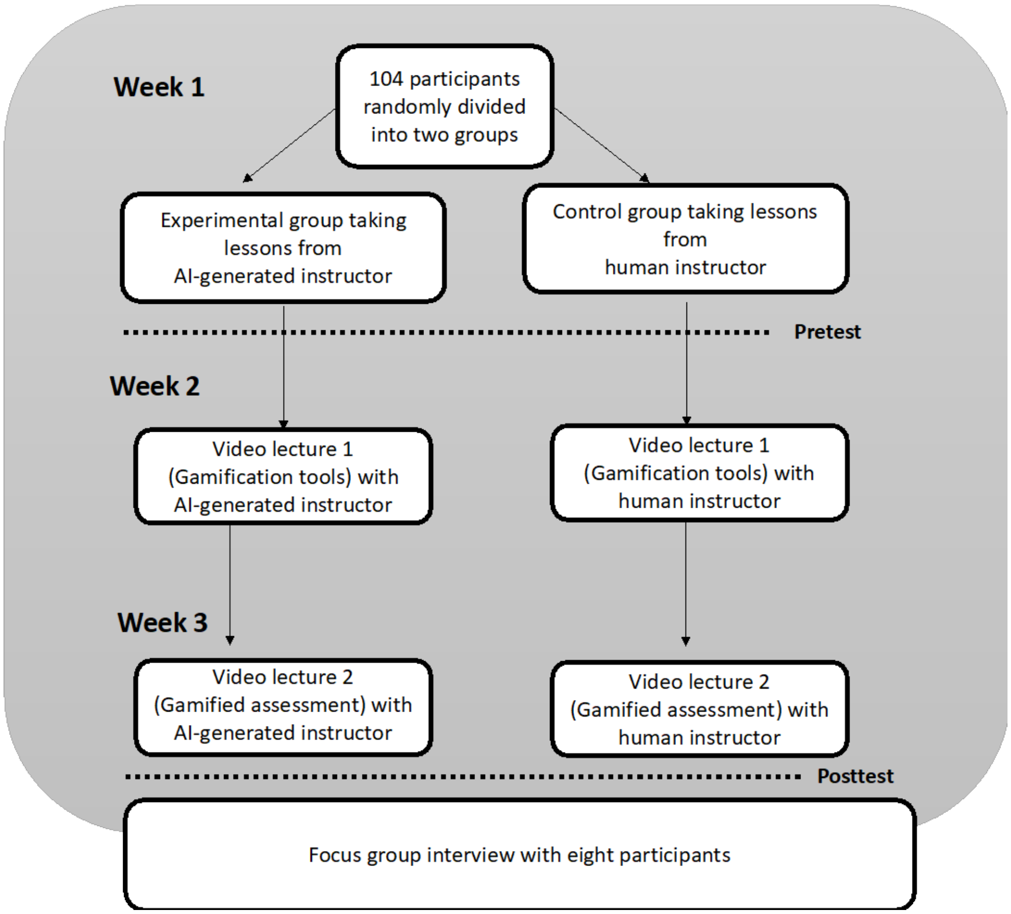
Note. VES = Video Engagement Scale. There was no pretest given to the focus group.

After different video types were developed, implementation began. The research included two different processes: the experimental process and the focus group interviews. The experimental process covered 3 weeks. In the first week, students were assigned to experimental and control groups. An academic performance test was applied to both groups. In the second and third weeks, videos featuring an AI-generated instructor were shown to participants in the experimental group, and videos featuring a human instructor were shown to the control group. Both groups watched the videos in the computer laboratory while wearing headphones. After completing the videos, the academic performance test and video engagement scale were administered to participants.

Then, a focus group consisting of eight people was formed. This focus group was shown both types of videos in the computer laboratory. At the end of the application, data regarding participants' opinions about the two different video types were collected. The implementation process is summarized in Figure 1.

Figure 1

Implementation Process of the Study Comparing Student Responses to Human and AI-Generated Instructors in Video Lectures

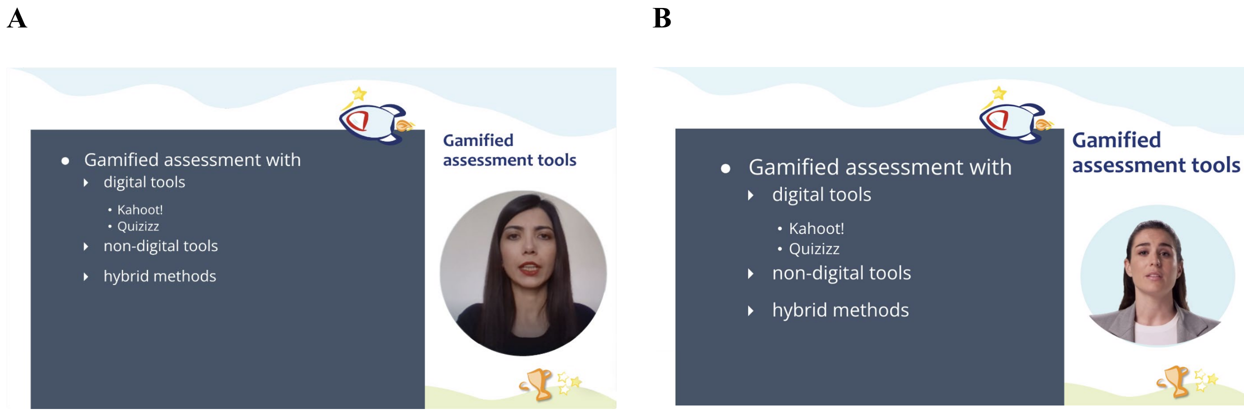


Videos Lectures

This study used two different videos explaining gamified assessment and gamification tools as the lecture material. Within the scope of the study, two types of videos were developed: one with a human instructor and the other with an AI-generated instructor. During the video development process for both types of video lectures, the text was first scripted and then the visuals and texts were combined in a presentation form. Then, an AI-based instructor video was added to the prepared presentation for the experimental group, and a human instructor video for the control group. The videos are of the talking-head type. Lessons were prepared in Turkish. The screenshots, shown in Figure 2, have been translated into English in this article for the purpose of reader comprehension.

Figure 2

Screenshots from the Control and Experimental Groups Video Lecture on Gamified Assessment Tools



Note. Panel A: Screenshot from the control group video with a human instructor. Panel B: Screenshot from the experimental group video with an AI-generated instructor.

The video with the AI-based instructor prepared for the experimental group was created with Vidext. Vidext is a generative AI tool that can produce videos in 27 languages from text by selecting different avatars (Ferre, 2023). The avatar chosen in video lectures created with Vidext exhibits various body movements during the lecture, just like real instructors. Their facial expressions differ, and they can display different postures (Ferre, 2023). Vidext is a commercial software, and only the trial version is free. It contains many design templates, and it is stated that it produces videos 10 times faster than other video generators (Marín, 2024).

In the control group, the prepared learning material was presented by a human instructor. The same script used in the experimental group was voiced, and the human instructor recorded the video. The human instructor is a lecturer from a different university who, therefore, is not known to the students who took part in the research. In summary, the same presentation and the same script were used in both video types. The instructors added to the presentations were similar in terms of their location on the screen and design appearance. Taking into consideration the studies in the literature (e.g., Hew, 2018; Manasrah et al., 2021; Ozan & Ozarslan, 2016), short videos (approximately 10 minutes) were prepared to engage the learners more easily.

Instruments

In this study, the Video Engagement Scale (VES) was administered to the participants after the experimental process, and data were collected by applying the Academic Performance Test before and after the experimental process. After this, a focus group interview was held with students to explore their opinions about the two different types of videos. Information about these data collection tools is explained under the subheadings that follow.

VES

Within the scope of this study, the Video Engagement Scale developed by Visser et al. (2016) and adapted into Turkish by Deryakulu et al. (2019) was used to determine the level of engagement of learners while watching the educational videos described in the video lecture section. This 7-point Likert-type scale consists of 15 questions. The lowest score that can be obtained from the scale is 15 and the highest is 105.

The scale has a 5-factor structure: (a) attention, (b) going into a narrative world, (c) identity, (d) empathy, and (e) emotion. In the original study group, the internal consistency coefficiency (Cronbach's alpha) for the factors was 0.57, 0.73, 0.87, 0.78, and 0.69, respectively. The scale's overall consistency was 0.90. The attention factor measures how much individuals focus on the video by disconnecting from the outside world. The going into a narrative world factor measures the extent to which one is immersed in the narrative world of the video. The identity factor measures the extent to which the identity of the character in the video, that is, the teacher, is adopted by the learners. The empathy factor measures the extent to which learners experience similar feelings to the teacher in the video. The emotion factor measures what kind of emotions the video lecture evokes in the individual. In this study, the Cronbach's alpha of the factors was 0.64, 0.68, 0.88, 0.89, and 0.81, respectively.

Academic Performance Test

As explained previously, the video lectures cover gamification tools and gamified assessment topics. In order to determine the performance of students in these subjects, we prepared an academic performance test that included 10 questions, five related to each topic in the video. The items were multiple-choice, and each question had four possible answers. The test was structured so that each correct answer equaled 1 point. The lowest score that could have been obtained from the test was 0, and the highest score was 10.

For the academic performance test, the researchers prepared a multiple-choice question pool containing 14 questions based on the video lectures' content. A subject matter expert was asked to evaluate the scope and structure of the draft academic performance test. Following expert opinion, four questions were removed, and expressions that could lead to misunderstandings were edited.

Semi-Structured Interview Form

As Roulston (2010) suggested, the main interview questions were limited in number, and the opinions of learners were detailed with follow-up questions. The form included three basic questions for each type of video:

1. What do you think about the video you watched?
2. How did the video you watched make you feel?
3. How did the video you watched affect your learning?

These were followed by another question to compare the types:

4. When you compare video lectures, what would you say?

Data Analysis

In this study, the academic performance pretest was first compared between the groups to evaluate whether the experimental and control groups had similar knowledge levels about gamification. According to the results of the independent samples *t*-test, there was no significant difference between the pretest academic performances of the groups ($t_{(98)}$ = 0.18, p > .05). Then, posttests for the variables were compared. Parametric tests were preferred since the data for the groups were normally distributed and homogeneous.

In this study, independent samples *t*-test and analysis of covariance (ANCOVA) were used to compare the video engagement and academic performance of the experimental and control groups. ANCOVA was applied to the data of the academic performance test used for the pretest and posttest. In this way, the effect of the pretest on the posttest was intended to be controlled (Büyüköztürk, 1998). The data collected with the VES used only for the posttest was analyzed with an independent samples *t*-test, and the differences between the groups were reported.

In the qualitative dimension of the research, participants were asked for their opinions about two different types of videos. Qualitative data were examined with content analysis. Content analysis focuses on making repeatable and valuable inferences from a text (Krippendorff, 2018). The analysis is carried out in four stages: (a) data coding, (b) finding themes, (c) organizing the codes and themes, and (d) interpreting the findings (Yıldırım & Şimşek, 2006). The data were analyzed and, independently, an external researcher also reviewed the figures. Reliability between coders was calculated with Cohen's kappa coefficient, and it was found that a significant level of agreement was reached (Cohen's κ = .61, p < .05). The results were then confirmed by asking additional questions to the focus group participants. To increase trustworthiness, data from interviews were reported with quotations and participant IDs.

Findings

Comparison of Findings on Video Engagement

The first research question was "Does the type of instructor (AI-generated/human) affect participants' engagement in the video lectures?" To answer the question, the groups' responses to the VES were analyzed with an independent sample *t*-test.

In Table 2, the averages in the experimental and control groups for the subdimensions of video engagement, attention, going into a narrative world, identity, empathy, and emotion are compared. In this context, a significant difference was found between the average of the experimental group (X_E = 11.13) and the average of the control group (X_C = 13.12) in the attention dimension ($t_{(98)}$ = 3.04, p < .05). A significant difference was also found between the average of the experimental group (X_E = 10.81) and the average of the control group (X_C = 12.27) in the going into a narrative world dimension ($t_{(98)}$ = 1.99, p < .05). A significant difference was found as well between the average of the experimental group (X_E = 7.60) and the average of the control group (X_C = 9.73) in the identity dimension ($t_{(98)}$ = 2.51, p < .05). For the empathy dimension, a significant difference was found between the experimental group (X_E = 7.50) and control group averages (X_C = 11.90), where $t_{(98)}$ = 5.22 and p < .05. Finally, for the emotion dimension, the difference between the experimental group (X_E = 9.73) and control group averages (X_C = 12.42) is statistically significant ($t_{(98)}$ = 3.89, $p < .05$). A significant difference was seen in all subdimensions, and it was found that the total video engagement score of the control group (X_C = 59.46) was significantly higher than the experimental group (X_E = 46.77), where $t_{(98)}$ = 3.89 and $p < .05$.

Table 2

Comparing Video Engagement Between Groups

VES subdimension	Group	*X*	*df*	*t*	*p*
Attention	Experimental	11.13	98	3.04	.03
	Control	13.12			
Going into a narrative world	Experimental	10.81	98	1.99	.049
	Control	12.27			
Identity	Experimental	7.60	98	2.51	.014
	Control	9.73			
Empathy	Experimental	7.50	98	5.22	.00
	Control	11.90			
Emotion	Experimental	9.73	98	3.30	.001
	Control	12.42			
Total (video engagement)	Experimental	46.77	98	3.89	.00
	Control	59.46			

Note. Experimental n = 48. Control n = 52. VES = video engagement scale.

Comparison of Findings on Academic Performance

The academic performance test scores of the groups were analyzed with ANCOVA to answer, "Does the type of instructor (AI-generated/human) affect the academic performance of the participants?" Table 3 shows the adjusted average academic performance scores. According to the analysis results displayed in Table 4, no significant difference was observed between the adjusted academic performance averages of the groups ($F_{(\text{group-error})}$ = 0.638, $p > .05$).

However, it was determined that the posttest average of the experimental group (X_E = 7.29) was higher than the pretest average (X_E = 5.53), and the posttest average of the control group (X_C = 7.35) was higher than the pretest average (X_C = 5.47). Students learned from both types of video lessons; there was no significant difference between them in terms of performance.

Table 3

Adjusted Average Academic Performance Scores for Control and Experimental Groups

Group	Pretest			Posttest		
	N	*X*	*SD*	*N*	*X*	*SD*
Experimental	48	5.53	1.92	48	7.29	1.82
Control	52	5.47	1.42	52	7.35	1.69

Table 4

Comparing Academic Performance Between Groups

Variance source	*SS*	*df*	*MS*	*F*	*p*
Pretest	75.053	1	75.053	31.892	.000
Group	1.501	1	1.501	0.638	.426
Error	225.922	96	2.353		
Total	5660	100			

Students' Views on the Different Types of Video Lectures

When the qualitative data were analyzed, it was seen that the opinions could be gathered under themes of distraction, discomfort, disconnectedness, and learning. The themes were named from the perspective of AI-generated instructor lectures, and while the first three are negative, learning is a theme that can be interpreted as positive. These are shown in Table 5 and discussed below.

Distraction

Six students expressed that the AI-generated instructor's speech and animation caused distraction due to low human-likeness. They mentioned that the voice was digital, there was a lack of emphasis, and that there were pronunciation mistakes in the speech. There were also complaints about the AI-generated instructor's animation, which included movements, gestures, and expressions. Participant 1 said, "The voice in the AI version was very digital, it sounded like I was listening to the woman in the navigation system. I thought it was too synthetic. It was hard for me to focus."

Participant 2 also found it hard to focus:

> It is more difficult to focus to the AI instructor, the voice always goes in the same rhythm. Besides, there were some pronunciation mistakes. That's why I couldn't pay full attention to the subject. Also, the teacher animation created with AI was distracting for me.

While two students mentioned their displeasure just about the speech, and four students mentioned similar issues with both the speech and animation, Participant 6 was not bothered by the speech but was unsatisfied with the animation of the AI-generated instructor.

> I was also distracted by the AI, if there was no animation of the AI-generated instructor, just the synthetic voice, it might not have felt so different from the human version. Even though the voice sounded human-like in the animated videos, the movements, gestures, and facial expressions on the screen felt unnatural. I think it would be more effective without the animation.

Discomfort

Two students mentioned a discomfort similar to the disturbance caused by the *uncanny valley effect* in the literature. One even used this expression directly. The uncanny valley describes a situation in which a machine appears to be almost human, yet not quite, which can cause a sense of unease (Cambridge Dictionary, n.d.). The AI-generated instructor was very similar to a human, but there was also an unfamiliarity, making students uncomfortable. For example, Participant 7 said, "The AI-instructor made me uncomfortable at first, but then I got used to it. ... Our brains tend to fear things that are very human-like but not a human, that might be the reason that I felt uncomfortable." Participant 8 added, "I had a bit of an uncanny valley effect, I felt uncomfortable, the AI instructor acts like a human, but she is not."

Disconnectedness

Although a recorded human video was used in the study, and the instructor in the videos was not recognized by the students, the AI-generated instructor was considered to be insincere, not having human characteristics, and, therefore, not giving a sense of personal connection. Participant 6 discussed this phenomenon.

> The real instructor gives a more intimate feeling, more personalized than the AI. It is not possible to get this personalization feeling in the AI version. When there is a real person in the video, even if it is a recording, I can empathize as if someone is talking to me considering my previous learning experiences.

Participant 8 agreed with this assessment.

> Seeing the teacher's face helped me to focus on the video. I followed her voice by watching her face. I know there is a real person behind her, it is easy to follow her. In AI, I got detached, there was no connection feeling.

Learning

In addition to the negative opinions about AI-generated videos, four students made positive comments about being able to learn from the AI instructor. Some students stated that they could learn when the subject was interesting enough or that they could learn if they focused on learning instead of AI-human instructor comparison. These views seem to be meaningful in explaining the significant improvement in the academic performance in both groups. Participant 7 said, "The subject of the video was very interesting for me. As such, my initial bad feeling about the AI-generated instructor disappeared, and I was able to watch the AI video and learn the topic."

Participant 8 expressed a similar sentiment:

> The AI instructor seemed synthetic, but the fact that I had this knowledge and that I had also watched the real human version may have been effective in making me feel this way. If I had only watched the AI version instead of watching both and making a comparison, I think I would have learned the subject easily because my focus would have been on learning.

Table 5

Participant Opinions About the AI-Generated Instructor

Participant	Negative theme				Positive theme
	Distraction		Discomfort	Disconnectedness	Learning
	Speech	Animation			
1	x				
2	x	x			x
3	x				
4	x	x			x
5	x	x			
6		x		x	
7	x	x	x		x
8			x	x	x
Total *n*	6	5	2	2	4
Total %	87.5		25	25	50

Students' opinions support the quantitative findings. The reason for the low video engagement in the AI-generated instructor could be the distraction, which is directly related to the attention dimension of the VES. In addition, it seems easier for students to engage in videos with the human instructor, since they feel more connected and the video is more personalized. In the disconnectedness theme, one student mentioned empath, which might indicate the empathy dimension of the VES. Another theme that emerged was related to students' uncomfortable feelings about the AI-generated instructor. The reason for this was expressed as both having an avatar very similar to a human and knowing that it was not human. On the other hand, when they were interested in the content or watched the video intending to learn, they did not have difficulty in learning. This is also a consistent finding with the academic performance test results.

Discussion

According to the experimental part of the study, the engagement of learners in the video course with the human instructor was higher than in the course with the AI-generated instructor. However, instructor type did not have a significant effect on academic performance. In the focus group interview, it was emphasized that the AI-generated virtual instructor made participants feel distracted, uncomfortable, and disconnected. In this context, like in the study of Vallis et al. (2023), the most frequently cited negative aspect of AI-generated instructors is that they cause distraction. It has been stated that the source of this situation is speech and animation. In fact, it is known that first impressions are more positive for a human than for a virtual instructor (Miller et al., 2023). It is thought that this problem can be solved with technological developments, thanks to synthesis engines that are getting better at creating reliable faces that don't cause uncanny valleys (Nightingale & Farid, 2022). On the other hand, the discomfort and disconnectedness felt in the video lecture with the AI-generated instructor might be corrected by improving the emotional characteristics of the avatar. Indeed, studies have found that adding emotional expressions to the virtual instructor positively affected empathy and the uncanny valley effect (Higgins et al., 2023). However, it was determined that these negative situations can be ignored by learners who find the video lecture's subject matter interesting or simply focus on learning.

The internal validity of the attention dimension of the VES used in this study is not at the desired level, which can be considered a limitation. However, this study indicated that the engagement in video lectures with the AI-generated virtual instructor is lower than that of the human instructor, and the main source of this is the synthetic voice, speech, and animation. For learners' engagement in a video lecture, they are expected to be able to focus, adopt the identity of the teacher, identify the teacher's emotions and feel similar feelings, and experience various emotions while being immersed in the narrative world of the lecture (Visser et al., 2016). In this context, although the AI-generated instructor used human-like gestures, her unnatural voice, speech, and gestures may have caused participants to have difficulty experiencing emotions, empathizing, adopting the instructor's identity, and focusing on the video. The close relationship of emotional factors with engagement has been shown in different studies. Pan et al. (2023) found that affective scaffolding predicted engagement positively. In a study examining the effect of the instructor's presence and absence on engagement in a video lecture, it was concluded that the instructor's presence gives an intimate and personal feel, which leads to a higher level of engagement (Guo et al., 2014). Similarly, a teacher's closeness in communication with a student in the classroom, that is, immediacy behavior and sincere behavior, have a high level of impact on engagement (Hu & Wang, 2023; Wang & Kruk, 2024). The results of these studies, that a human instructor increases engagement by providing an intimate feeling, seem consistent with our findings. In addition to these situations, participants stated that the learners who found the subject interesting could focus on the video without being affected by the instructor's type. According to Shoufan's (2019) study, the factors that affect university-level students' liking of an educational video are technical presentation, content, efficiency, the speaker's voice, and the interestingness of the video. Therefore, the content itself and subject being interesting may cause the instructor's type (AI-generated or human) to be ignored by learners.

Additionally, this study found that instructor type affected video engagement but did not affect academic performance. Leiker et al. (2023) compared those who watched a video lecture featuring an AI-generated virtual instructor with those who watched the traditionally produced instructor video and found that the learning of both groups improved, while no difference was found between the two groups in terms of learning. Studies investigating the effect of the human versus synthesized voice on academic performance in AI-based lectures are frequently encountered (e.g., Chiou et al., 2020). A systematic review conducted between 2010–2021 examining the impact of pedagogical agent features on academic performance and motivation found that the voice type did not affect academic performance (Dai et al., 2022). Another study found that the impact of using human voice and synthesized voice on retention did not differ (Davis et al., 2019). It is known that academic performance does not depend only on instructor characteristics. Within the scope of this study, the learning content prepared for both video lectures, the visuals used, and the additional resources recommended were all the same. Therefore, whether the channel transmitting the information was AI-generated or human, instructor type did not prevent performance improvement and did not cause a difference between the groups.

Conclusion

In this study, the effect of AI-generated virtual and human instructors in video lectures on university students' video engagement and academic performance was investigated. One hundred and twelve university students participated in the study conducted with mixed methods. In summary, this study found that engagement differs in favor of the human instructor in video lectures where an AI-generated instructor and a human instructor are used. Despite this, it was found that both types of instructors improved academic performance, and it was concluded that academic performance did not differ between groups. This shows that AI-generated video lectures, produced quickly and at a low cost, can be used instead of videos with human instructors provided certain conditions and improvements are met and made.

Implications and Recommendations

Many studies conducted on the multimedia principle to improve the conditions in question show that many features, such as instructors' presence in the video lecture, their position on the screen, gestures, tone of voice, and whether they are loved or familiar, are related to behavioral, cognitive, and emotional factors of learning. In this context, in AI-generated video lectures, which have an increasing interest in educational research, structural equation modeling studies can be carried out to see the effect of instructor characteristics on the students' learning from a holistic perspective. In particular, without forgetting the close relationship of engagement with learning, it can be investigated how affective factors such as empathy, emotion, and identity that the human instructor evokes in the learner can be achieved in video lectures, independent of the instructor's characteristics. In this context, AI-generated instructors can be added to video lectures based on emotional design and their effects can be discovered. Additionally, in this study, the importance of the instructor sharing a common denominator with the learner was ignored. Considering this, AI-generated instructors can be used in future studies and their effects on engagement can be examined. Finally, it should not be forgotten that content delivery is a very limited part of the learning process and interaction is the main element. Therefore, investigating the situations in which AI-generated instructors interact with students could be meaningful for future studies.

References

Alemdag, E. (2022). Effects of instructor-present videos on learning, cognitive load, motivation, and social presence: A meta-analysis. *Education and Information Technologies, 27*(9), 12713–12742. https://doi.org/10.1007/s10639-022-11154-w

Beege, M., Schroeder, N. L., Heidig, S., Rey, G. D., & Schneider, S. (2023). The instructor presence effect and its moderators in instructional video: A series of meta-analyses. *Educational Research Review*, Article 100564. https://doi.org/10.1016/j.edurev.2023.100564

Bozkurt, A. (2023). Unleashing the potential of generative AI, conversational agents and chatbots in educational praxis: A systematic review and bibliometric analysis of GenAI in education. *Open Praxis, 15*(4), 261–270. https://doi.org/10.55982/openpraxis.15.4.609

Büyüköztürk, Ş. (1998). Analysis of covariance (A comparative study with analysis of variance). *Ankara University Journal of Faculty of Educational Sciences (JFES)*, *31*(1), 93–100.

Cambridge Dictionary. (n.d.). Uncanny valley. In Cambridge online dictionary. Retrieved May 3, 2024, from https://dictionary.cambridge.org/dictionary/english/uncanny-valley.

Chen, H.-T. M., & Thomas, M. (2020). Effects of lecture video styles on engagement and learning. *Educational Technology Research and Development*, *68*(5), 2147–2164. https://eric.ed.gov/?id=EJ1269149

Chiou, E. K., Schroeder, N. L., & Craig, S. D. (2020). How we trust, perceive, and learn from virtual humans: The influence of voice quality. *Computers & Education*, *146*, Article 103756. https://doi.org/10.1016/j.compedu.2019.103756

Creswell, J. W., & Clark, V. P. (2011). *Mixed methods research*. SAGE Publications.

Crook, C., & Schofield, L. (2017). The video lecture. *The Internet and Higher Education*, *34*, 56–64. https://doi.org/10.1016/j.iheduc.2017.05.003

Dai, L., Jung, M. M., Postma, M., & Louwerse, M. M. (2022). A systematic review of pedagogical agent research: Similarities, differences and unexplored aspects. *Computers & Education*, Article 104607. https://doi.org/10.1016/j.compedu.2022.104607

Daniels, D., & Lee, J. S. (2022). The impact of avatar teachers on student learning and engagement in a virtual learning environment for online STEM courses. In P. Zaphiris & A. Ioannou (Eds.), *Learning and collaboration technologies. Novel technological environments. HCII 2022. Lecture Notes in Computer Science, Vol. 13329* (pp. 158–175). Springer International Publishing. https://doi.org/10.1007/978-3-031-05675-8_13

Dao, X.-Q., Le, N.-B., & Nguyen, T.-M.-T. (2021). AI-powered MOOCs: Video lecture generation. In X. Jiang (Chair), *IVSP '21: Proceedings of the 2021 3rd International Conference on Image, Video* *and Signal Processing* (pp. 95–102). Association for Computing Machinery. https://doi.org/10.1145/3459212.3459227

Davis, R. O., Vincent, J., & Park, T. (2019). Reconsidering the voice principle with non-native language speakers. *Computers & Education*, *140*, Article 103605. https://doi.org/10.1016/j.compedu.2019.103605

Deryakulu, D., Sancar, R., & Ursavaş, Ö. F. (2019). Adaptation, validity and reliability study of the video engagement scale. *The Journal of Educational Technology Theory and Practice, 9*(1), 154–168. https://doi.org/10.17943/etku.439097

Ferre, A. G. (2023, November 27). Videos con inteligencia artificial: mejores herramientas para impactor. [Videos with artificial intelligence: better tools to impact]. *Vidext Technologies*. https://blog.vidext.io/videos-con-inteligencia-artificial-mejores-herramientas-para-impactar

Fiorella, L., & Mayer, R. E. (2018). What works and doesn't work with instructional video. *Computers in Human Behavior*, *89*, 465–470. https://doi.org/10.1016/j.chb.2018.07.015

García-Peñalvo, F. J., & Vázquez-Ingelmo, A. (2023). What do we mean by GenAI? A systematic mapping of the evolution, trends, and techniques involved in Generative AI. *International Journal of Interactive Multimedia and Artificial Intelligence, 8*(4), 7–16. DOI:10.9781/ijimai.2023.07.006

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, *13*(7), Article 692. https://doi.org/10.3390/educsci13070692

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In M. Sahami (Chair), *L@S '14: Proceedings of the first ACM Conference on Learning@Scale* (pp. 41–50). Association for Computing Machinery. https://doi.org/10.1145/2556325.2566239

Henderson, M. L., & Schroeder, N. L. (2021). A systematic review of instructor presence in instructional videos: Effects on learning and affect. *Computers and Education Open*, *2*, Article 100059. https://doi.org/10.1016/j.caeo.2021.100059

Hew, K. F. (2018). Unpacking the strategies of ten highly rated MOOCs: Implications for engaging students in large online courses. *Teachers College Record*, *120*(1), 1–40. https://eric.ed.gov/?id=EJ1162815

Higgins, D., Zhan, Y., Cowan, B. R., & McDonnell, R. (2023, August). Investigating the effect of visual realism on empathic responses to emotionally expressive virtual humans. In A. Chapiro, A. Robb, F. Durupinar, Q. Sun, & L. Buck (Eds.), *SAP '23: ACM Symposium on Applied Perception 2023* (pp. 1–7). Association for Computing Machinery. https://doi.org/10.1145/3605495.3605799

Horovitz, T., & Mayer, R. E. (2021). Learning with human and virtual instructors who display happy or bored emotions in video lectures. *Computers in Human Behavior*, *119*, Article 106724. https://doi.org/10.1016/j.chb.2021.106724

Hu, L., & Wang, Y. (2023). The predicting role of EFL teachers' immediacy behaviors in students' willingness to communicate and academic engagement. *BMC Psychology*, *11*, Article 318. https://doi.org/10.1186/s40359-023-01378-x

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage Publications.

Lambert, J., & Stevens, M. (2023, September 9). ChatGPT and generative AI technology: A mixed bag of concerns and new opportunities. *Computers in the Schools,* 1–25. https://doi.org/10.1080/07380569.2023.2256710

Leiker, D., Gyllen, A. R., Eldesouky, I., & Cukurova, M. (2023). Generative AI for learning: Investigating the potential of learning videos with synthetic virtual instructors. In N. Wang, G. Rebolledo-Mendez, N. Matsuda, O. C. Santos, & V. Dimitrova (Eds.), *Artificial Intelligence in Education: 24th International Conference proceedings. Posters and late breaking results, workshops and tutorials, industry and innovation tracks, practitioners, doctoral consortium and blue sky.* (pp. 523–529). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-36336-8_81

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*(4), Article 410. https://doi.org/10.3390/educsci13040410

Manasrah, A., Masoud, M., & Jaradat, Y. (2021). Short videos, or long videos? A study on the ideal video length in online learning. In K. M. Jaber (Ed.), *2021 International Conference on Information Technology (ICIT)* (pp. 366–370). IEEE. https://www.doi.org/10.1109/ICIT52682.2021.9491115

Marín Orozco, D. (2024). *Aplicación de inteligencias artificiales en los procesos de diseño y creación en la malla curricular del programa de Diseño Visual* [Application of artificial intelligence in design and creation processes in the curriculum of the visual design program] [Doctoral dissertation, University of Caldas]. Universidad de Caldas Repositorio Institucional. https://repositorio.ucaldas.edu.co/handle/ucaldas/19777

Miller, E. J., Foo, Y. Z., Mewton, P., & Dawel, A. (2023). How do people respond to computer-generated versus human faces? A systematic review and meta-analyses. *Computers in Human Behavior Reports*, Article 100283. https://doi.org/10.1016/j.chbr.2023.100283

Nightingale, S. J., & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. *Proceedings of the National Academy of Sciences*, *119*(8), Article e2120481119. https://doi.org/10.1073/pnas.2120481119

Ozan, O., & Ozarslan, Y. (2016). Video lecture watching behaviors of learners in online courses. *Educational Media International*, *53*(1), 27–41. https://doi.org/10.1080/09523987.2016.1189255

Pan, Z., Wang, Y., & Derakhshan, A. (2023). Unpacking Chinese EFL students' academic engagement and psychological well-being: The roles of language teachers' affective scaffolding. *Journal of Psycholinguistic Research*, *52*(5), 1799–1819. https://doi.org/10.1007/s10936-023-09974-z

Pellas, N. (2023). The influence of sociodemographic factors on students' attitudes toward AI-generated video content creation. *Smart Learning Environments*, *10*(1), Article 57. https://doi.org/10.1186/s40561-023-00276-4

Pelletier, K., Robert, J., Muscanell, N., McCormack, M., Reeves, J., Arbino, N., & Grajek, S. (with Birdwell, T., Liu, D., Mandernach, J., Moore, A., Porcaro, A., Rutledge, R., & Zimmern, J.). (2023). *EDUCAUSE horizon report, teaching and learning edition*. EDUCAUSE. https://library.educause.edu/resources/2023/5/2023-educause-horizon-report-teaching-and-learning-edition

Roulston, K. (2010). *Reflective interviewing: A guide to theory and practice*. Sage Publications. https://doi.org/10.4135/9781446288009

Seo, K., Dodson, S., Harandi, N. M., Roberson, N., Fels, S., & Roll, I. (2021). Active learning with online video: The impact of learning context on engagement. *Computers & Education*, *165*, Article 104132. https://doi.org/10.1016/j.compedu.2021.104132

Shoufan, A. (2019). What motivates university students to like or dislike an educational online video? A sentimental framework. *Computers & Education*, *134*, 132–144. https://doi.org/10.1016/j.compedu.2019.02.008

Soffer, T., & Cohen, A. (2019). Students' engagement characteristics predict success and completion of online courses. *Journal of Computer Assisted Learning*, *35*(3), 378–389. https://doi.org/10.1111/jcal.12340

Urhan, S., & Kocadere, S. A. (2024). The effect of video lecture types on the computational problem-solving performances of students. *Educational Technology & Society, 27*(1), 117–133. https://doi.org/10.30191/ETS.202401_27(1).RP08

Vallis, C., Wilson, S., Gozman, D., & Buchanan, J. (2023, June 6). Student perceptions of AI-generated avatars in teaching business ethics: We might not be impressed. *Postdigital Science and Education*, 1–19. https://doi.org/10.1007/s42438-023-00407-7

Verma, N., Getenet, S., Dann, C., & Shaik, T. (2023). Characteristics of engaging teaching videos in higher education: A systematic literature review of teachers' behaviours and movements in video conferencing. *Research and Practice in Technology Enhanced Learning*, *18*, Article 040. https://doi.org/10.58459/rptel.2023.18040

Visser, L. N. C., Hillen, M. A., Verdam, M. G. E., Bol, N., de Haes, H. C. J. M., & Smets, E. M. A. (2016). Assessing engagement while viewing video vignettes; Validation of the Video Engagement Scale

(VES). *Patient Education and Counseling*, *99*(2), 227–235. https://doi.org/10.1016/j.pec.2015.08.029

Wang, Y., & Kruk, M. (2024). Modeling the interaction between teacher credibility, teacher confirmation, and English major students' academic engagement: A sequential mixed-methods approach. *Studies in Second Language Learning and Teaching*, 1–31. https://doi.org/10.14746/ssllt.38418

Yıldırım, A., & Şimşek, H. (2006). *Qualitative research methods in social sciences*. Seçkin Press.

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

The Use of Deep Learning in Open Learning: A Systematic Review (2019 to 2023)

Odiel Estrada-Molina[1], Juanjo Mena[2], and Alexander López-Padrón[3]
[1]University of Valladolid, [2]University of Salamanca, [3]Technical University of Manabí

Abstract

No records of systematic reviews focused on deep learning in open learning have been found, although there has been some focus on other areas of machine learning. Through a systematic review, this study aimed to determine the trends, applied computational techniques, and areas of educational use of deep learning in open learning. The PRISMA protocol was used, and the Web of Science Core Collection (2019–2023) was searched. VOSviewer was used for networking and clustering, and in-depth analysis was employed to answer the research questions. Among the main results, it is worth noting that the scientific literature has focused on the following areas: (a) predicting student dropout, (b) automatic grading of short answers, and (c) recommending MOOC courses. It was concluded that pedagogical challenges have included the effective personalization of content for different learning styles and the need to address possible inherent biases in the datasets (e.g., socio-demographics, traces, competencies, learning objectives) used for training. Regarding deep learning, we observed an increase in the use of pre-trained models, the development of more efficient architectures, and the growing use of interpretability techniques. Technological challenges related to the use of large datasets, intensive computation, interpretability, knowledge transfer, ethics and bias, security, and cost of implementation were also evident.

Keywords: open learning, deep learning, MOOC, systematic review

The Use of Deep Learning in Open Learning: A Systematic Review (2019 to 2023)

Conceptions about open learning are vast and wide. In a general sense, it refers to educational modalities to broaden education and training by breaking the barriers of time and space. Open learning has provided both students and teachers with greater flexibility, professional development, productivity, culture and socialisation in learning communities (Tzeng et al., 2022; Uddin et al., 2021). This pedagogical conception has been complemented by distributed learning in which teachers, students, and the content to be taught and learned are not centralised but can occur at any time and place (Zakaria et al., 2022; Zheng et al., 2022). In both open and distributed learning, the teacher is a facilitator of learning, the learner assumes an increasingly active and interactive role, and teaching and learning are reinforced and mediated by the use of digital technologies promoting open, online, and ubiquitous distributed learning environments (Mardini et al., 2023; Zheng et al., 2022).

Distance education has gained momentum with the expansion of open educational offerings and online vocational training. This has led to the enrolment of large numbers of students and has reinforced and integrated the use of information and communication technology (ICT). Therefore, the generation of new models and patterns of teaching and learning has been closely linked to new ICT trends creating the inter-, multi-, and trans-disciplinary space of emerging technologies (Mrhar et al., 2021).

Over the last 10 years, common open learning tools have included open educational resources (OER), collaborative teaching platforms, virtual learning environments, and educational social networks. With the evolution of open and distributed learning, distributed learning ecosystems (DLE) have emerged. These ecosystems utilize distributed learning infrastructures that bring together various tools and technologies related to OER, services, resources, and open learning environments (Otto et al., 2023). From an educational and open learning point of view, DLEs have focused on the diversity and interactions of actors and (re)use activities, allowing the creation of solutions such as resource aggregation mechanisms and open learning repositories. In this context, open pedagogy, as well as software advances based on artificial intelligence (AI) and the standardisation of OER metadata have all developed.

DLEs have promoted the effectiveness of open and online learning, although highly dependent on platforms, the Internet, interaction and interactivity, as well as teacher and learner empowerment. Significant progress has been made in the design and development of DLEs; however, there is still a latent lack of theoretical and empirical analysis of how emerging technologies such as AI, virtual reality, and augmented reality have influenced open learning (Otto et al., 2023).

AI has been grounded in various disciplines, such as natural language processing (NLP), artificial neural networks, computer vision, robotics, knowledge engineering, and machine learning (ML), among others (Hassan et al., 2019). The development of DLE has enabled the use of artificial intelligence in education (AIEd), notably, as expressed by Chen, Feng, et al. (2020), in adapting content, designing virtual tutors, automated assessment, data analysis, the use of virtual and augmented reality, creating recommender systems, and developing specific skills. These have all sought to improve the accessibility, personalisation, and efficiency of learning (Alruwais, 2023; Goel & Goyal, 2020).

AIEd can be learner-centred (e.g., adaptive or personalised learning management systems), teacher-centred (e.g., automating tasks such as administration, assessment, learning progress, and detecting plagiarism) or system-centred, providing administrators with decision-making information related to monitoring dropout patterns Chen, Feng, et al. (2020).

ML is one of the most widely used disciplines of AI. It enables, among other applications, the design of intelligent tutoring systems and performance prediction. However, in education, unstructured data such as images, text, and voice have often been handled (An et al., 2019). Conventional ML models may not be as effective in extracting useful features from these types of data, and therefore, it has been necessary to use more powerful models such as deep learning (DL). DL, a subset of ML, refers to the use of deep neural networks, configured with multiple successive layers of neurons (LeCun et al., 2015). It has represented the most advanced machine learning technique for solving problems with large sets of structured training data (Chassagnon et al., 2019) such as the analysis of traces and data from massive open online courses (MOOCs) for predicting school performance. While DL can address some of the limitations of ML, it is not a universal solution and also has its considerations, such as the need for large amounts of training data and computational resources. The choice between ML and DL depends on the specific nature of the problem and the data available in the educational context (El-Rashidy et al., 2023).

Recent studies on AI in open learning have focused on providing a learning experience for each learner by influencing motivation and online participation (Salas-Rueda, 2023). In a general sense, these systems must ensure the ability to provide feedback and structure adaptive learning content according to the individual capabilities of each learner. The usefulness of these tools depend on the design and development of more efficient intelligent tutoring systems, as shown in the teaching of mathematics, languages, and programming (Liang et al., 2023; Su et al., 2023).

In AI, DL as a subset of ML is based on the use of artificial neural networks (ANN), whose typology can be convolutional, recurrent, generative adversarial, deep, or modular neural networks. This area of AI has made inroads in education, mainly in analysing learning interactions in MOOCs, determining the chronological sequence of each student's interactions, predicting academic dropouts, and designing new and more efficient learning courses based on user experience, learning habits, and interactions (Verma et al., 2023). The most widely used technological models have been long short-term memory (LSTM) algorithms, sequential interaction rule mining process, and temporal interaction analysis (Yu et al., 2021).

Another recent application of DL has been automated online discussion message categorization based on convolutional neural network (CNN) and random forest classifiers. This advancement allowed for the analysis of interactions in online and open learning contexts. Utilizing the community of inquiry (CoI) framework, this application of DL has delineated the dimensions of cognitive presence (e.g., knowledge (re)construction, problem-solving), social presence (e.g., social interactions), and teaching presence (e.g., course design, interaction, interactivity) (Hu et al., 2021).

To enhance learning outcomes, DL has facilitated the creation of adaptive e-learning systems that analyze the behavior of individual learners in their interactions with learning objectives. Additionally, deep autoencoders have been utilized to learn and predict learner behaviors. DL has also supported the development of video analysis classification systems aimed at generating engaging video learning reports (Verma et al., 2023).

While there have been numerous studies on DL applications in education, few have focused on open learning. Consequently, there has been a lack of systematic reviews analyzing these applications. This research addresses this gap by examining the scientific literature.

Gaps in the Analysis of Studies on DL in Open Learning

Table 1 shows that some reviews, mappings, and bibliometric studies on AIEd have been published in Scopus and the Web of Science (WoS). Only three of these explicitly included DL studies, (Chen, Xie, et al. (2020); Pan et al., 2023; Vanitha & Jayashree, 2023). The remainder focused on other branches of AI.

Pan et al. (2023) performed a generic mapping of the use of DL in education, Chen, Feng, et al. (2020) analysed common errors in terminologies and the semantic forest of AIEd, and Vanitha and Jayashree (2023) focused on educational time series. However, none of these discussed the use of DL in open learning per se.

Table 1

Systematic Reviews

Research study	Period	Database	Number of studies
Pan et al. (2023)	1992 to 2002	WoS (SSCI)	2,827
Crompton and Burke (2023)	2016 to 2022	EBSCO, Wiley Online Library, JSTOR, Science Direct, and WoS	138
Su et al. (2023)	2016 to 2022	WoS, BSCO, IEEE, ACM, Scopus, and Google Scholar	16
Koong Lin et al. (2023)	2018 to 2023	Scopus	217
Vanitha and Jayashree (2023)	2018 to 2022	Google Scholar and IEEE Xplore	22
Liang et al. (2023)	1990 to 2020	WoS	16
Alhothali et al. (2022)	2017 to 2021	Scopus, Web of Science, Springer, IEEE, Elsevier, and Sage	67
Shafiq et al. (2022)	2017 to 2021	Google Scholar, IEEE Xplorer, ScienceDirect, Springer, ResearchGate, MDPI, Taylor & Francis, ACM Library, Emerald Insight, IOPscience, and Wiley	75
Hwang et al. (2021)	1996 to 2020	WoS (SSCI)	43
Uddin et al. (2021)	2013 to 2021	IEEE Xplore, ACM Digital Library, Science Direct, and Google Scholar	116

Talan (2021)	2001 to 2021	WoS	2,686
Chen, Xie, et al. (2020)	1970 to 2019	WoS (SSCI)	45
Chen, Xie, and Hwang (2020)	1999 to 2019	WoS	9,560

Several points are worth highlighting regarding these studies. Crompton and Burke (2023) discussed how, in higher education, AIEd contributed to learning assessment and prediction from AI assistants, intelligent tutoring systems, and learning management. Along these lines, others have focused on (a) algorithms and systems for learning prediction and student retention (Alhothali et al., 2022; Shafiq et al., 2022); (b) MOOC recommender systems (Uddin et al., 2021); and (c) analysing the impact of deep learning on educational time series datasets (Vanitha & Jayashree, 2023).

Su et al. (2023) focused on artificial intelligence digital literacy and AIEd challenges in the context of K–12 to higher education, while Koong Lin et al. (2023) focused on the unique use of ChatGPT in education.

In their systematic review, Liang et al. (2023) analysed research methods, and the role of AI in language teaching and its learning outcomes. Similarly, other mapping studies and reviews have explained the applications of engineering and computational techniques at certain levels of education, such as higher education (Hwang et al., 2021) and early childhood education (Su et al., 2023).

Although several authors (Chen, Xie, et al., 2020; Pan et al., 2023; Talan, 2021) conducted comprehensive analyses of more than 900 articles, they focused on scholarly output and its bibliometric analysis only, without delving into the advantages of AI in open learning.

In their extensive systematic review of influential AIEd studies, Chen, Xie, et al. (2020) stated that only two studies explicitly identified the use of DL, referring to the study of advanced neural network architecture and the achievements of optimising study strategies from parallel robot instruction. In their study, they concluded that this was understandable, as DL was a newer area of research compared to general AI and machine learning.

In general, the reviews argued that these AIEd studies focused on profiling and dropout prediction, content evaluation, adaptive system design, and intelligent tutoring systems; there is still a lack of studies that systematised the use of DL. This theoretical research reaffirmed the importance of analysing empirical studies on the unique use of DL and its relationship with open learning, as MOOC dropout prediction and course recommendation require the use of powerful computational models (Wang et al., 2024).

To fill this gap, this study analysed AIEd-related scientific articles focusing on DL and open learning published between 2019 and 2023 to explore the important questions that remain to be investigated.

Objectives and Research Questions

This study aimed to identify, through a systematic review, the trends, applied computational techniques, and areas of educational use of deep learning in open learning. For this purpose, we examined the scientific literature from 2019 to 2023, present in one of the main bibliographic reference database collections, namely the Web of Science (WoS).

Aligned with this objective and addressing theoretical gaps, we sought to answer the following questions:

> How does scientific collaboration relate to the study of DL in open learning, highlighting its application areas? (This question was addressed through the bibliometric dimension, specifically co-authorship networks, keyword networks, and main clusters.)
>
> What are the dependent variables studied and their main findings? (This was answered through the pedagogical dimension by analyzing the content of each study.)
>
> What are the DL techniques or algorithms used in open learning (independent variable), and what is their level of accuracy? (This was addressed through the technological dimension, identifying, for each study, DL techniques or algorithms, data sources used, and levels of accuracy.)

Method

We utilized the updated PRISMA statement guidelines to search for and select scientific information (Page et al., 2021). A quantitative procedure was employed for coding to ensure the validity of the study (Zawacki-Ritcher et al., 2020).

Search Strategy and Criteria

Only articles from peer-reviewed journals were selected to ensure a high level of quality. The search parameters were narrowed to the period from 2019 to 2023 to ensure the currency of the literature, which is essential in the AIEd area. Mendeley was used to extract articles and eliminate duplicates.

The electronic search protocol included the WoS databases, namely the Social Sciences Citation Index (SSCI) and the Science Citation Index Expanded (SCIE), housing more than 12,000 journals. A full-text search was conducted in line with the research objectives and questions. The Boolean search included terms related to DL, open learning, and distributed learning.

Since open education, distance education, online education, and distributed learning are related but distinct, they were included as keywords to verify later whether each result was related to open education. Similarly, DL and ML are different, but as some authors mention them interchangeably (Chen, Xie, et al. (2020) they were used as keywords, and subsequently, only works referring to DL were included in the analysis. The techniques used were checked to identify whether they referred to ML (e.g., supervised learning, unsupervised learning, and reinforcement learning) or DL (e.g., convolutional, recurrent, generative adversarial, deep, or modular artificial neural networks).

The initial Boolean search string focused on "deep learning OR DL OR machine learning OR ML," which yielded generic AIEd results. Subsequently, these were filtered according to the search string "open learning OR OL OR distributed learning OR DL" to obtain documents related to the research topic. In summary, the final search string focused on "(deep learning OR machine learning) AND (open education OR distance education OR online education OR distributed learning) AND (educational technologies OR artificial intelligence in education)."

The inclusion and exclusion criteria are detailed in Table 2.

Table 2

Inclusion and Exclusion Criteria

Inclusion criteria	Exclusion criteria
Published in the period 2019 to 2023	Published before 2019
English language	Languages other than English
Original articles	Purely theoretical studies, systematic reviews,
Empirical studies on DL use in open learning	conference proceedings, and editorial letters,
	among others of equal magnitude

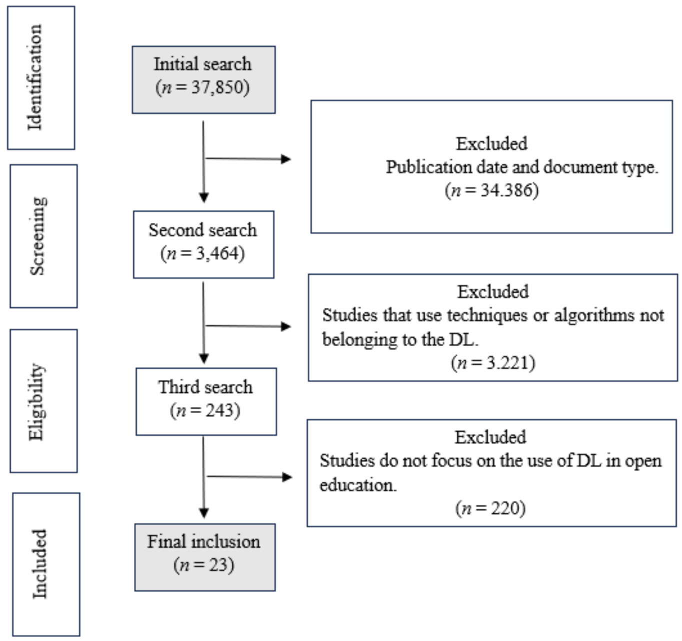
Screening and Validation Strategy

After carrying out the search of papers, two independent researchers used a standard checklist form to exclude irrelevant articles and determine their eligibility. In the process, any discrepancies between reviewers were resolved.

Subsequently, bibliographic data were extracted, and key findings and results were synthesised and recorded. The PRISMA diagram and its checklist for determining study quality were used to assess and carry out the selection process (Page et al., 2021). Finally, the two independent researchers read the selected articles to extract relevant information and answer the research questions. Any inconsistencies between the two researchers' results were resolved by a third reviewer. The process revealed a few articles on the use of DL in open learning (Figure 1

Figure 1

Selecting Studies (PRISMA Protocol)



Coding and Visualisation Tools

The selected studies were coded for deductive aspects and internal validity with a set of three criteria: keywords, authors, and first authors' countries. Concerning inductive coding for conclusion validity, the focus was on the influence of DL on open learning to identify trends. Finally, grounded coding for external validity focused the findings on how DL was used rather than how it could be used. VOSviewer version 1.6.19 was used for bibliometric data visualisation and analysis.

Results

Bibliometric Output Information

In the literature analysis, 23 articles were finally selected (Figure 2) of which two were highly cited (Onan, 2021; Xing & Du, 2019). The analysis showed a trend of three to six manuscripts published annually. The low scientific output regarding the use of DL in open learning was highlighted.

Figure 2

Scientific Production

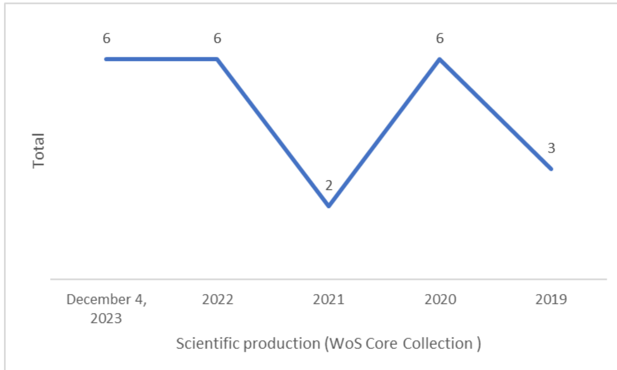
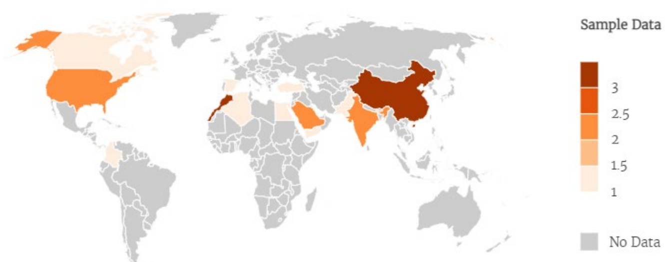


Figure 3 was created with mapsinseconds.com to illustrate the first authors' countries in our sample. The two countries with the highest scientific production were China (n = 10) and Morocco (n = 2).

Figure 3

Origin of the Works Analysed (N = 23)



Bibliometric Dimension (Question 1)

A co-authorship network of the 147 authors was created. Of these, collaboration was shown in 83 (Figure 4). The most cited names were those with the highest presence, represented by 23 clusters.

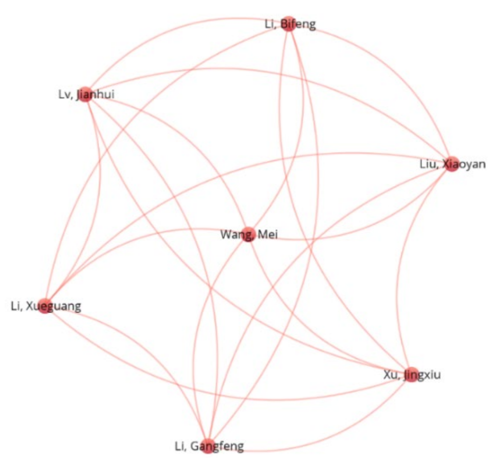
Figure 4

Co-Authorship Network

Of the 83 authors, only seven showed strong collaboration (Figure 5).

Figure 5

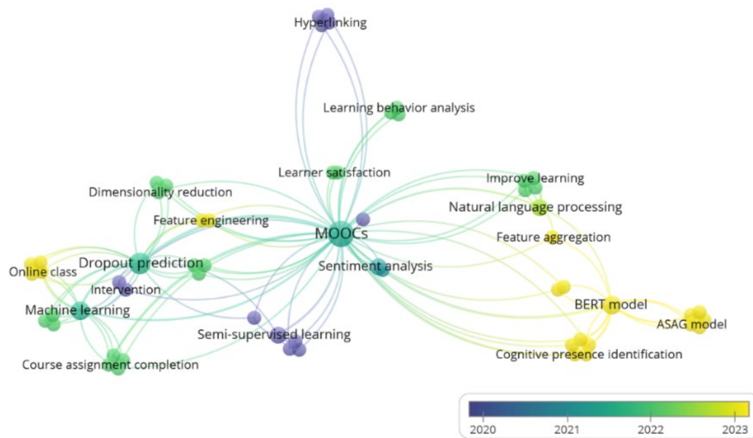
Co-Authorship Network with Consistent Connections



A total of 81 author keywords were identified. After debugging and standardising common keywords and abbreviations, these were reduced to 68 (Figure 6).

Figure 6

Author Keyword Network



In total, 14 clusters were identified, highlighting the following: In the period of 2022–2023, research focused on the utilization of feature aggregation and BERT and ASAG models to analyze students' cognitive presence in Massive Open Online Courses (MOOC) courses.

During the period of 2021–2022, studies emphasized predicting dropout through sentiment analysis, the integrated use of machine learning (ML) and deep learning (DL) techniques, and the analysis of learner traces of interaction and interactivity. Similarly, in the period of 2019–2021, work also concentrated on the prediction of MOOC courses, mainly based on models that integrated DL and semi-supervised learning. The results indicated that the primary areas of application of DL in open learning were MOOC course recommendation, student dropout prediction, cognitive presence analysis, and sentiment analysis. These aspects are further explored in the discussion of the pedagogical dimension.

Pedagogical Dimension (Question 2)

Each study was analyzed in-depth, and the results are presented in Table 3.

Table 3

Dependent Variables and Main Findings (N = 23)

ID	Citation	Dependent variable	Findings
ID 1	Li et al. (2023)	Recommending MOOCs	Automatic detection of user needs
ID 2	Liu et al. (2023a)	Cognitive presence	Cognitive presence assessment
ID 3	El-Rashidy et al. (2023)	Performance of MOOC posts classification	Automatic quality assessment of forums in MOOC courses
ID 4	Alruwais (2023)	Predicting MOOC dropout	Dropout determinants in MOOCs (e.g., video clickstream, forum interaction)

ID 5	Liu et al. (2023b)	Predicting MOOC dropout	Analysed video interactions based on learner and course characteristics
ID 6	Mardini et al. (2023)	Reading comprehension assessment	Automatic grading of short answers
ID 7	Nithya and Umarani (2022)	Predicting MOOC dropout	Relationship between learner behaviour (i.e., interaction and interactivity) and MOOC course dropout
ID 8	Zheng et al. (2022)	Predicting MOOC dropout	Determinants of dropout in MOOCs (i.e., videos, task completion, and interactivity in forums)
ID 9	Lemay and Doleck (2022)	Predicting MOOC dropout	Determinants of video features in MOOCs
ID 10	Zakaria et al. (2022)	Predicting MOOC dropout	Relationship between interactivity and interaction time, and MOOC course dropout
ID 11	Jiang (2022)	Teaching modern and contemporary literature in Chinese	Identifying possible causes of dropout in relation to language and literature learning
ID 12	Tzeng et al. (2022)	MOOC student experiences	Predicting student satisfaction
ID 13	Fan et al. (2022)	Learning behaviours and MOOC recommendation	Relationship between didactic description of the MOOC and personal learning goals
ID 14	Hamal and El Faddouli (2022)	Answering learner questions in a MOOC	Answering learner questions related to French language learning
ID 15	Mubarak et al. (2021)	Predicting learners' performance	Determinants of dropout in MOOCs (analysis of video interactions)
ID 16	Mrhar et al. (2021)	Sentiment analysis on forum interactivity in MOOCs	Correlation between sentiment (forum interactions) and MOOC dropout rate
ID 17	Onan (2021)	Sentiment analysis on assessments in MOOCs	Correlation between sentiment and MOOC dropout rate
ID 19	Goel and Goyal (2020)	Predicting MOOC dropout	Correlations among possible friends, their closeness levels, and the probability of dropout
ID 20	Yin et al. (2020)	Predicting weekly MOOC dropout	Probability of weekly dropout, based on interaction and interactivity
ID 21	Xing and Du (2019)	Predicting weekly MOOC dropout	Probability of weekly dropout, based on analysis of forum-type activities and quizzes

ID 22	Hassan et al. (2019)	Predicting weekly MOOC dropout	Relationship between video click-through rate and likelihood of dropout
ID 23	An et al. (2019)	Predicting MOOC dropout	Probability of dropout, based on analysis of forum-type activities

The main findings focused on the analysis and prediction of learner behaviour, interaction, and interactivity in MOOCs, although some specific cases evaluated other open learning systems (Hamal & El Faddouli, 2022; Mardini et al., 2023).

Technological Dimension (Question 3)

The DL techniques and algorithms used, and their level of accuracy, are shown in Table 4.

Table 4

Techniques, Level of Accuracy, and Data Sources

ID	DL technique or algorithm	Data source	Accuracy
ID 1	Bidirectional encoder representations from transformers (BERT)	Open datasets (MoocCube)	2.150 F1-score@10 0.2854 recall@10 0.172 precision@10
ID 2	MOOC-BERT (BERT model variant)	Datasets from two Chinese university MOOCs	85.8 % precision 86.1% recall 85.9% F1-score 88.1% accuracy
ID 3	BERT New model based on CNNs	Datasets present in the Stanford post-MOOC corpus	83.6 % precision 83% recall 83.3% F1 of urgent 92.7% F1-weighted
ID 4	Factorisation machine with DNN (deep neural network) models	Two datasets (HarvardX person-course academic year 2013 de-identified and MOOC)	99 % accuracy
ID 5	Learning network model (LBDL) and Bi-LSTM	Open datasets (MoocCube)	74.89% F1-score 82.39% ROC curve
ID 6	Deep-learning-based grading system BERT	Universidad del Norte datasets	BERT-1-ES [Pearson correlation (0.78) Root mean squared error (0.66)] BERT-2-ES [Pearson correlation (0.78) Root mean squared error (0.66)]
ID 7	FIAR-ANN Model	KDD Cup 2015 dataset	3.16 F1-score 92.42% accuracy.
ID 8	CNNs and bidirectional long short-term memory network (Bi-LSTM)	KDD Cup 2015 dataset	87.1% area under the receiver operating characteristic curve

			87.3 % area under the precision-recall curve (AUC) 86.8% F1-value 86.4% accuracy
ID 9	Logistic regression, SMO (Sequential minimal optimisation), Naïve Bayes, J48, JRip, IBK and WekaDeeplearning4J	Dataset analysis of MOOCs offered at the University of Pennsylvania	High accuracy (~80%) Variance (26.9%) Kappa > 0 and AUC > 0.5
ID 10	Deep neural network (DNN)	KDD Cup 2015 dataset	0.943 accuracy 0.876 AUC
ID 11	Basic teaching and learning optimisation algorithm	Questionnaires to students enrolled in MOOC courses	No validation of the algorithm presented, only students' opinions
ID 12	Deep neural network (DNN)	Eight MOOC courses from National Tsing Hua University (NTHU) Video analysis. Questionnaire analysis	Cronbach's alpha (0.842) Mean absolute error (0.41 to 0.55)
ID 13	Multi-attention deep learning model	Records of 6,628 students from 1,789 MOOCs	0.90 Hit ratio-20. 0.58 Normalised discounted cumulative gain-20
ID 14	DL in NLP	Proposal validated by two datasets (FQuAD, SQuAD-FR)	FCuAD: 79.81 F1-score SquaD-FR: 80.61 F1-score
ID 15	LSTM	Stanford University MOOCs	89%–95% accuracy 89% automata theory accuracy 90.30% in the "Mining of Massive Datasets"
ID 16	Bayesian CNN-LSTM Model	100,000 Coursera course reviews	90% precision 85% recall 88% F1-score 91.27% accuracy
ID 17	Three word embedding schemes (word2vec, fastText and GloVe) Long short-term memory networks (LSTM)	66,000 course reviews on coursetalk.com	95.80% accuracy
ID 18	Semi-supervised deep learning (SSDL) framework	Stanford MOOC posts dataset	89.73% accuracy 93.55% F1-score
ID 19	Self-training model	XuetangX (Datasets from China)	94.29% average F1-score
ID 20	Deep neural network model	KDD Cup 2015 dataset	Weekly average accuracy: Week 1 (0.84%) Week 2 (0.73%) Week 3 (0.87%) Week 4. (0.91%)

			Week 5 (0.84%)
ID 21	Temporal prediction mechanism	MOOC course dataset	Accuracy range 0.928 to 0.981 Dropout precision 93%
ID 22	Long short-term memory (LSTM deep)	Open University learning analytics	97.25% learning accuracy, 92.79% precision 85.92% recall
ID 23	New incremental model of LSTM–CRF	Real-world forum posts from Coursera	Dropout precision 65.6% 3.16 F1-score

As can be seen, the most commonly used DL techniques or algorithms were related to the classical use or modern variants of BERT, LSTM, DNN, and NLP. Their levels of accuracy are adequate to measure the dependent variables (refer to Table 3).

Discussion and Conclusions

This study aimed to identify, through a systematic review, the trends, applied computational techniques, and areas of educational use of deep learning in open learning. To this end, the WoS Core Collection, including both SCIE and SSCI, was searched. The PRISMA protocol guided the selection of 23 articles, reflecting a low academic production related to the research objective, which indicates that deep learning (DL) in education is relatively recent.

Concerning the first research question, the application of DL has mainly focused on predicting student dropout, automatic grading of short answers, and recommending MOOC courses. Technological underpinnings were based on (a) the flow of student clicks on videos, (b) student interaction and interactivity, (c) the quality of responses in interactive activities such as forums, and (d) the length of time spent on activities. The low academic research output reflects that there is still insufficient scientific collaboration in this research area, which may be a consequence of how recent the use of DL is in open learning (Pan et al., 2023; Vanitha & Jayashree, 2023).

For the second and third research questions, dealing with pedagogical and technological dimensions, respectively, we identified trends in the use of DL in open learning in several key directions.

Pedagogical Dimension

Predicting Dropout or Attrition From MOOCs

There is agreement that the quality and length of videos in MOOCs influence dropout (Alruwais, 2023; Goel & Goyal, 2020; Hassan et al., 2019; Lemay & Doleck, 2022; Liu et al., 2023a; Mubarak et al., 2021; Nithya & Umarani, 2022; Tzeng et al., 2022; Zakaria et al., 2022; Zheng et al., 2022). In this regard, several researchers (Nithya & Umarani, 2022; Zheng et al., 2022) found that browsing and closing pages had no effect on dropout; completing tasks, watching videos, and discussing problems in forums did. On the other hand, Zakaria et al. (2022) successfully elucidated the relationships among access, video engagement, homework completion, and discussion participation in predicting dropout. Others (Alruwais, 2023; Goel & Goyal, 2020; Hassan et al., 2019; Liu et al., 2023b; Mubarak et al., 2021; Zheng et al., 2022) agreed that the lower the clickstream on videos, the higher the probability of dropping out of a MOOC course.

As an interesting note, we agreed with Goel and Goyal (2020) that AI studies in general have identified that there is a possible relationship between personal or friendship relationships and the likelihood of dropout— in other words, virtually all likely friends among all enrolled students showed the same behaviour in both video interaction and dropout likelihood.

A relationship has been established between the course content, the characteristics of open educational resources, their complexity, and student cognitive fatigue (Jiang, 2022; Zakaria et al., 2022). In addition, analysing these factors as well as student interactivity and interaction, through logs and traces, has established the probability of weekly dropout (Yin et al., 2020). Similarly, analysis of interaction in forums (An et al., 2019; El-Rashidy et al., 2023) and quizzes (Xing & Du, 2019) in a MOOC course can help predict the probability of weekly dropout.

Analysis of students' cognitive presence in sustained discourse in a virtual community (e.g., integration, problem-solving, and intuition) is a high predictor of academic performance in MOOCs (Liu et al., 2023a).

Sentiment analysis or opinion mining of individual student responses allows for some assessment and prediction of retention or dropout in a MOOC course. In this regard, research has been conducted on mass assessments (Onan, 2021) and forum interactions (Chen, Feng, et al., 2020; Mrhar et al., 2021). There has been some hesitation from students regarding mass assessments related to quality and veracity, although the computational results provided good reliability rates (Onan, 2021). A hybrid procedure between automated and teacher-led assessments was suggested in several reviews (Mrhar et al., 2021). Furthermore, the analysis of interaction and interactivity in activities posted in forums within MOOCs showed that there was a correlation between learner sentiment and engagement, and the likelihood of dropping out. As the number of students dropping out of the MOOC course decreased, the feelings and motivation towards the course increased (Chen, Feng, et al., 2020; Mrhar et al., 2021).

In this area (MOOC dropout), the results can be grouped into the following clusters:

- Cluster 1: Dropout prediction using artificial neural networks, association rules mining, data analytics, ML, and personalisation.

- Cluster 2: Sentiment classification using asymmetric data, co-training, self-training, and semi-supervised learning.

- Cluster 3: The identification of cognitive presence through the community of inquiry model, online discussions, pre-trained language model, and text analysis.

- Cluster 4: Students' dropout through deep-neural networks and the deepfm model.

Automatic Grading of Short Answers

Scoring short answer reading comprehension questions is effective if a comparison is made between the student's answer and the target answer, yet it is a complex process that has not yet been fully resolved (Hamal & El Faddouli, 2022; Mardini et al., 2023). Research related to this topic has been based on the use of feature aggregation, intelligent systems, and the use of DL in NLP.

Recommending MOOCs Personalized recommendations of Massive Open Online Courses (MOOCs) have been based on two trends: (a) the use of big data and deep learning (DL) through the analysis of content features (Li et al., 2023); or (b) through learning logs, content, and course descriptions (Fan et al., 2022). These authors agreed that the accurate wording of learning objectives and didactic description of a course influenced the effectiveness of the recommendations.

Related research has been focused on automated grading, big data, reading comprehension assessment, sentence embedding, the ASAG model, and the skip-thoughts model. Although significant results have been achieved, three techno-pedagogical challenges associated with these systems have remained: (a) ensuring pedagogical usability, (b) the design of quality computational models, and (c) confidence in communicating and grading learning outcomes.

Synthesis The literature review allowed us to identify some pedagogical challenges of using DL in open learning, including (a) effective personalization of content for different learning styles; (b) transparent interpretation and explanation of model decisions; and (c) the need to address possible inherent biases in the datasets (e.g., socio-demographics, traces, competencies, learning objectives) used for training. In addition, continuous adaptation as technologies evolve and ethical integration of artificial intelligence are key aspects to consider in educational settings (Tzeng et al., 2022).

Technological Dimension The following technological challenges were apparent in the literature we analyzed (N = 23):

- Large datasets: DL models often require massive datasets for optimal performance, which can be difficult to obtain in open learning environments where data availability may be limited.

- Intensive computing: DL algorithms are computationally intensive, which implies the need for powerful hardware resources. This can be a challenge in environments where access to high-end computational resources is limited.

- Interpretability: DL models are often perceived as black boxes due to their complexity. Understanding how they make decisions can be crucial, especially in contexts where transparency is essential.

- Knowledge transfer: Adapting pre-trained models to new tasks can be challenging, as knowledge transfer is not always straightforward and may require sensitive fine-tuning.

- Ethics and bias: The presence of biases in datasets can lead to biased and discriminatory results. Addressing these ethical issues is essential for inclusive and fair open learning.

- Safety: DL models can be vulnerable to adversarial attacks, where carefully designed inputs can mislead the model. Ensuring the robustness of the model is a constant challenge.

- Cost of implementation: Developing and implementing DL solutions can be costly in terms of human resources, hardware, and time. This can limit its adoption in resource-constrained contexts.

Despite these challenges, research advances have continued to address these concerns and improve the applicability of DL in open learning environments. In this educational domain, various DL algorithms have been employed, such as convolutional neural networks (CNNs) for processing visual data, recurrent neural networks (RNNs) for temporal sequences, and transformers for natural language processing (NLP) tasks (Hamal & El Faddouli, 2022). Techniques have included transfer learning, data augmentation, and personalized optimization. In terms of trends, there has been an increase in the application of pre-trained models, the development of more efficient architectures, and the growing use of interpretability techniques. In this sense, the integration of Artificial Intelligence (AI) in the personalization of the learning experience is an emerging trend reiterated in the Horizon Reports (EDUCAUSE, 2023).

Conclusion

This study aimed to contribute to the educational community by identifying the pedagogical potential of Deep Learning (DL) in open learning, such as:

- Content recommendation and automatic grading of short answers with feedback as key tools for meaningful personalized learning, achieved through constructive alignment between objectives, activities, and learning assessment.

- Predicting student dropout based on levels of student interactivity and interaction with digital educational resources, and the quality of responses to self-assessment as well as formative and summative assessment activities. This would be linked to an adequate multidirectional synchronous and asynchronous pedagogical communication and interaction process, providing support and tutoring services to motivate the student to learn in an autonomous, personalized, and collaborative way.

The results obtained in the application of DL in open learning have influenced the efficiency of higher education administrations, early counseling, and mentoring, as well as the design and implementation of educational interventions. However, there is agreement on the need to delve deeper into the ethical and moral issues of artificial intelligence (AI) concerning cultural differences, inclusion, and student emotions, as well as the pedagogical use of AI by teachers (Mouta et al., 2023).

At the algorithmic level, the most commonly employed DL algorithms were variants of artificial neural networks such as DNN (Alruwais, 2023; Tzeng et al., 2022; Yin et al., 2020; Zakaria et al., 2022), LSTM (An et al., 2019; Hassan et al., 2019; Liu et al., 2023b; Mrhar et al., 2021; Mubarak et al., 2021; Onan, 2021; Zheng et al., 2022), and BERT (El-Rashidy et al., 2023; Li et al., 2023; Liu et al., 2023a; Mardini et al., 2023; Zheng et al., 2022).

The studies reviewed demonstrated their contribution to education; however, there was a lack of research to follow up on these results by answering questions such as: Have Massive Open Online Courses (MOOCs) been redesigned based on the results of DL application in open learning? (See Table 3); and Is the dropout rate maintained?

Learning as an educational, cultural, and psychosocial process depends on a variety of cognitive, motivational, affective, communicative, sociological, pedagogical, didactic, and technological factors. In terms of technology, DL and AI in education have brought us closer to identifying some criteria for approaching success in open learning. The algorithms and methods used have offered a high cognitivist weight of pedagogical value but could be enriched with other pedagogical foundations. It is interesting that the description of the DL methods used has hardly described the pedagogical basis, which to some extent obscures educational assessment.

Limitations of the Study

This study was limited to consulting only the Web of Science (WoS) Core Collection, specifically SCIE and SSCI. Therefore, it is possible that interesting results published in journals indexed in Scopus or other databases were omitted. Additionally, only articles in English were analyzed, overlooking articles published in other languages that could have enriched the results obtained in this research.

Future Lines of Research

There has been a lack of studies that have analyzed and compared the results obtained in the use of DL and machine learning (ML) in open learning, which would help to make decisions and consequently define the most efficient techniques and algorithms. In this sense, a systematic review and meta-analysis are recommended.

References

Alhothali, A., Albsisi, M., Assalahi, H., & Aldosemani, T. (2022). Predicting student outcomes in online courses using machine learning techniques: A review. *Sustainability, 14*(10), 6199. https://doi.org/10.3390/su14106199

Alruwais, N. M. (2023). Deep FM-based predictive model for student dropout in online classes. *IEEE Access, 11*, 96954–96970. https://doi.org/10.1109/ACCESS.2023.3312150

An, Y. H., Pan, L., Kan, M. Y., Dong, Q., & Fu, Y. (2019). Resource mention extraction for MOOC discussion forums. *IEEE Access, 7*, 87887–87900. https://doi.org/10.1109/ACCESS.2019.2924250

Chassagnon, G., Vakalopolou, M., Paragios, N., & Revel, M. (2019). Deep learning: Definition and perspectives for thoracic imaging. *European Radiology*, *30*(4), 2021–2030. https://doi.org/10.1007/s00330-019-06564-3

Chen, J., Feng, J., Sun, X., & Liu, Y. (2020). Co-Training semi-supervised deep learning for sentiment classification of MOOC forum posts. *Symmetry, 12*(1). https://doi.org/10.3390/sym12010008

Chen, X., Xie, H., & Hwang, G. J. (2020). A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers. *Computers and Education: Artificial Intelligence, 1*, 100005. https://doi.org/10.1016/J.CAEAI.2020.100005

Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). Application and theory gaps during the rise of artificial intelligence in education. C*omputers and Education: Artificial Intelligence, 1*. https://doi.org/10.1016/J.CAEAI.2020.100002

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*(1), 1–22. https://doi.org/10.1186/S41239-023-00392-8/FIGURES/11

EDUCAUSE. (2023). *2023 EDUCAUSE Horizon Report® Teaching and learning edition*. *https://library.educause.edu/-/media/files/library/2023/4/2023hrteachinglearning.pdf*

El-Rashidy, M. A., Farouk, A., El-Fishawy, N. A., Aslan, H. K., & Khodeir, N. A. (2023). New weighted BERT features and multi-CNN models to enhance the performance of MOOC posts classification. *Neural Computing & Applications, 35*(24), 18019–18033. https://doi.org/10.1007/s00521-023-08673-z

Fan, J., Jiang, Y., Liu, Y., & Zhou, Y. (2022). Interpretable MOOC recommendation: A multi-attention network for personalized learning behavior analysis. *Internet Research, 32*(2), 588–605. https://doi.org/10.1108/INTR-08-2020-0477

Goel, Y., & Goyal, R. (2020). On the effectiveness of self-training in MOOC dropout prediction. *Open Computer Science, 10*(1), 246–258. https://doi.org/10.1515/comp-2020-0153

Hamal, O., & El Faddouli, N. E. (2022). Intelligent system using deep learning for answering learner questions in a MOOC. *International Journal of Emerging Technologies in Learning, 17*(2), 32–42. https://doi.org/10.3991/ijet.v17i02.26605

Hassan, S., Waheed, H., Aljohani, N. R., Ali, M., Ventura, S., & Herrera, F. (2019). Virtual learning environment to predict withdrawal by leveraging deep learning. *International Journal of Intelligent Systems, 34*(8), 1935–1952. https://doi.org/10.1002/int.22129

Hu, Y., Ferreira Mello, R., & Gašević, D. (2021). Automatic analysis of cognitive presence in online discussions: An approach using deep learning and explainable artificial intelligence. *Computers and Education: Artificial Intelligence, 2*, 100037. https://doi.org/10.1016/J.CAEAI.2021.100037

Hwang, G.-J., Tu, Y.-F., Hwang, G.-J. , & Tu, Y.-F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics, 9*(6). https://doi.org/10.3390/MATH9060584

Jiang, H. (2022). Modern and contemporary literature courses in colleges and universities using the teaching mode of deep learning. *Mobile Information Systems, 2022*, 1–13. https://doi.org/10.1155/2022/3517022

Koong Lin, H.-C., Ye, J.-H., Hao, Y.-W., Wu, Y.-F., Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis. *Sustainability, 15*(17). https://doi.org/10.3390/SU151712983

LeCun, Y., Bengio, Y., & Hinton, G. E. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lemay, D. J., & Doleck, T. (2022). Predicting completion of massive open online course (MOOC) assignments from video viewing behavior. *Interactive Learning Environments, 30*(10), 1782–1793. https://doi.org/10.1080/10494820.2020.1746673

Li, B., Li, G., Xu, J., Li, X., Liu, X., Wang, M., & Lv, J. (2023). A personalized recommendation framework based on MOOC system integrating deep learning and big data. *Computers and Electrical Engineering, 106*, 108571. https://doi.org/10.1016/j.compeleceng.2022.108571

Liang, J. C., Hwang, G. J., Chen, M. R. A., & Darmawansah, D. (2023). Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments, 31*(7), 4270–4296. https://doi.org/10.1080/10494820.2021.1958348

Liu, Z., Kong, X., Chen, H., Liu, S., & Yang, Z. (2023a). MOOC-BERT: Automatically identifying learner cognitive presence from MOOC discussion data. *IEEE Transactions on Learning Technologies, 16*(4), 528–542. https://doi.org/10.1109/TLT.2023.3240715

Liu, H., Chen, X., & Zhao, F. (2023b). Learning behavior feature fused deep learning network model for MOOC dropout prediction. *Education and Information Technologies, 29*, 3257–3278. https://doi.org/10.1007/s10639-023-11960-w

Mardini, G., Quintero, C., Viloria, C., Percybrooks, W., Robles, H., & Villalba, K. (2023). A deep-learning-based grading system (ASAG) for reading comprehension assessment by using aphorisms as open-answer-questions. *Education and Information Technologies, 29*, 4565–4590. https://doi.org/10.1007/s10639-023-11890-7

Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2023). Uncovering blind spots in education ethics: Insights from a systematic literature review on artificial intelligence in education. *International Journal of Artificial Intelligence in Education, 2023*. https://doi.org/10.1007/s40593-023-00384-9

Mrhar, K., Benhiba, L., Bourekkache, S., & Abik, M. (2021). A Bayesian CNN-LSTM model for sentiment analysis in massive open online courses MOOCs. *International Journal of Emerging Technologies in Learning, 16*(23), 216–232. https://doi.org/10.3991/ijet.v16i23.24457

Mubarak, A. A., Cao, H., Zhang, W., & Zhang, W. (2021). Visual analytics of video-clickstream data and prediction of learners' performance using deep learning models in MOOCs' courses. *Computer Applications in Engineering Education, 29*(4), 710–732. https://doi.org/10.1002/cae.22328

Nithya, S., & Umarani, S. (2022). MOOC dropout prediction using FIAR-ANN model based on learner behavioral features. *International Journal of Advanced Computer Science and Applications, 13*(9), 607–617. https://dx.doi.org/10.14569/IJACSA.2022.0130972

Onan, A. (2021). Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach. *Computer Applications in Engineering Education, 29*(3), 572–589. https://doi.org/10.1002/cae.22253

Otto, D., Scharnberg, G., Kerres, M., & Zawacki-Richter, O. (2023). Distributed learning ecosystems: Concepts, resources, and repositories. In D. Otto, G. Scharnberg, M. Kerres, & O. Zawacki-Richter (Eds.), *Distributed learning ecosystems* (pp. 1–10). Springer VS. https://doi.org/10.1007/978-3-658-38703-7_1

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*(71). https://doi.org/10.1136/bmj.n71

Pan, Q., Zhou, J., Yang, D., Shi, D., Wang, D., Chen, X., & Liu, J. (2023). Mapping knowledge domain analysis in deep learning research of global education. *Sustainability, 15*(4), 3097. https://doi.org/10.3390/su15043097

Salas-Rueda, R. A. (2023). Use of deep learning to analyze Facebook and Google Classroom in the educational field. *Pixel-Bit. Revista De Medios y Educación, 67*, 87–122. https://doi.org/10.12795/pixelbit.96994

Shafiq, D. A., Marjani, M., Habeeb, R. A. A., & Asirvatham, D. (2022). Student retention using educational data mining and predictive analytics: A systematic literature review. *IEEE Access, 10*, 72480–72503. https://doi.org/10.1109/ACCESS.2022.3188767

Su, J., Ng, D. T. K., & Chu, S. K. W. (2023). Artificial intelligence (AI) literacy in early childhood education: The challenges and opportunities. *Computers and Education: Artificial Intelligence, 4*, 100124. https://doi.org/10.1016/J.CAEAI.2023.100124

Talan, T. (2021). Artificial intelligence in education: A bibliometric study. *International Journal of Research in Education and Science, 7*(3), 822–837. https://doi.org/10.46328/IJRES.2409

Tzeng, J.-W., Lee, C.-A., Huang, N.-F., Huang, H.-H., & Lai, C.-F. (2022). MOOC evaluation system based on deep learning. *The International Review of Research in Open and Distributed Learning, 23*(1), 21–40. https://doi.org/10.19173/irrodl.v22i4.5417

Uddin, I., Imran, A. S., Muhammad, K., Fayyaz, N., & Sajjad, M. (2021). A systematic mapping review on MOOC recommender systems. *IEEE Access, 9*, 118379–118405. https://doi.org/10.1109/ACCESS.2021.3101039

Vanitha, S., & Jayashree, R. (2023). Towards finding the impact of deep learning in educational time series datasets: A systematic literature review. *International Journal of Advanced Computer Science and Applications, 14*(3), 804–815. http://dx.doi.org/10.14569/IJACSA.2023.0140392

Verma, N., Getenet, D. S., Dann, D. C., & Shaik, T. (2023). Designing an artificial intelligence tool to understand student engagement based on teacher's behaviours and movements in video conferencing. *Computers and Education: Artificial Intelligence, 5*, 100187. https://doi.org/10.1016/J.CAEAI.2023.100187

Wang, D., Tao, Y., & Chen, G. (2024). Artificial intelligence in classroom discourse: A systematic review of the past decade. *International Journal of Educational Research, 123*, 102275. https://doi.org/10.1016/J.IJER.2023.102275

Xing, W., & Du, D. (2019). Dropout prediction in MOOCs: Using deep learning for personalized intervention. *Journal of Educational Computing Research, 57*(3), 547–570. https://doi.org/10.1177/0735633118757015

Yin, S., Lei, L., Wang, H., & Chen, W. (2020). Power of attention in MOOC dropout prediction. *IEEE Access, 8*, 202993–203002. https://doi.org/10.1109/ACCESS.2020.3035687

Yu, J., Alrajhi, L., Harit, A., Sun, Z., Cristea, A. I., & Shi, L. (2021). Exploring Bayesian deep learning for urgent instructor intervention need in MOOC forums. In A. I. Cristea & C. Troussas (Eds.), *Intelligent tutoring systems, ITS 2021: Lecture notes in computer science* (pp. 78–90). Springer. https://doi.org/10.1007/978-3-030-80421-3_10

Zakaria, A., Anas, B., & Oucamah, C. M. M. (2022). Intelligent system for personalised interventions and early drop-out prediction in MOOCs. *International Journal of Advanced Computer Science and Applications, 13*(9). https://doi.org/10.14569/IJACSA.2022.0130983

Zawacki-Richter, O., Kerres, M., Bedenlier, S., Bond, M., & Buntins, K. (2020). *Systematic reviews in educational research*. Springer VS. https://doi.org/10.1007/978-3-658-27602-7

Zheng, Y., Shao, Z., Deng, M., Gao, Z., & Fu, Q. (2022). MOOC dropout prediction using a fusion deep model based on behaviour features. *Computers and Electrical Engineering, 104*, 108409. https://doi.org/10.1016/j.compeleceng.2022.108409

International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 25, Number 3

August – 2024

Potentials and Implications of ChatGPT for ESL Writing Instruction

Dr. Karim Ibrahim and Dr. Robert Kirkpatrick
Gulf University for Science and Technology

Abstract

The release of ChatGPT has marked the dawn of a new information revolution that will transform how people communicate and make meaning. However, to date, little is known about the implications of ChatGPT for L2 composition instruction. To address this gap, the present study uses a systematic review design to synthesize available research on the educational potentials of ChatGPT as an instructional assistant, outline the implications of these potentials for L2 writing instruction, and discuss their practical applications. The findings, based on a meta-analysis of 42 research articles, demonstrate that ChatGPT can enhance L2 writing instruction by boosting learners' motivation, automating instructional tasks, and offering instantaneous, personalized feedback to learners. These findings have important implications for harnessing the instructional potential of generative AI in L2 writing classes.

Keywords: artificial intelligence, artificial neural network, L2 writing, AI applications, chatbots

Introduction

The public release of ChatGPT in November 2022 marked the dawn of a new information literacy revolution in which generative AI may come to dominate the generation, classification, and dissemination of knowledge in information and communication technologies (ICT). The impressive abilities of ChatGPT to interact verbally with users, perform complex natural language processing tasks, and generate coherent, intelligible, and humanlike texts (Lund & Wang, 2023; Radford et al., 2018; Radford et al., 2019; Solaiman et al., 2019) attracted 100 million users within two months of its release, making it the fastest-growing Internet app ever (Hu, 2023). These revolutionary abilities have transformational effects on education as they facilitate the automation of various mundane tasks that are integral to education (Ibrahim, 2023).

A sub-field of education that has been deeply influenced by the rise of generative AI is second language (L2) writing/composition instruction, perhaps because of the wide overlap between the cognitive skills taught in writing classes and generative AI capabilities. The unprecedented capabilities of ChatGPT to process and generate texts hold both great potentials and challenges to L2 writing instruction, but most of the existing research on the implications of ChatGPT for L2 writing has emphasized ethical challenges to traditional L2 writing assessments (Cotton et al., 2024; Francke & Alexander, 2019; Gao et al., 2023; Haque et al., 2022; King & ChatGPT, 2023; Khalil & Er, 2023; Susnjak, 2022; Yeadon et al., 2022). As a result, little is known about the potential of ChatGPT to support L2 writing instruction and development.

To address this gap in the literature and shed light on the practical implications of this promising new technology, the present study will review available research on the potentials and implications of ChatGPT for L2 writing instruction. It is hoped that this review can underline the instructional benefits of ChatGPT for L2 writing, provide guidelines for integrating ChatGPT in L2 writing classes, and outline essential areas and questions for future research in this new field. These findings can help L2 writing instructors harness the instructional potentials of ChatGPT and integrate AI literacies in their classes. The paper uses a systematic review approach to summarize and synthesize research findings on the applications of ChatGPT for L2 writing instruction. Research articles were gathered through an extensive search on Web of Science, EBSCOHost, Google Scholar, and arXiv databases using the following keywords: “ChatGPT and L2 writing,” “ChatGPT and writing instruction,” “ChatGPT and automatic feedback,” “generative AI and instruction automation,” and “ChatGPT and academic writing motivation.” The search results generated 58 relevant research articles, but 16 articles were excluded over methodological issues; thus, 42 articles were finally included in the literature review.

Literature Review

ChatGPT

The generative pre-trained transformer model (GPT) is an auto-regressive large language model (LLM) equipped with a chatbot, ChatGPT, that allows Internet users to interact verbally with the model via a virtual agent called “Assistant” (Radford et al., 2018). GPT is a machine-learning neural network trained to model natural language systems (MacNeil et al., 2022) and perform natural language processing (NLP) tasks, including writing articles, composing music, answering questions, and generating computer code (OpenAI, n.d.-b). GPT was trained using an innovative machine-learning approach that combines unsupervised pre-training (Radford et al., 2018) and reinforcement learning from human feedback approaches (Christiano et al., 2017) on a massive corpus of one billion words of web texts (Radford et al., 2019). This unique training approach allows the model to recognize patterns in the training data and develop different NLP skills without task-specific training (Brown et al., 2020). This approach and rich training data allowed GPT to grow into a 175-billion-parameter model that can generate fluent, coherent, and humanlike texts on various topics (Ibrahim, 2023). The latest version of the model, GPT-4, is multimodal and can interact with users verbally and interpret audiovisual content (OpenAI, n.d.-b).

ChatGPT and L2 Education

As an intelligent chatbot, ChatGPT can leverage its natural language processing capabilities, deep learning dynamics, and extensive training to support L2 education (Huang & Li, 2023; Hong, 2023; Kohnke et al., 2023; Rahman & Watanobe, 2023). ChatGPT can be used as an intelligent L2 tutor to offer L2 learners extended L2 practice opportunities (Rahman & Watanobe, 2023) and corrective feedback on their L2 output (Rudolph et al., 2023). ChatGPT can use its machine-learning capabilities to analyze L2 learners' performance and personalize their L2 learning trajectories to match their individual learning needs (Huang & Li, 2023). These potentials would boost autonomous L2 learning and learners' engagement (Qadir, 2022). ChatGPT could also offer L2 educators rich L2 learning resources that support autonomous and collaborative L2 instruction and learning, including practice activities, grammar worksheets, and vocabulary illustration (Rahman & Watanobe, 2023).

A few exploratory studies have demonstrated the potential of ChatGPT to support L2 education (Rakhmonov & Kurbonova, 2023; Shaikh et al., 2023; Young et al., 2023). Rakhmonov and Kurbonova (2023) investigated the L2 learning potentials of ChatGPT with 20 L2 learners and instructors. Analysis of survey and interview data revealed that ChatGPT could generate personalized L2 learning content that matches learners' needs. Similarly, Shaikh et al. (2023) explored the usability of ChatGPT in L2 learning contexts with 10 Norwegian ESL learners. The researchers engaged the participants in ChatGPT-mediated ESL learning activities and collected their responses via a questionnaire. Data analysis revealed that the participants perceived ChatGPT to be a valuable and accessible resource for L2 practice. In a similar vein, Young et al. (2023) investigated the appropriateness of ChatGPT output for ESL education. Using a readability matrix analysis, the researchers analyzed the appropriateness of English dialogues generated by ChatGPT for ESL learners. Data analysis revealed that ChatGPT-generated content is deemed appropriate for the learning needs of ESL learners.

ChatGPT as an L2 Writing Instructional Assistant

ChatGPT can support L2 writing instruction and enhance teachers' efficiency by automating instructional planning, materials design, assessment, and course management (Qadir, 2022; Rudolph et al., 2023). Firstly, AI can help writing instructors reduce their administrative and planning workload so they can spend more time designing L2 writing opportunities by automating instructional design tasks (Baidoo-anu & Ansah, 2023). They could use ChatGPT to design course outlines and plans (Rahman & Watanobe, 2023) tailored to specific L2 writing objectives and goals (Baskara & Mukarto, 2023); generate specific, well-articulated, and measurable intended learning outcomes based on Bloom's taxonomy for specific L2 writing sub-skills (Sridhar et al., 2023); and create custom-designed lesson plans that follow specific pedagogical approaches or methods (Hong, 2023). For instance, in an exploratory study, Sridhar et al. (2023) examined the potential of GPT-4 to support instructional design by authoring high-quality learning outcomes (LOs) for a university course on AI. Using Bloom's taxonomy as a frame of reference, the researchers prompted GPT-4 by providing specific guidelines for the designed LOs, including a course description, instructional design guidelines, design specifications for the LOs, and examples of well-designed LOs. Analysis of the generated LOs by human reviewers and AI classifiers demonstrated that they were sensible, appropriate for the intended cognitive processes, and consistent with Bloom's taxonomy. This study demonstrates that ChatGPT can assist writing instructors by automating instructional design processes.

Second, L2 writing instructors can use ChatGPT to generate rich instructional resources tailored to specific learning objectives, learning preferences, and pedagogical approaches (Baskara & Mukarto, 2023; Huang & Li, 2023; Hong, 2023; Kohnke et al., 2023; Qadir, 2022; Rahman & Watanobe, 2023). For instance, the model can support writing instruction by generating writing prompts (Baskara & Mukarto, 2023), writing practice activities, and learning tips and suggestions (Yan, 2023). It can also support research writing more specifically by locating and summarizing previous research on a topic (Huang & Li, 2023), assisting students in research design and planning, and offering suggestions for new research directions (Rahman & Watanobe, 2023). Instructors can also use the deep learning capabilities of ChatGPT to develop adaptive learning activities that adjust their pedagogical approaches and complexity levels to students' L2 proficiency levels, learning needs, and preferences (Baidoo-anu & Ansah, 2023). That is, ChatGPT would analyze students' performance data, identify areas of weakness, monitor their progress, and adjust learning design as needed to address the specific learning needs of each student (Rahman & Watanobe, 2023). Also, instructors can use ChatGPT to instill innovative instructional practices, such as using flipped classroom approaches (Lage et al., 2000) by automating L2 writing instruction to be delivered out of class and using class time for intensive L2 writing practice (Rudolph et al., 2023). This way, students could learn foundational constructs outside class and maximize L2 writing practice in class, which should boost teaching and learning effectiveness (Hong, 2023). In addition, ChatGPT can support experiential learning by generating different scenarios and problem-solving activities that could foster collaborative L2 interaction between learners (Rudolph et al., 2023).

A few empirical studies have explored the potential of ChatGPT to automate instruction design (MacNeil et al., 2022; Pardos & Bhandari, 2023). MacNeil et al. (2022) reported on a case study where the Codex and GPT-3 LLMs were used to generate practice programming assignments and automatic explanations (explanation, definitions, hints, and feedback) on learners' code. The researchers were able to use LLMs to automate the processes of designing personalized assignments and offering rich explanations to students, which is typically time-consuming. In another study, Pardos and Bhandari (2023) examined the potential of ChatGPT to offer personalized learning hints to students in an algebra course. The researchers used an online GPT-powered tutoring system (GPT-3.5) to generate automatic learning hints and explored their implications for learning. Using a pre-test–post-test design, they compared the efficacy of ChatGPT-generated and human-generated hints from tutors with 77 college algebra students. The results demonstrated that both groups experienced positive learning gains. Even though these studies did not investigate L2 contexts, their findings suggest that ChatGPT can offer custom-designed and personalized instructional support to L2 writers.

Third, ChatGPT can automate the procedures of generating and implementing assessment and feedback processes (Baidoo-anu & Ansah, 2023; Hong, 2023; Rahman & Watanobe, 2023). ChatGPT can generate assessment tools, including writing prompts, graded reading passages, open-ended or multiple-choice questions, and grading rubrics that align with specific learning objectives (Kohnke et al., 2023; Qadir, 2022). Also, ChatGPT can function as an automated L2 assessment platform that can grade students' work automatically and provide students with instant feedback on their performance (Barrot, 2023), ensuring timely feedback, which is crucial for learners' development and can be challenging for instructors (e.g., Wang et al., 2023). This is especially beneficial in L2 writing practice (Rahman & Watanobe, 2023), where previous research suggests that using Grammarly to offer L2 learners feedback on their writing was an effective intervention for engaging students in writing practice (Koltovskaia, 2020). Automating assessment would also integrate assessment feedback into the learning process. For example, automated essay scoring can be used to identify patterns in students' responses and suggest revisions to students. Intelligent feedback could also support students' autonomy and improve their writing skills through error analysis: recognizing errors, identifying correct patterns, and reformulating their writing (Rudolph et al., 2023).

Several exploratory studies have demonstrated the potential of ChatGPT to automate assessment and grading (Altamimi, 2023; Jia et al., 2022; Kortemeyer, 2023; Kınık & Çetin, 2024). Altamimi (2023) evaluated the performance of ChatGPT (both GPT-3.5 and 4.0 versions) as automatic essay grading and feedback systems across several academic domains. Specifically, the study examined the accuracy, efficacy, and reliability of ChatGPT grading in comparison to human raters. The results indicated that ChatGPT offered an efficient automatic grading platform and that grading accuracy was higher for GPT-4 compared to GPT-3.5. Similarly, Jia et al. (2022) examined the feasibility of a BART-based pre-trained LLM to generate instant feedback on students' report projects. The researcher used a dataset of 484 project reports to train and test the performance of the BART-based feedback system in comparison to human experts according to specific criteria (e.g., readability, helpfulness, and specificity). Analysis of the data revealed that BART could generate feedback on par with expert human feedback. In another study, Kortemeyer (2023) examined the performance of GPT-4 on automatic grading of short-answer questions. Using standardized automatic grading evaluation benchmarks, the researcher evaluated the precision and recall of GPT-4 as a general-purpose automatic grader. The analysis revealed that the performance of ChatGPT-4 was comparable to that of earlier automatic grading models (custom-designed automatic graders) without the need for reference answers, but it underperformed compared to deep learning models that received task-specific pre-training for automatic grading, a limitation that can be rectified with additional training. In another study, Kınık and Çetin (2024) compared the scoring of 20 descriptive essays by student teachers of English and ChatGPT-3.5. They found the human raters were generally more accurate, but they also suggested that ChatGPT had great potential in this area.

Fourth, ChatGPT technology can assist instructors by automating learning management systems and reducing manual administrative interventions (Huang & Li, 2023). It could automatically post activities, grade assignments, monitor students' performance, supervise teacher assistants/tutors, identify learning challenges, and take remedial action (Huang & Li, 2023). Automation of mundane administrative tasks would save instructor's time and help them focus on improving their teaching effectiveness. At least one study has examined the potential of ChatGPT to automate administrative learning management tasks.

Hirunyasiri et al. (2023) examined the potential of GPT-4 to automate the evaluation of tutors' performance and offer them feedback. Using a framework of effective tutoring feedback criteria (e.g., timely delivery, process-focused praise, sincerity, etc.), the researchers investigated the ability of GPT-4 to understand evaluation criteria, analyze tutor-student interactions, accurately evaluate tutors' comments, and offer them feedback on their effectiveness in delivering praise to students. The researchers used 30 synthetic tutor-student dialogues to assess the ability of ChatGPT to offer feedback on tutors' praise of learners' work based on a specific set of criteria. They compared the performance of GPT-4 to that of three experienced and trained human reviewers. Analysis of the data revealed that GPT-4's performance was consistent with that of human reviewers on most criteria, but it underperformed on evaluation criteria that required integrating contextual information.

Implications of ChatGPT for ESL Writing Instruction

Some studies have examined the implications of ChatGPT for ESL writing instruction and demonstrated that it can foster L2 writing practice and development in many ways.

Affective Factors

First, a few studies have examined the effects of ChatGPT use on the affective factors in writing classes and reported that integrating ChatGPT in writing classes can foster student motivation (Fuchs, 2023) and promote writing practice. For example, in an experimental study examining the impact of ChatGPT on student anxiety in writing classrooms with 73 undergraduate English students, Hawanti and Zubaydulloevna (2023) compared the anxiety levels of writing students with and without access to ChatGPT. The study found that anxiety levels were significantly lower for the group with access to ChatGPT and concluded that integrating ChatGPT in writing classes could increase students' optimism and deliver real-time improvements in writing. In another study, Wambsganss et al. (2022) explored the impact of using an AI-powered social comparison nudge on writing students' performance with 71 writing students. The researchers divided the students into two groups; both groups received automatic feedback, but only the experimental group received a social comparison nudge. Analysis of students' essays revealed that the experimental group wrote more convincing and better-developed argumentative essays. The researchers concluded that the use of AI-generated nudges motivated students to write more effective argumentative texts by triggering basic motivational mechanisms. Similarly, Han et al. (2023) explored the use of ChatGPT in English writing courses at a Korean university with over 200 participants. The study did not reach clear conclusions about the usefulness of AI for writing, but it demonstrated that participants were highly satisfied with the experience. In another study, Ali et al. (2023) used a questionnaire to gather data about the implications of ChatGPT for English writing instruction. The results demonstrated that ChatGPT was highly useful and motivational for writing students.

Instructional Assistance

A number of studies also suggested that ChatGPT can provide instructional assistance and scaffolding to students' writing practice. For instance, Marzuki et al. (2023) used interview data to explore how four ESL teachers utilized ChatGPT in English writing classes in Indonesia. The study revealed that ChatGPT improved students' writing output by suggesting effective wording and detecting logical inconsistencies. Similarly, using questionnaire data, Wulandari et al. (2024) explored the benefits of using ChatGPT with 20 English teachers from junior high schools in Indonesia. The study revealed that ChatGPT assisted teachers in enhancing students' writing skills. In a similar vein, Ljujić et al. (2023) compared ChatGPT-written essays about multimedia with highly ranked students' essays. They found that ChatGPT could provide instructional assistance to teachers with regard to grading and instruction, and it could also scaffold students' drafting of writing projects. In another study, Kasneci et al. (2023) conducted a scoping review of the literature on the benefits of AI for writing instruction and reported that it can help instructors teach research and writing skills and train students on mundane writing processes. Similarly, Imran and Almusharraf (2023) reviewed 30 articles examining the use of ChatGPT as a writing assistant and concluded that it can assist learners in generating drafts, brainstorming ideas, and summarizing research articles. In another review, Mhlanga (2023) synthesized the findings of 23 articles on the implications of ChatGPT for education and reported that ChatGPT can enhance students' creativity and analytical thinking by allowing them to explore different writing techniques.

Feedback

In addition to its ability to support L2 writing practice and instruction, several studies have demonstrated that ChatGPT could offer L2 writing students automated feedback on their writing (Bonner et al., 2023). Dai et al. (2023) compared ChatGPT-generated and instructors' feedback on 103 students' reports. The study revealed that ChatGPT provided more detailed, readable, and consistent feedback to students. Similarly, Loem et al. (2023) evaluated the ability of ChatGPT-3 to correct grammatical errors using a variety of prompt designs. The results suggested that ChatGPT was responsive to individual needs, provided the prompts are well designed, and demonstrated that the suggested corrections were accurate and beneficial for learners. In a similar vein, Hawanti and Zubaydulloevna (2023) examined the potential of ChatGPT to support writing instruction in classroom settings. They found that ChatGPT provided instant feedback that enabled students to promptly fix their mistakes and mitigated students' concerns over errors. In another study, Üstünbaş (2024) used semi-structured interviews to explore the ability of ChatGPT to offer corrective feedback to five English writing students at a Turkish university. The study revealed that ChatGPT's feedback was perceived by learners to be useful and accessible, but it lacked the social dimension of human feedback. Also, Davishi et al. (2024) examined the quality of feedback offered by an AI platform, RIPPLE, with 1,625 students across 10 courses using 16,007 peer reviews at an Australian university. The results showed that RIPPLE offered students high-quality feedback and that the quality of their reviews dropped when access to RIPPLE was removed. In addition, Athanassopoulos et al. (2023) examined the impact of ChatGPT-generated feedback on writing quality with a small group of teenage German writing students. Analysis of participants' writing before and after ChatGPT use demonstrated that automated feedback improved participants' vocabulary choice and sentence structure. On the other hand, a study by Yoon et al. (2023) compared the quality of human and GPT-4 feedback on 50 students' essays. They found that GPT-4 feedback was generic and abstract and it failed to identify substantial issues in participants' writings. They suggested that a trained and optimized version of ChatGPT may give more useful and specific feedback.

Discussion

The present review of the available literature demonstrates that ChatGPT has impressive instructional potential that can foster L2 writing pedagogy. One of the essential trends that emerged from the literature review is the potential of ChatGPT integration to boost L2 writers' motivation (Fuchs, 2023). It appears that the availability of ChatGPT to L2 writers as an intelligent tutor can lower their affective filters, which fosters their motivation to experiment with L2 writing and use ChatGPT's feedback to develop their L2 writing skills through trial and error (Wambsganss et al., 2022). A potential explanation for this finding is that the use of machine-generated assistance and feedback removes the social awkwardness that emerges from the face-threatening practice of exposing one's weaknesses and mistakes to others. As a result, students do not experience high levels of anxiety when receiving guidance from ChatGPT (Hawanti & Zubaydulloevna, 2023), which results in a positive learning experience that fosters L2 writers' motivation to experiment with their writing skills (Wambsganss et al., 2022). Incorporating positive affect into the normally face-threatening task of practicing writing in a foreign language should result in extended L2 writing practice and considerable development in students' writing skills (Kasneci et al., 2023; Ljujić et al., 2023; Mhlanga, 2023; Wulandari et al., 2023).

Another potential of ChatGPT to support L2 writing instruction that emerged from the literature review was its ability to provide automated instructional support to L2 writers and scaffold their L2 writing practice (Ljujić et al., 2023; Wulandari et al., 2023). Specifically, the current literature demonstrates that ChatGPT can guide students' practice on standard writing processes, such as brainstorming and outlining ideas, by offering explanations, tips, and examples (Imran & Almusharraf, 2023; Kasneci et al., 2023; Marzuki et al., 2023). This potential appears to (a) offer L2 writers autonomous and individualized learning trajectories tailored to each learner's specific learning needs, (b) provide them with rich opportunities for exploring different writing styles and genres, and (c) promote extended L2 writing practice, all of which are conducive to improving their L2 writing skills and enhancing their creativity and critical thinking (Kasneci et al., 2023; Mhlanga, 2023; Wulandari et al., 2023). Also, the ability of the system to learn from human feedback (Christiano et al., 2017) could allow ChatGPT to adapt and adjust these learning trajectories to learners' needs based on performance data, ensuring the relevance of the learning experience to learners. In addition, using ChatGPT as an instructional assistant would allow writing instructors to delegate instructional design and administration tasks to ChatGPT so they could use most of their time to design and manage active learning opportunities. Freeing instructors' time from mundane administrative responsibilities would give them the time and energy to transform traditional L2 writing classrooms into communities of practices focused on the dissemination of L2 writing literacies through collaborative interaction and apprenticeship (Lave & Wenger, 1991), which was previously infeasible because instructors were hampered by their administrative workloads.

Another transformative potential that ChatGPT brings to L2 writing classrooms, according to the literature, is the ability to automate the processes of offering students on-demand, personalized feedback on their L2 writing output (Bonner et al, 2023; Dai et al., 2023; Davishi et al., 2024). Traditionally, one of the main challenges faced by L2 writers was the scarcity of personalized feedback on their writing output due to limited instructional time and resources. The ability of ChatGPT to offer automatic, instant, and frequent feedback on L2 writers' output would overcome this limitation by expanding the volume and frequency of corrective feedback they can receive on their writing. This should help them capitalize on their personal strengths and address their individual learning needs (Athanassopoulos et al., 2023) and, as a result, expedite their L2 writing skills development trajectories (Hawanti & Zubaydulloevna, 2023). Also, the automation of feedback delivery would offer L2 writers a safe environment for experimental learning (Üstünbaş, 2024) and motivate them to engage in L2 writing practice more frequently, which should boost their L2 writing skills (Kasneci et al., 2023).

Conclusion

The present study attempted to shed some light on the underexplored potentials and implications of ChatGPT for L2 writing instruction. To this end, it reviewed existing research on the potentials and implications of using ChatGPT on various aspects of writing and L2 instruction. It was hoped that such an account would (a) underline the potentials and implications of ChatGPT for L2 writing instruction, (b) promote the integration of ChatGPT in L2 writing classes, and (c) identify important directions for future research in this new research area. Analysis of the available literature to date demonstrated that ChatGPT has great potential as an instructional assistant in L2 writing classes because it could automate instructional design and assessment processes, allowing writing instructors to focus their energy on creating a community of practice centered around L2 writing skills and literacies. The literature further revealed that ChatGPT could personalize L2 writers' learning trajectories, adjust instructional resources to their individual learning needs, and adapt learning activities to their learning progress based on their performance data. The literature also demonstrated that ChatGPT can overcome instructional problems in traditional writing classes, especially limited access to corrective feedback on students' writing, as ChatGPT can produce regular and frequent feedback, which is crucial for students to develop their writing skills. On the other hand, the review demonstrated that the system is not without limitations and that it needs specialized training on instructional design and assessment processes to perform these tasks more effectively.

Practical Applications

These findings have a number of important practical applications. First, automating the design of instructional materials will change the role of the L2 writing instructor from that of a lecturer to that of a mentor. The instructor will move from the role of a teacher who imparts knowledge and offers feedback to that of a mentor who supports and monitors the learning process by making effective pedagogical decisions (based on learning analytics from AI) and guiding learners along their learning trajectories (Huang & Li, 2023). Second, given the proliferation of generative AI and the growing dependency of young learners on new technologies in communication and meaning-making activities, it is fair to assume that generative AI skills are going to be key digital literacies that college composition courses will have to integrate in order to equip learners with the necessary skills to use AI ethically and effectively to compose texts. Therefore, it is vital for L2 writing courses to include a new digital literacies component that emphasizes generative AI skills such as prompt design and fair use of AI-generated content (Baskara & Mukarto, 2023). Third, the proliferation of AI-generated content and the increasingly humanlike nature of AI-generated texts necessitates a growing emphasis on critical thinking and research skills to equip learners with the skills to locate and evaluate information in the age of AI-dominated content generation (Qadir, 2022). Fourth, the demand for integrating generative AI literacies in L2 writing courses and the potential the technology can bring to instructional contexts necessitates that L2 educators develop their digital literacy skills so they can manipulate AI tools and take full advantage of them in their classrooms (Huang & Li, 2023). This would involve acquiring the digital competencies needed to use these tools pedagogically, capitalize on their affordances, and develop a critical awareness of their limitations and risks (Kohnke et al., 2023). Finally, the rapid proliferation of generative AI technologies suggests that educational institutions will have to consider their potentials and limitations (Qadir, 2022) and develop clear guidelines for integrating these tools into their instruction and assessment to prepare students for a new world where AI-powered tools are becoming mainstream communication tools (Kohnke et al., 2023).

Limitations and Future Research Directions

The present study has a number of limitations that point to important future research directions. First, the study offered a systematic review of the literature on ChatGPT's implications for L2 writing instruction, but the empirical investigations to date have been limited and have been primarily descriptive, theoretical, or exploratory in nature, and thus, any conclusions drawn from the literature are tentative. Future research should investigate specific instructional potentials of ChatGPT in depth using rigorous quantitative and qualitative research methods. For instance, future research could explore the accuracy and reliability of ChatGPT's grading of academic essays according to a specific rubric. Second, the present study focused primarily on the positive implications of ChatGPT for instruction and did not cover the potential drawbacks of ChatGPT for L2 writing. Therefore, future research should consider how L2 writers use ChatGPT in real life and the potential drawback of generative AI for L2 writing instruction, especially with regard to violations of academic integrity. Third, the present review demonstrates that ChatGPT can offer promising instructional potential in L2 writing classes based on generic training and general skills that it elicited from general training data. In light of that, it is fair to suggest that these instructional potentials can improve dramatically with specialized training on highly curated datasets exhibiting desired instructional performance (i.e., fine-tuning). Therefore, future research should investigate the training of a specialized GPT version on L2 writing instruction and the impact of specialized training on its performance as an instructional assistant.

References

Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. (2023). Impact of ChatGPT on learning motivation: Teachers and students' voices. *Journal of English Studies in Arabia Felix*, *2*(1), 41–49. https://doi.org/10.56540/jesaf.v2i1.51

Altamimi, A. B. (2023, August). Effectiveness of ChatGPT in essay autograding. In *2023 International Conference on Computing, Electronics & Communications Engineering (iCCECE)* (pp. 102–106). IEEE. https://doi.org/10.1109/iCCECE59400.2023.10238541

Athanassopoulos, S., Manoli, P., Gouvi, M., Lavidas, K., & Komis, V. (2023). The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom. *Advances in Mobile Learning Educational Research*, *3*(2), 818–824. https://doi.org/10.25082/AMLER.2023.02.009

Ibrahim. K (2023). Using AI-based classifiers to detect AI-generated texts in ESL composition. *Language Testing in Asia,13,* 46. https://doi.org/10.1186/s40468-023-00260-2

Baidoo-anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52–62. https://doi.org/10.61969/jai.1337500n

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, *57*, 100745. https://doi.org/10.1016/j.asw.2023.100745

Baskara, R., & Mukarto, M. (2023). Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics*, *7*(2), 343–358. https://files.eric.ed.gov/fulltext/EJ1391490.pdf

Bonner, E., Lege, R., & Frazier, E. (2023). Large language model-based artificial intelligence in the language classroom: Practical ideas for teaching. *Journal of Teaching English with Technology*, *23*(1), 23–41. http://dx.doi.org/10.56297/BKAM1691/WIEO1749

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam P., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877-1901. https://doi.org/10.48550/arXiv.2005.14165

Christiano, P. F., Leike, J., Brown, T. B., Martic, M., Legg, S., & Amodei, D. (2017). Deep reinforcement learning from human preferences. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.), *Advances in Neural Information Processing Systems: Vol. 30* (pp. 4299–4307). https://proceedings.neurips.cc/paper_files/paper/2017/file/d5e2c0adad503c91f91df240d0cd4e49-Paper.pdf

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y.-S., Gašević, D., & Chen, G. (2023, July). Can large language models provide feedback to students? A case study on ChatGPT. In *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)* (pp. 323–325). IEEE. https://doi.org/10.1109/ICALT58122.2023.00100

Darvishi, A., Khosravi, H., Abdi, S., Sadiq, S., & Gašević, D. (2022, June). Incorporating training, self-monitoring and AI-assistance to improve peer feedback quality. In *Proceedings of the ninth ACM conference on learning@ scale* (pp. 35-47). https://doi.org/10.1145/3491140.3528265

Francke, E., & Alexander, B. (2019). The potential influence of artificial intelligence on plagiarism a higher education perspective. In Proceedings of the European Conference on the Impact of Artificial Intelligence and Robotics (pp. 131–140). EM Normandie Business School, Oxford. https://doi.org/10.1186/s40468-023-00260-2

Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: Is ChatGPT a blessing or a curse? *Frontiers in Education*, *8*, 1–4. https://doi.org/10.3389/feduc.2023.1166682

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digital Medicine*, *6*, article 75. https://doi.org/10.1038/s41746-023-00819-6

Han, J., Yoo, H., Myung, J., Kim, M., Lee, T. Y., Ahn, S.-Y., & Oh, A. (2023). Exploring student-ChatGPT dialogue in EFL writing education. In *Thirty-Seventh Annual Conference on Neural Information Processing Systems*. Neural Information Processing Systems Foundation. https://gaied.org/neurips2023/files/19/19_paper.pdf

Haque, M. U., Dharmadasa, I., Sworna, Z. T., Rajapakse, R. N., & Ahmad, H. (2022). *"I think this is the most disruptive technology": Exploring sentiments of ChatGPT early adopters using Twitter data*. arXiv. https://doi.org/10.48550/arXiv.2212.05856

Hawanti, S., & Zubaydulloevna, K. M. (2023). AI chatbot-based learning: Alleviating students' anxiety in English writing classroom. *Bulletin of Social Informatics Theory and Application*, *7*(2), 182–192. https://doi.org/10.31763/businta.v7i2.659

Hirunyasiri, D., Thomas, D. R., Lin, J., Koedinger, K. R., & Aleven, V. (2023). *Comparative analysis of GPT-4 and human graders in evaluating praise given to students in synthetic dialogues*. arXiv. https://doi.org/10.48550/arXiv.2307.02018

Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, *5*(1). https://jeti.thewsu.org/index.php/cieti/article/view/103

Hu, K. (2023, February 2). *ChatGPT sets record for fastest-growing user base: Analyst note*. Reuters. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

Huang, J., & Li, S. (2023). Opportunities and challenges in the application of ChatGPT in foreign language teaching. *International Journal of Education and Social Science Research*, *6*(4), 75–89. http://dx.doi.org/10.37500/IJESSR.2023.6406

Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, *15*(4), ep464. https://doi.org/10.30935/cedtech/13605

Jia, Q., Young, M., Xiao, Y., Cui, J., Liu, C., Rashid, P., & Gehringer, E. (2022). Insta-reviewer: A data-driven approach for generating instant feedback on students' project reports. *Proceedings of the 15th International Conference on Educational Data Mining* (pp. 5–16). International Educational Data Mining Society. https://zenodo.org/doi/10.5281/zenodo.6853098

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... & Kasneci, G. (2023). *ChatGPT for good? On opportunities and challenges of large language models for education* [Position paper]. Department of Educational Sciences, Technische Universität München. https://www.edu.sot.tum.de/fileadmin/w00bed/hctl/_my_direct_uploads/ChatGPT_for_Good_.pdf

Khalil, M., & Er, E. (2023). *Will ChatGPT get you caught? Rethinking of plagiarism detection*. arXiv. https://doi.org/10.48550/arXiv.2302.04335

Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, *44*, 100450.

King, M. R., & ChatGPT. (2023). Editorial: A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and Molecular Bioengineering*, *16*, 1–2. https://doi.org/10.1007/s12195-022-00754-8

Kınık, B., & Çetin, H. (2024). Human vs. AI: The use of ChatGPT in writing assessment. In Z. Çetin Köroğlu & A. Çakır (Eds.), *Fostering foreign language teaching and learning environments with contemporary technologies* (pp. 194–215). IGI Global. https://doi.org/10.4018/979-8-3693-0353-5.ch009

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Kortemeyer, G. (2023). *Performance of the pre-trained large language model GPT-4 on automated short answer grading*. arXiv. https://doi.org/10.48550/arXiv.2309.09338

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *Journal of Economic Education*, *31*(1), 30–43. https://doi.org/10.1080/00220480009596759

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Ljujić, B., Miljković, J., & Grozdić, V. (2023). ChatGPT and academic writing in higher education. In M. Radulović & M. Trajković, *Towards a more equitable education: From research to change* (pp. 104–109). Institute for Educational Research. https://reff.f.bg.ac.rs/handle/123456789/5759

Loem, M., Kaneko, M., Takase, S., & Okazaki, N. (2023). *Exploring effectiveness of GPT-3 in grammatical error correction: A study on performance and controllability in prompt-based methods*. arXiv. https://doi.org/10.48550/arXiv.2305.18156

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News*. https://dx.doi.org/10.2139/ssrn.4333415

MacNeil, S., Tran, A., Mogil, D., Bernstein, S., Ross, E., & Huang, Z. (2022, August). Generating diverse code explanations using the GPT-3 large language model. *Proceedings of the ACM Conference on International Computing Education Research, 2*, 37–39. https://doi.org/10.1145/3501709.3544280

Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, *10*(2), 2236469. https://doi.org/10.1080/2331186X.2023.2236469

Mhlanga, D. (2023). OpenAI in education, the responsible and ethical use of ChatGPT towards lifelong learning. SSRN. http://dx.doi.org/10.2139/ssrn.4354422

OpenAI. (n.d.-b). *Overview*. Retrieved January 5, 2024, from https://openai.com/chatgpt

Pardos, Z. A., & Bhandari, S. (2023). *Learning gain differences between ChatGPT and human tutor generated algebra hints*. arXiv. https://doi.org/10.48550/arXiv.2302.06871

Qadir, J. (2022). *Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education*. TechRxiv. https://doi.org/10.36227/techrxiv.21789434.v1

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pre-training*. Retrieved June 30, 2023, from https://www.mikecaptain.com/resources/pdf/GPT-1.pdf

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, *1*(8), 9. https://insightcivic.s3.us-east-1.amazonaws.com/language-models.pdf

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, *13*(9), 5783 https://doi.org/10.3390/app13095783

Rakhmonov, I. U., & Kurbonova, R. S. (2023). The pedagogical principles and effectiveness of utilizing ChatGPT for language learning. *Research and Education*, *2*(9), 226–243. https://researchedu.org/index.php/re/article/view/4899

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1), 342–363. https://doi.org/10.37074/jalt.2023.6.1.9

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. https://doi.org/10.3390/ejihpe13090140

Solaiman, I., Brundage, M., Clark, J., Askell, A., Herbert-Voss, A., Wu, J., Radford, A., Krueger, G., Kim, J. W., Kreps, S., McCain, M., Newhouse, A., Blazakis, J., McGuffie, K., & Wang, J. (2019). *Release strategies and the social impacts of language models*. arXiv. https://doi.org/10.48550/arXiv.1908.09203

Sridhar, P., Doyle, A., Agarwal, A., Bogart, C., Savelka, J., & Sakr, M. (2023). *Harnessing LLMs in curricular design: Using GPT-4 to support authoring of learning objectives*. arXiv. https://doi.org/10.48550/arXiv.2306.17459

Susnjak, T. (2022). *ChatGPT: The end of online exam integrity?* arXiv. https://doi.org/10.48550/arXiv.2212.09292

Üstünbaş, Ü. (2024). EFL learners' views about the use of artificial intelligence in giving corrective feedback on writing: A case study. In Z. Çetin Köroğlu & A. Çakır (Eds.), *Fostering foreign language teaching and learning environments with contemporary technologies* (pp. 115–133). IGI Global. https://doi.org/10.4018/979-8-3693-0353-5.ch006

Wambsganss, T., Janson, A., & Leimeister, J. M. (2022). Enhancing argumentative writing with automated feedback and social comparison nudging. *Computers & Education*, *191*, 104644. https://doi.org/10.1016/j.compedu.2022.104644

Wang, Y., Derakhshan, A., Pan, Z., & Ghiasvand, F. (2023). Chinese EFL teachers' writing assessment feedback literacy: A scale development and validation study. *Assessing Writing*, *56*, 100726. https://doi.org/10.1016/j.asw.2023.100726

Wulandari, F., Astuti, M. T., & Marhamah, M. (2024). Enhancing writing literacy teachers through AI development. *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra*, *10*(1), 246–256. https://doi.org/10.30605/onoma.v10i1.3175

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*, 13943–13967. https://doi.org/10.1007/s10639-023-11742-4

Yeadon, W., Inyang, O.-O., Mizouri, A., Peach, A., & Testrow, C. (2022). *The death of the short-form physics essay in the coming AI revolution*. arXiv. https://doi.org/10.48550/arXiv.2212.11661

Yoon, S.-Y., Miszoglad, E., & Pierce, L. R. (2023). *Evaluation of ChatGPT feedback on ELL writers' coherence and cohesion*. arXiv. https://doi.org/10.48550/arXiv.2310.06505

Young, J. C., & Shishido, M. (2023). Investigating OpenAI's ChatGPT potentials in generating chatbot's dialogue for English as a foreign language learning. *International Journal of Advanced Computer Science and Applications*, *14*(6). https://doi.org/10.14569/IJACSA.2023.0140607