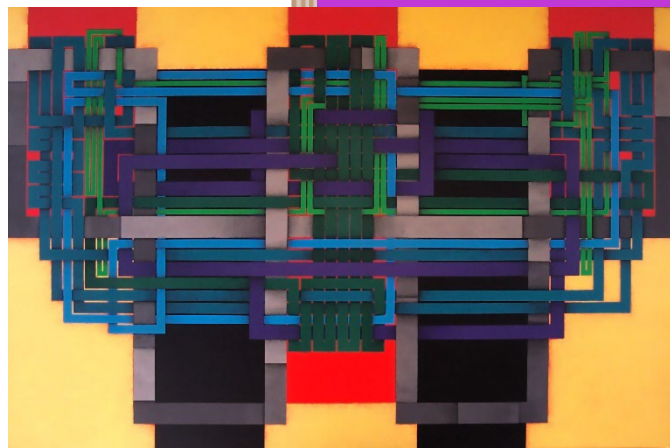




2026

EJEL Volume 24, Issue 4



Editors

Heinrich Söbke and Alessandro Pagano

Published by Academic Publishing
International Limited
Curtis Farm, Kidmore End, Nr Reading, RG4
9AY, United Kingdom
karen.harris@academic-publishing.org

eISSN: 1479-4403

Cover artwork by By Luciano de Liberato -
<https://www.flickr.com/photos/141714512@N07/27222879871/in/photostream/>,
CC BY 2.0,
<https://commons.wikimedia.org/w/index.php?curid=49087342>

EJEL Volume 24, Issue 4

Self-efficacy as a Mediator Between AI-dialogic Scaffolding, Language Anxiety, and Speaking Confidence in Saudi EFL Context <i>Shadi Majed Alshraah, Amani BinJwair, Ahmad Subhi Salem Mufleh, Ashwaq A. Aldaghri</i>	01-17
Generative AI in University Mathematics: Attitudes and Academic-Leisure Use Patterns <i>Hassan Hossein-Mohand, Hossein Hossein-Mohand, Manuel García-Alonso, María del Carmen Olmos-Gómez</i>	18-31
Pedagogical Alignment, Adaptivity, and Analytics in Intelligent Tutoring Systems: A Systematic Review <i>Ilyass Houssam, Zouhair Chiba, Mounia Miyara</i>	32-43
Modeling AI Tool Adoption in Higher Education: The Role of Authentic Learning and Prompting Competence <i>Sultan Hammad Alshammari, Mohammed Habib Alshammari</i>	44-59
Validating an Instrument of Adult Learning Needs in Online and Blended Learning <i>Anh Nguyet Diep, Khuyen Dinh, Anne-Françoise Donneau, Thanh Ngoc Phuong Phan, Minh Hien Vo, Céline Cocquyt, Maurice de Greef, Tom Vanwing, Chang Zhu</i>	60-91
Multi-level Emergency Remote Teaching Framework for K–12 Distance Education: A Stakeholder-based Qualitative Study <i>Bünyami Kayalı, Selçuk Karaman</i>	92-105
A Comprehensive Psychometric Validation of Student Online Readiness <i>Hugh Kellam, Jin Liu, Tranell Gilmore</i>	106-118
Technology-enhanced Simulation for Developing Nursing Competencies and Digital Intelligence <i>Prakob Koraneekij, Sresuda Wongwiseskul</i>	119-133
Ubuntu-informed Critical Thinking in First-year Online Accounting: Academics' Perceptions <i>Rika Dry, Annelien van Rooyen, Geesje van den Berg</i>	134-149
University Students' Use of Generative AI for English Language Learning: AI Readiness, Challenges, and Perceived Impact on Language Skills <i>Rastislav Metruk</i>	150-163

Self-efficacy as a Mediator Between AI-dialogic Scaffolding, Language Anxiety, and Speaking Confidence in Saudi EFL Context

Shadi Majed Alshraah¹, Amani BinJwair², Ahmad Subhi Salem Mufleh³ and Ashwaq A. Aldaghri⁴

¹English Department, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, Saudi Arabia

²Curriculum and Instruction, College of education, Prince Sattam Bin Abdulaziz University, Saudi Arabia

³Basic Science Department, Preparatory Year Deanship Prince Sattam bin Abdulaziz University, Saudi Arabia

⁴Department of English Language and Literature, College of Languages and Translation, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

s.alshraah@psau.edu.sa

dr.aasj2007@gmail.com

a.mufleh@psau.edu.sa (corresponding author)

Aaaldaghri@imamu.edu.sa

<https://doi.org/10.34190/ejel.24.4.4675>

An open access article under [CC Attribution 4.0](#)

Abstract: Obstacles remain in the form of inconsistent outcomes of AI applications in English as a foreign language (EFL) speaking instruction, particularly in Saudi contexts, where language anxiety and a sense of insecurity prevent learners from becoming more empowered through technological exposure. In this study, a mediator variable, self-efficacy, was postulated in the interaction between AI dialogic scaffolding, language anxiety, and speaking confidence. The current research employed a quantitative cross-sectional design, with data analyzed using partial least squares structural equation modelling (PLS-SEM) among 243 Saudi students studying at EFL universities. The findings established that AI-dialogic scaffolding had positive effects on speaking confidence and self-efficacy, and negative effects on language anxiety were very high. These relationships were partially mediated by self-efficacy, which is a critical psychological mediator. The results present a new model that incorporates technological and affective factors, offering meaningful theoretical and practical implications for the creation of AI-based language-learning contexts that facilitate psychological stability and skills acquisition. The originality of this research lies in empirically verifying complex mediating pathways in the context of Saudi EFL and extending the models of direct effects that most other researchers have previously explored. This study offers a rational framework that combines technological, cognitive, and affective features, thus contributing to the theoretical knowledge of both applied linguistics and educational technology. It extends beyond the examination of immediate impacts and shapes models of how scaffolding procedures align with the wavy paths through which they exert their influence. The studies suggest using an evidence-based approach in the Saudi context, and researchers should focus on enhancing self-efficacy to break anxiety and lack of self-confidence. Finally, the study illuminates that the true potential of AI in ed-tech is not just its ability to copy an interaction; rather, its capacity to be organized in a way that instills psychological strength and confidence in the messages it delivers. Thus, the study provides a distinct path for the evolution of a better, more holistic, and learner-focused digital language-learning environment.

Keywords: AI dialogic scaffolding, Language anxiety, Speaking confidence, Self-efficacy, Saudi EFL learners

1. Introduction

The adoption of artificial intelligence (AI) in English as a foreign language (EFL) teaching has radically changed oral communication instruction (Ekizer, 2025). Saudi Arabian learning groups often face severe challenges in achieving proficiency in speaking, with significant challenges including large-scale language anxiety (Alshakhi, 2025) and communicative confidence. Typical classroom environments do not usually provide individuals with the non-judgmental and persistent feedback required to mitigate these affective challenges. This leaves students afraid of spontaneous communication, crippling their language development and empowerment. AI dialogic scaffolding (ADS) is a transformative solution that provides adaptive and interactive scaffolding by simulating a conversation with a human (Ma, 2025). In this study, AI dialogic scaffolding is operationally defined as a pedagogical process in which an AI tool (e.g., a conversational chatbot or adaptive dialogue system) provides structured, turn-by-turn support for oral practice. This support includes: (a) modelling target utterances, (b)

offering corrective feedback on pronunciation/grammar, (c) prompting learners to elaborate or reformulate responses, and (d) gradually reducing hints as competence increases. The specific AI tool used was a custom-built dialogue system integrated into the university's learning management system. Learners engaged in five 20-minute sessions over two weeks, with each session focusing on a different everyday topic (e.g., introducing oneself, making requests). Interaction logs showed that 92% of participants completed all sessions. This operationalization distinguishes ADS from generic chatbot use or non-dialogic automated feedback.

However, the use of AI tools cannot guarantee learner empowerment. This technology is closely connected to the learner's psychological state, specifically their self-efficacy (SE). Even though language production structural support is only assisted by ADS, language anxiety (LA) continues to affect the courage to talk (SC) perniciously. The problem is that AI integration does not yield results that are quite similar; some learners find it pleasant, whereas others experience stress. What is very important is that one should know the interaction between ADS and LA that alters SC, but more to the point, SE is a very important psychological bridge in that interaction. Without a clear view of such mediation processes, the future of AI-driven learner empowerment can be regarded as a mere theoretical construct, and Saudi EFL learners have little ground to resist the fear they experience when speaking online. This study addresses this pressing question and concerns the multidimensional relationship between technological support and mental stability. Existing studies have either (a) examined AI exposure without measuring psychological mediators (e.g., Ekizer, 2025), (b) studied self--Alshakhi, 2025), or (c) used qualitative designs that cannot quantify indirect effects (Zhang, 2025). No study has modelled self-efficacy as a simultaneous mediator between AI scaffolding and language anxiety on speaking confidence in a Saudi EFL sample using PLS-SEM. Thus, it remains unknown whether AI's benefits of AI operate through self-efficacy or only directly, and whether anxiety damages confidence partly by eroding self-belief. This study closes this gap by testing the specific indirect paths shown in Figure 1.

Although educational technology is currently gaining popularity, many research gaps remain in understanding the multifaceted interactions among AI dialogic scaffolding, language anxiety, and speaking confidence through the self-efficacy lens. The overall positive effects of applying AI to language learning have been widely reported in the existing literature (Ekizer, 2025); however, little empirical evidence is available on specific dialogic scaffolding processes in the context of Saudi EFL. Although Alshakhi (2025) emphasized the importance of anxiety among Saudi learners, they did not delve deeply into the role of AI-guided scaffolding procedures that are likely to reduce the burden. Also, even though recent studies indicate that AI tools have the potential to enhance speaking fluency as proposed by some studies (Wang and Wang 2025), the fact that self-efficacy has a mediating effect on the connection between these two is disregarded or is considered briefly (Yan et al., 2025). Earlier research has mostly focused on two extremes (Ebadi et al., 2025) and overlooked the subtle psychological mechanisms that guide translation to empower learners (Dai et al., 2025). In addition, the precise effects of interactive situations and personal feedback on self-evaluation remain unexplored (Cui, Yang and Xu, 2025). Most studies employ qualitative or simple correlational designs, creating a methodological gap and necessitating more rigorous structural modelling (Zhang, 2025).

This study aimed to examine the mediating effect of self-efficacy on the connection between AI dialogic scaffolding, language anxiety, and speaking confidence in Saudi EFL learners. This study examined the interactions among these variables to empower learners using a partial least squares structural equation modelling (PLS-SEM) approach. This study is of great value in two ways: it provides a critical theoretical and methodological contribution to the dynamics of Applied Linguistics. In principle, it is a further extension of Social Cognitive Theory, theorizing AI dialogic scaffolding as a contemporary environmental factor that restructures the self-efficacy of EFL learners (Ren, Stephens and Lee, 2026). It offers a new framework for the interaction between digital scaffolding and affective variables, such as language anxiety and communicative outcomes. Methodologically, the researchers used PLS-SEM, which provides strong statistical support for complex mediating mechanisms that are inaccessible to traditional regression modelling.

2. Literature Review

2.1 Theoretical Foundation

The theoretical base draws on sociocultural, social cognitive, and affective perspectives to explain how students' conditions and psychology influence oratory confidence in AI-facilitated EFL settings. The conceptualization of speaking confidence in Figure 1 proposes that AI dialogic scaffolding and language anxiety will alter the proximal communicative outcome of speaking confidence, and that the central mediation between the two will be self-efficacy as a belief system. The application of AI dialogic scaffolding is justified by sociocultural theory, which conceptualizes learning as a mediated activity in which interaction, feedback, and regulation facilitate the

internalization of communicative competence, whether mediated by humans or technology (Vygotsky, 1978; Lantolf et al., 2014). From this perspective, AI-free dialogic scaffolding entails a teaching instruction system that emphasizes meaning-making, learner agency, and gradual advancement of oral skills. The explanatory logic for attributing evidence of speaking confidence from these instructional experiences to social cognitive theory is that it underlies self-efficacy as an agentic judgment that controls effort, persistence, and communicative risk-taking (Bandura, 2018). In this regard, scaffolded interaction can indirectly improve speaking confidence by enhancing learners' perceived communicative ability. This explanation is supplemented by the affective perspective, which explains the restraining effect of language anxiety, which disrupts attentional control and leads to negative self-assessment in an interactive situation (Gegersen & MacIntyre, 2020). Thus, it is hypothesized that anxiety has a direct and indirect impact on speaking confidence by undermining self-efficacy beliefs. Figure 1 combines all these theoretical perspectives into a consistent causal framework in which self-efficacy, in interaction with instructional support and emotional constraints, results in variability in speaking confidence. This framework is reflected in current theoretical developments and will provide a rigorous foundation for testing the hypotheses suggested in technology-based language-learning settings. While the present study emphasizes recent work from 2024 to 2026, the role of self-efficacy in language learning was established decades earlier (Woodrow, 2011). The current model extends these foundational studies by substituting human scaffolding with AI-dialogic scaffolding. Consequently, the contribution is not the discovery of new psychological mechanisms but rather the application of well-established theory to an emerging technological context. Sociocultural theory explains how AI scaffolds learning (mediation). Social cognitive theory explains why self-efficacy changes (mastery experiences and vicarious observation). Affective perspective explains when anxiety overrides both (attentional control, negative self-talk). The novel integration in Figure 1 is that self-efficacy is positioned as the common pathway: AI operates through it (sociocultural → cognitive), and anxiety operates through it (affective → cognitive). This synthesis predicts that any intervention (AI or anxiety reduction) will have a limited effect on speaking confidence unless it first changes self-efficacy. Figure 1 is derived from Bandura's (2018) triadic reciprocal determinism (person – environment – behaviour). AI dialogic scaffolding represents the environmental factor; self-efficacy is the *personal* belief; speaking confidence and language anxiety are behavioural/affective outcomes. The specific arrows are based on: (a) Vygotsky's (1978) mediation principle (AI → self-efficacy), (b) Gegersen and MacIntyre's (2020) affective filter model (anxiety → speaking confidence), and (c) previous meta-analyses (Ren, Stephens and Lee, 2026) showing self-efficacy as a mediator. No prior study has tested this exact configuration in Saudi EFL with AI scaffolding.

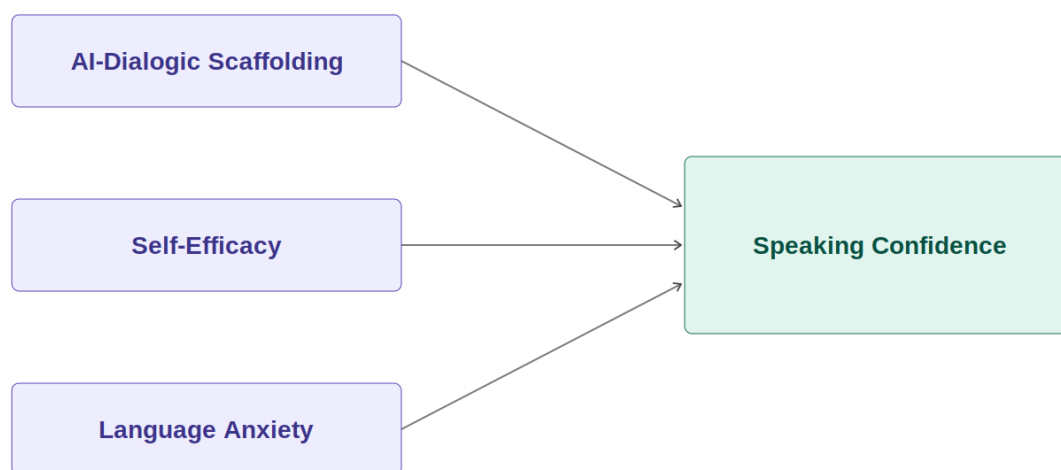


Figure 1: Research Framework

2.2 Previous Studies and Hypotheses Development

Effect of AI-Dialogic Scaffolding on Speaking Confidence

Recent literature suggests the growing importance of the AI dialogic scaffold in helping English as a Foreign Language (EFL) students feel more confident in speaking their language, often in the context of social cognitive theory. This theoretical approach assumes that learning occurs as a social process involving dynamic reciprocal interactions among the person, environment, and behavior. Interactive and personalized feedback from AI

dialogic scaffolding helps develop an environment in which observational learning and mastery experiences are inviting and can contribute to the formation of self-efficacy as the central construct of social cognitive theory, thereby directly affecting speaking confidence. For example, Cui, Yang and Xu (2025) found that AI can significantly improve EFL learners' speaking, provided that a system is used as a communicative scaffold and co-constructs with scenario-specific interactions. Similarly, Ma and Chen (2025) observed that language games involving AI did not demotivate or decrease the interest of Chinese EFL learners, which indirectly resulted in the acquisition of English and consequently increased confidence in speaking. Wang and Wang (2025) narrowed the study to the use of individually interactive scaffolding using AI and its impact on the performance of secondary school students in speaking, setting goals, and self-critical reflection, which are the main components of confidence in speaking. Ebadi et al. (2025) also noted the benefits of AI-enhanced scaffolding procedures in increasing confidence during the speaking process, Tian et al. (2026) investigated the effects of a GenAI-assisted dialogic model of peer feedback on self-efficacy among EFL writers, with the implication that it propagates to speaking confidence. Finally, Grab (2025) demonstrated that ESL undergraduates learned more through AI chatbots with built-in practice to improve speaking. These studies lead to the conclusion that AI dialogic scaffolding is a beneficial, interactive learning environment that enhances EFL learners' speaking confidence, skill acquisition, and self-belief development. Although we use causal language ("influences," "affects") for consistency with social cognitive theory, all hypotheses refer to statistical associations from cross-. Based on these findings, we formulated our first hypothesis as follows:

H1: AI-Dialogic scaffolding positively influences Speaking Confidence in Saudi EFL learners.

Effect of Language Anxiety on Speaking Confidence

Language anxiety significantly influences the development of speaking confidence among EFL learners, a phenomenon frequently attributed to sociocultural theory, which emphasizes social interaction and cultural context in the learning process. Too much anxiety may construct a filter of emotions, making learners unable to take part in communicative practices and speaking, thereby negatively affecting their confidence. For example, Lin et al. (2025) validated the negative predictive role of foreign language anxiety on learners' willingness to communicate, which is a direct antecedent of speaking confidence. Septiawan et al. (2025) outlined the extent, attitudes, and perceptions of speaking anxiety among university students, indicating its ubiquity. Alharbi, (2025) established a strong negative correlation between EFL anxiety and academic performance in speaking. Cancino and Cabello (2024) investigated the links between language learning strategies and self-efficacy and implicitly related the negative effects of anxiety on strategy use and future confidence. The triarchic nature of the relationship between L2 learners' cognition, emotion, and language performance was also explored by Ma (2022). A crucial point in these studies is how the concept of anxiety (emotion) interferes with performance (i.e., speaking). Radjuni, Sahraeny and Latief (2025) analyzed the correlation between English-speaking self-efficacy and students' English-speaking, in which anxiety was identified as an inhibitor of self-efficacy. As these studies have consistently shown, language anxiety is a significant barrier that lowers the speaking confidence of EFL learners and hinders their participation and performance. Drawing on the negative role of anxiety, we formulated hypothesis two as follows:

H2: Language Anxiety negatively influences Speaking Confidence in Saudi EFL learners.

The relationship based on social cognitive theory is that of English as a Foreign Language (EFL) learners. This model assumes that mastery experiences and verbal persuasion foster self-efficacy through a triadic interplay. Empirical observations of recent times suggest that scaffolds based on AI offer a framework for such experiences, a low-stakes, low-distraction environment in which they thrive. For example, Almusharraf et al. (2025) showed that generative AI chatbots help learners develop a sense of competence through immediate, non-judgmental feedback that builds on an understanding of communicative power. Similarly, Apriani et al. (2024) found that self-efficacy, a key indicator of the development of interactive support for AI, increased substantially after artificial task-based support was provided. This was also quantified by Ren, Stephens and Lee, (2026), who reported that academic self-belief in AI-mediated learning environments was significantly positive. Moreover, Xiao et al. (2025b) argue that AI-guided systems promote cognition, thereby reducing cognitive overload and increasing self-efficacy when performing challenging tasks. According to Zhang (2026), AI dialogue provides a safe environment for mastering verbal communication. In contrast, Wang and Wang (2025) stated that AI interactive scaffolding directly facilitates the establishment of goals and self-assessment, which are key components of effective behavior. Considering the central role of self-efficacy in social cognitive theory, we formulated hypothesis three as follows:

H3: AI-Dialogic Scaffolding positively influences Self-Efficacy in Saudi EFL learners. These results highlight the empowering nature of technology.

Effect of Language Anxiety on Self-Efficacy

The effects of language anxiety on the self-efficacy of EFL students are also severe, and they are well explained with reference to the social cognitive theory, which focuses on the mutual interaction among personal, behavioral, and environmental factors. Extreme anxiety causes a person to lose trust in their abilities, especially in performance-related activities, such as speaking a foreign language. This loss of self-efficacy may lead to avoidance tendencies, which, in turn, limits opportunities to develop skills and build confidence in their abilities. In this manner, for example, the study by Heydarnejad (2025) on the roles of academic mindfulness and teacher support in educational self-efficacy implicitly implied that anxiety, as an emotional constraint, would harm all of them. Similarly, Pan and Li (2025) observed that acceptance of generative AI was associated with EFL learners' well-being and self efficacy. Therefore, negative emotional states, such as anxiety, have the opposite effect. The research published by Li et al. (2024) on increasing learners' confidence in English learning emphasized that highly effective learners are more motivated, which, in turn, would decrease anxiety and motivation through lowered self-efficacy. Wang (2024) developed and confirmed scale measures of speaking self-efficacy, and it is important to note that perceived capability plays an important role in language production and is directly affected by anxiety. Moreover, the connection between English-speaking self-efficacy and performance perception indicates that factors such as anxiety can reduce self-efficacy (Radjuni, Sahraeny and Latief, 2025). Finally, the systematic review by Wang and Wang (2025) on anxiety and self-efficacy among English language learners further solidifies that anxiety is a major hindrance to the development of powerful self-efficacy. Altogether, these studies show that language anxiety plays a crucial role in reducing EFL learners' self-efficacy, as it creates self-doubt and prevents them from engaging in language learning. Given the inverse relationship between anxiety and perceived competence, we formulated hypothesis four as follows:

H4: Language Anxiety negatively influences Self-Efficacy in Saudi EFL learners.

Effect of Self-Efficacy on Speaking Confidence

Self-efficacy is a fundamental construct of social cognitive theory that determines EFL speaking confidence. This theory assumes that the beliefs individuals hold about their ability to plan and implement the courses of action needed to achieve their goals provide the basis for their motivation, behavior, and emotional states. Increased self-efficacy in speaking results in greater effort, persistence, and resilience in the face of challenges, thereby directly increasing confidence. Wang (2024) formulated and tested scales of speaking self-efficacy, explaining that it is directly connected to perceived speaking ability and confidence. Li et al. (2024) found that positive EFL students are more motivated, and their motivation is closely intertwined with their confidence in their ability to speak. According to Chang, Zhou and Zhang (2024), EFL learners' participation in speaking activities is positively affected by their task-specific self-efficacy. When learners have greater confidence in their speaking skills, they tend to engage more actively, which enhances their speaking confidence. A direct test of the relationship between English-speaking students' self-efficacy and English-speaking performance was conducted by Radjuni, Sahraeny and Latief (2025), who found a positive direct correlation between these two variables. This notion is also supported by Wang and Wang (2025) on anxiety and self-efficacy among English language learners, as self-efficacy levels are inversely related to anxiety and confidence in speaking. Furthermore, the article by Ma and Chen (2025) on the contribution of human touch to AI-optimized language learning and self-efficacy implicitly suggests that self-efficacy is a major factor in explaining general language skills and confidence, based on research on the role of human touch in language learning. This study has consistently reaffirmed that a strong sense of self-efficacy is the highly desired element that can assist EFL students in building their speaking confidence, leading them to become more active and consistent in speaking. Considering the central role of self-efficacy in social cognitive theory, we formulated hypothesis five as follows:

H5: Self-efficacy positively influences Speaking Confidence in Saudi EFL learners.

Self-Efficacy as a mediator

The relationship between the AI dialogue scaffold, language anxiety, and speaking confidence was mediated by self-efficacy, as best explained by social cognitive theory. The theory stresses that people's perceptions of their abilities are influential factors that shape their motivation, behavior, and emotional reactions. When applied to EFL learning, AI dialogic scaffolding could enhance speaking confidence, both directly and indirectly, by supporting learners' self-efficacy. For example, Pan and Li (2025) proposed that acceptance of generative AI is correlated with the well-being and self-efficacy of EFL learners, suggesting that a positive attitude toward AI

tools could boost self-belief. According to Cui, Yang and Xu (2025), AI, as a communicative scaffold, enhances speaking skills, and the subsequent improvement in speaking skills leads to a stronger sense of self-efficacy, which, in turn, boosts confidence. Moreover, self-efficacy tends to mediate the negative effects on speaking confidence in language anxiety. When language anxiety is high, it takes away the self-efficacy of the learners, such that they feel incompetent to speak effectively, and hence this reduces their speaking confidence. Alharbi (2025) indicated good negative relationships between EFL anxiety and digital self-efficacy, which means that anxiety directly diminishes self-belief. Heydarnejad (2025) investigated whether academic mindfulness and teacher support can affect educational self-efficacy, suggesting that educational self-efficacy would be positively influenced by factors that reduce anxiety. The role of anxiety and self-efficacy among English language learners aligns with the findings of Wang and Wang (2025) who revealed that higher self-efficacy is associated with lower anxiety and confidence. These studies indicate that self-efficacy is an influential psychological process and that AI-dialogic scaffolding and language anxiety can mediate the effect of EFL learners' speaking confidence. Integrating the previous paths, we formulated hypothesis six on mediation as follows:

H6: Self-efficacy mediates the relationship between AI-Dialogic scaffolding, language anxiety, and speaking confidence in Saudi EFL learners.

3. Research Methodology

3.1 Participants and Setting

This study used a quantitative design and a cross-sectional survey to examine the correlations among ADS, LA, SE, and SC among Saudi EFL learners. The study population consisted of students studying English as a Foreign Language at the university level in Saudi Arabia, an industry that enrolled approximately 2.2 million students in 2024 (Alanazi, 2024). To determine the sample size, a table based on Krejcie and Morgan's (1970) method was used, and a sample of 384 participants was selected to ensure statistical representativeness of the sample. A convenience (accessible) sampling method was used, meaning that the researchers invited all EFL students in the six cooperating universities who were present during the data-collection period. The target of 384 was based on Krejcie and Morgan's table for a population of 2.2 million, but this table assumes a simple random sampling. Because our sample was non-probability, the sample achieved ($N = 243$) was not statistically representative of all Saudi EFL students in the strict sense. We report a 63.3% response rate, but the target was 384 (Krejcie & Morgan for 2.2 million), and we obtained 243 valid responses (63.3% of the target). A total of 243 valid responses were obtained out of 384 questionnaires sent, with a response rate of 63.28%. Table 1 shows the Saudi EFL participants in this study.

Table 1: Demographic Profile of the Respondents ($N = 243$)

Variable	Category	Frequency	Percentage (%)
Gender	Male	108	44.4
	Female	135	55.6
Age Group	18–20 years	99	40.7
	21–23 years	132	54.3
	24 years and above	12	4.9
Year of Study	Foundation Year	40	16.5
	Year 1	48	19.8
	Year 2	55	22.6

Variable	Category	Frequency	Percentage (%)
	Year 3	52	21.4
	Year 4	48	19.8
Institution Type	Public University	168	69.1
	Private University	75	30.9
Field of Study / Major	English Language / Applied Linguistics	84	34.6
	Education (EFL/TESOL)	53	21.8
	Other Disciplines	106	43.6
English Proficiency Level	Beginner	40	16.5
	Intermediate	101	41.6
	Upper-Intermediate	47	19.3
	Advanced	55	22.6
Experience with AI Tools	No prior experience	60	24.7
	Limited experience	91	37.4
	Moderate experience	60	24.7
	Extensive experience	32	13.2
Experience with Dialogic Scaffolding Activities	None	52	21.4
	Occasional	98	40.3
	Frequent	93	38.3
Confidence in Speaking Activities (Self-Rated)	Low	50	20.6
	Moderate	122	50.2

Variable	Category	Frequency	Percentage (%)
	High	71	29.2

The demographics of the 243 Saudi EFL participants in this study are presented in Table 1, which reflects the study's results. It was mainly a women's sample (55.6%), with the majority (54.3%) in the 21-23 age group, which is typical of university enrolment (Alanazi, 2024). The respondents were spread across all academic years, with the highest percentage (22.6) in Year 2, and most were from public universities (69.1). Interestingly, although 56.4% of the participants were English majors, a significant number (43.6%) were from other majors, and most of them showed a general interest in EFL. Regarding proficiency, most participants rated themselves as intermediate (41.6) or advanced (22.6). The general experience with AI tools and dialogic scaffolding was also ambivalent; 37.4% had little experience with AI, and 40.3% had a low frequency of scaffolding. Overall, there was a moderate (50.2 percent) and high (29.2 percent) proportion of 70.8 percent speaking with confidence, 20.6 percent speaking with low confidence, and progressing to affective in the Saudi environment (Alshakhi, 2025). This variety of students demonstrates that this research is applicable to students from diverse backgrounds and experiences.

3.2 Data Collection

The study items were a structured questionnaire containing adjusted field questions from current scales. The ADS was assessed using six items based on Tuan (2025), focusing on AI-based pedagogical support. LA was measured using the 5-item Foreign Language Classroom Anxiety Scale (Horwitz, Horwitz and Cop, 1986). SE was measured using five items modified from the Speaking Skills Self-Efficacy Beliefs questionnaire (Wang and Sun, 2024), and SC was assessed using seven adapted Griffiee (1997) items. A pretest with 5 participants (different from the main sample) was used only to check for obvious wording errors and survey flow. This was followed by a pilot study with 30 participants (10% of the target sample), who completed the same survey and then were interviewed for 5 minutes about item clarity. Based on pilot feedback, two items (LA4 and SC6) were reworded slightly (e.g., "I feel anxious" → "I feel nervous"). The pilot data were not included in the main analysis. Construct validity was further assessed via confirmatory factor analysis (all loadings > 0.77; Table 3). No items were deleted. The original scales were in English. They were translated into Arabic by two bilingual applied linguists, then back-translated by two independent translators. Discrepancies (3 out of 23 items) were resolved by a third expert. For the Saudi context, we replaced culturally specific references (e.g., "speaking in front of the whole class" became "speaking in front of a group of 5–6 Saudi peers"). No items were deleted; all were retained to maintain comparability with the original validation studies. A pilot investigation involving 30 learners revealed that Cronbach's alphas for all constructs exceeded 0.70, thereby validating the internal reliability and ruling out the need for further adjustments. After conducting the pilot study, most statements underwent a review and refinement process to more accurately align with the academic environment of Saudi EFL learners, guaranteeing that the items were both linguistically suitable and culturally pertinent for the intended audience.

3.2.1 Procedure and ethical approval

Data were collected over 2 months, consecutive days (14th January–16th March 2026) using the university's licensed Qualtrics platform. Participants accessed the survey via a unique anonymous link distributed through course LMS announcements and WhatsApp groups of six English departments. Inclusion criteria were: (a) enrolled in a mandatory EFL speaking course, (b) had not previously used an AI dialogue system for speaking practice, (c) age ≥18 years, and (d) provided informed consent. Exclusion criteria: (a) incomplete responses (>10% missing), (b) straight-lining (identical answers for >80% of items), and (c) completion time <3 minutes (determined by pilot testing). The survey remained open until 243 valid responses were obtained after cleaning, a process that took 2 months. (a) Our model includes both direct and indirect paths (mediation), which PLS-SEM handles without distributional assumptions. (b) We prioritize prediction ($Q^2 = 0.622–0.700$) over covariance-based fit indices because our goal is to identify which variables predict speaking confidence for intervention design. (c) Preliminary Shapiro-Wilk tests indicated that 7 of 23 items violated normality assumptions ($p < 0.05$), rendering maximum likelihood estimation inappropriate. (d) The sample size ($N = 243$) exceeds the 10-times rule for the most complex regression (max 3 predictors → minimum 30 cases). Thus, PLS-SEM is methodologically justified, not merely convenient. The researchers received ethical approval from the Institutional Review Board at Prince Sattam bin Abdulaziz University (Approval No. SCBR-399/2026; January 2026). This research adhered to the Helsinki Declaration of 1964; it supported informed consent and did not violate confidentiality. All

participants were informed of their right to withdraw, and their data was handled securely to ensure privacy throughout the research.

3.3 Data Analysis

Data collection was conducted electronically using the university's licensed Qualtrics platform to ensure the participants' anonymity. This platform was chosen because it automatically anonymizes IP addresses and prevents the collection of identifying metadata, thereby reducing social desirability bias and encouraging honest responses regarding language anxiety, which is a sensitive topic in Saudi EFL contexts. Data analysis was performed using Partial Least Squares Structural Equation Modelling (PLS-SEM) with SmartPLS 4. PLS-SEM was selected for three reasons: (a) the research prioritizes the prediction of speaking confidence ($Q^2 = 0.700$) over covariance-based fit indices; (b) preliminary normality tests revealed that seven out of 23 items violated univariate normality (Shapiro-Wilk $p < 0.05$), making maximum likelihood estimation inappropriate; and (c) the model includes both direct and indirect mediation paths, which PLS-SEM estimates without distributional assumptions (Hair et al., 2022). This approach is methodologically appropriate for complex mediation in applied-linguistics research.

4. Results of the Study

Table 2 presents the results of the exploratory factor analysis for common method biases. The first accounts for 48.831 percent of the overall variation, with an initial eigenvalue of 14.161, which falls short of the standard 50 percent criterion (Hair et al., 2022). This implies that, although we do find common method variance, as in the case of simple single-source and self-report survey data, it is not high enough to seriously affect the findings. The fact that three further significant elements (eigenvalues greater than one) accounted for 67.111 percent of the variance also shows that the constructs were latent and not the product of a single methodological variable. Thus, despite its presence, common method bias was not considered a serious issue in this study.

Table 2: Common Method Bias

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	14.161	48.831	48.831	14.161	48.831	48.831	8.546	29.470	29.470
2	2.924	10.083	58.913	2.924	10.083	58.913	5.081	17.522	46.992
3	1.292	4.456	63.370	1.292	4.456	63.370	3.281	11.314	58.306
4	1.085	3.741	67.111	1.085	3.741	67.111	2.554	8.805	67.111

The findings from the model measurement are presented in Table 3, which indicates high reliability and convergent validity for all constructs. All item loadings were beyond the acceptable limit of 0.70 (Hair et al., 2022), ranging from 0.770 to 0.889. Internal consistency was good, with Cronbach's alpha values exceeding 0.80 (AI-Dialogic scaffolding $\alpha=0.802$, Language Anxiety $\alpha=0.856$, Speaking Confidence $\alpha=0.817$, self-efficacy $\alpha=0.822$). The reliability of the scales was further supported by composite reliability (rho_c) scores ranging from 0.825 to 0.841. The criteria used to statistically define convergent validity were an Average Variance Extracted (AVE) for each construct higher than 0.50, with an average of 0.668 to 0.762 (Hair et al., 2019). These findings suggest that the scaled tools, including ADS (Tuan, 2025), LA (Horwitz, Horwitz and Cop, 1986), SE (Wang and Sun, 2024), and SC (Griffiee, 1997), worked well in the Saudi EFL context and offered psychometrically adequate tools for testing the structural model. The six ADS items (e.g., "The AI dialogue system gave me step-by-step hints when I struggled") directly reflected the operationalization above. No other AI tool (e.g., a generative chatbot without scaffolding or an automated grammar checker) was used. This distinguishes ADS from generic "AI-assisted interaction".

Table 3: Measurement Model

Constructs	Factor Loadings	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
AI-Dialogic Scaffolding		0.802	0.813	0.825	0.672
ADS1	0.833				
ADS2	0.821				
ADS3	0.777				
ADS4	0.811				
ADS5	0.836				
ADS6	0.837				
Language Anxiety		0.856	0.803	0.827	0.718
LA1	0.826				
LA2	0.819				
LA3	0.845				
LA4	0.881				
LA5	0.865				
Speaking Confidence		0.817	0.819	0.834	0.668
SC1	0.847				
SC2	0.779				
SC3	0.845				
SC4	0.797				
SC5	0.770				
SC6	0.877				
SC7	0.801				
Self-Efficacy		0.822	0.823	0.841	0.762
SE1	0.860				
SE2	0.889				
SE3	0.863				
SE4	0.881				
SE5	0.871				

Table 4 shows the discriminant validity measurement based on the heterotrait-monotrait (HTMT) ratio and predictive model performance. Discriminant validity is supported by HTMT values for all constructs, which are all below the conservative threshold of 0.85 (Hair et al., 2022). Namely, the most significant correlation was between LA and SC at 0.798, indicating a strong but neither close nor weak relationship. The model is quite explanatory, as SE and SC have R-squared values of 0.630 and 0.711, respectively, indicating that the predictors explain a significant percentage of the variance in these endogenous variables and that the model satisfies the theoretical assumptions of Social Cognitive Theory (Bandura, 2018). Moreover, the Q-squares (0.622 for SE and 0.700 for SC) are also greater than 0, indicating that the model has predictive relevance (Hair et al., 2019). The overall findings support the uniqueness of the measured constructs and the model's ability to test hypotheses.

Table 4: Hetetrait-monotrait Discriminant Validity

Constructs	R-square	Q-square	ADS	LA	SC	SE
ADS						
LA			0.456			
SC	0.711	0.700	0.558	0.798		
SE	0.630	0.622	0.546	0.641	0.777	

As shown in Table 4a, all items loaded higher on their intended construct than on any other construct, with a minimum difference of 0.10 or more. For instance, ADS1 loaded at 0.833 on ADS but cross-loaded at only 0.368 on LA, 0.433 on SE, and 0.417 on SC. Similarly, LA3 loaded at 0.845 on LA, with cross-loadings on SC (0.488), SE (0.661), and ADS (0.306), which were all substantially lower. The highest cross-loading across all 23 items was 0.714 (LA4 on SE), which was well below the conventional threshold of 0.70 (Hair et al., 2022). No item exhibited a cross-loading that exceeded its primary load. This pattern confirms that each construct is empirically distinct from the others, supporting discriminant validity at the item level. The complete cross-loading matrix is shown in Table 4a.

Table 4a: Cross Loadings

Items	ADS	LA	SC	SE
ADS1	0.833	-0.368	0.417	0.433
ADS2	0.821	-0.335	0.438	0.412
ADS3	0.777	-0.378	0.435	0.457
ADS4	0.811	-0.287	0.397	0.363
ADS5	0.836	-0.341	0.400	0.398
ADS6	0.837	-0.309	0.408	0.384
LA1	-0.363	0.826	-0.566	-0.636
LA2	-0.367	0.819	-0.473	-0.617
LA3	-0.306	0.845	-0.488	-0.661
LA4	-0.355	0.881	-0.445	-0.689
LA5	-0.359	0.865	-0.691	-0.646
SC1	0.444	-0.503	0.847	0.620
SC2	0.400	-0.448	0.779	0.560
SC3	0.455	-0.475	0.845	0.596
SC4	0.420	-0.446	0.797	0.570
SC5	0.392	-0.425	0.770	0.529
SC6	0.419	-0.343	0.877	0.547
SC7	0.379	-0.331	0.801	0.568
SE1	0.399	-0.659	0.597	0.860
SE2	0.446	-0.676	0.649	0.889
SE3	0.445	-0.659	0.606	0.863
SE4	0.441	-0.714	0.672	0.881
SE5	0.451	-0.639	0.599	0.871

Table 5 presents the Variance Inflation Factors (VIFs) to assess multicollinearity among the predictor constructs in the structural model. None of the VIF values exceeded the generally established level of 5, and it was even stricter at 3 (Hair et al., 2019). In the case of SC prediction, the VIFs were 1.337, 2.442, and 2.703 for ADS, LA, and SE, respectively, indicating no multicollinearity. For the prediction of Self-Efficacy, the VIFs were 1.205 for ADS and 1.205 for LA. These low values suggest the absence of harmful multicollinearity, indicating that the

constructs of the predictors are well dispersed and that the individual effects of these predictors on the dependent variables (SC and SE) can be estimated consistently and without bias due to high intercorrelations.

Table 5: Variance Inflation Factor (VIF)

Constructs	SC	SE
ADS	1.337	1.205
LA	2.442	1.205
SE	2.703	

The results of the path analysis for the test of direct relationships in the hypothesized model are presented in Table 6 and Figure 2. The first hypothesis (H1) posits that ADS is positively associated with SC. This hypothesis is supported by a significant positive path ($\beta = 0.176$, $p < 0.001$). Sociocultural theory is consistent with this observation, as it assumes that tools and social interaction mediate learning (Lantolf et al., 2014). The AI tool served as a mediating artifact, providing organized, interactive assistance that could internalize communicative competence and, therefore, directly increase confidence, as demonstrated in comparable situations by Wang and Wang (2025). Although $\beta = 0.176$ for ADS $\rightarrow$ SC is small ($f^2 = 0.081$, below Cohen's 0.15 for medium), it is meaningful because (a) it operates alongside a very large anxiety effect ($\beta = -0.640$) and (b) the indirect effect (0.030) represents an additional 17% of the direct effect. In educational terms, Hypothesis H2, which states that LA has a negative impact on SC, is strongly supported by a significant negative path ($\beta = -0.640$, $p = 0.001$). Such a large contribution can be explained through the affective perspective of sociocultural and cognitive paradigms, with anxiety serving as a filter that disrupts performance and self-assessment (Gregersen and MacIntyre, 2020), a claim confirmed by other researchers in Saudi settings (Alharbi, 2025). The direct effect of AI on confidence ($\beta = 0.176$) was weaker than expected, given that prior meta-scaffolding procedures could yield stronger effects. Alternatively, the strong anxiety effect ($\beta = -0.640$) may suppress the AI- In H3, the impact of ADS on SE was positive ($\beta = 0.221$, $*p = .001$). This is because the social cognitive theory acknowledges mastery experiences and supportive feedback as sources of self-efficacy (Bandura, 2018). The AI scaffold facilitated such experiences, increasing learners' belief in their capabilities, as observed by Ren, Stephens and Lee (2026). The negative effect of LA on SE (H4) is also strongly supported (H4: $\beta = -0.677$, $p < 0.001$). This is explicated by social cognitive theory, which holds that anxiety weakens cognitive processing and the motivational beliefs required for strong self-efficacy (a point Heydarnejad (2025) observes). Lastly, H5, which suggests that SE has a positive effect on SC, is supported ($\beta = 0.137$, $*p = 0.024$). This direct relationship is key to social cognitive theory, in which efficacy beliefs determine effort and perseverance in activities such as speaking (Bandura, 2018), and was confirmed in EFL learning by Wang, Zhao and Li, (2024).

Table 6: Path Analysis

Hypotheses	Path Analysis	Beta	STDEV	T statistics	C.I. Bias			f-square	P values
					Bias	2.5%	97.5%		
H1	ADS $\rightarrow$ SC	0.176	0.039	4.519	0.001	0.093	0.248	0.081	0.000
H2	LA $\rightarrow$ SC	-0.640	0.054	11.953	-0.002	-0.736	-0.529	0.582	0.000
H3	ADS $\rightarrow$ SE	0.221	0.045	4.934	-0.000	0.139	0.313	0.110	0.000
H4	LA $\rightarrow$ SE	-0.677	0.035	19.425	-0.001	-0.736	-0.596	0.726	0.000
H5	SE $\rightarrow$ SC	0.137	0.061	2.261	-0.002	0.019	0.263	0.024	0.024

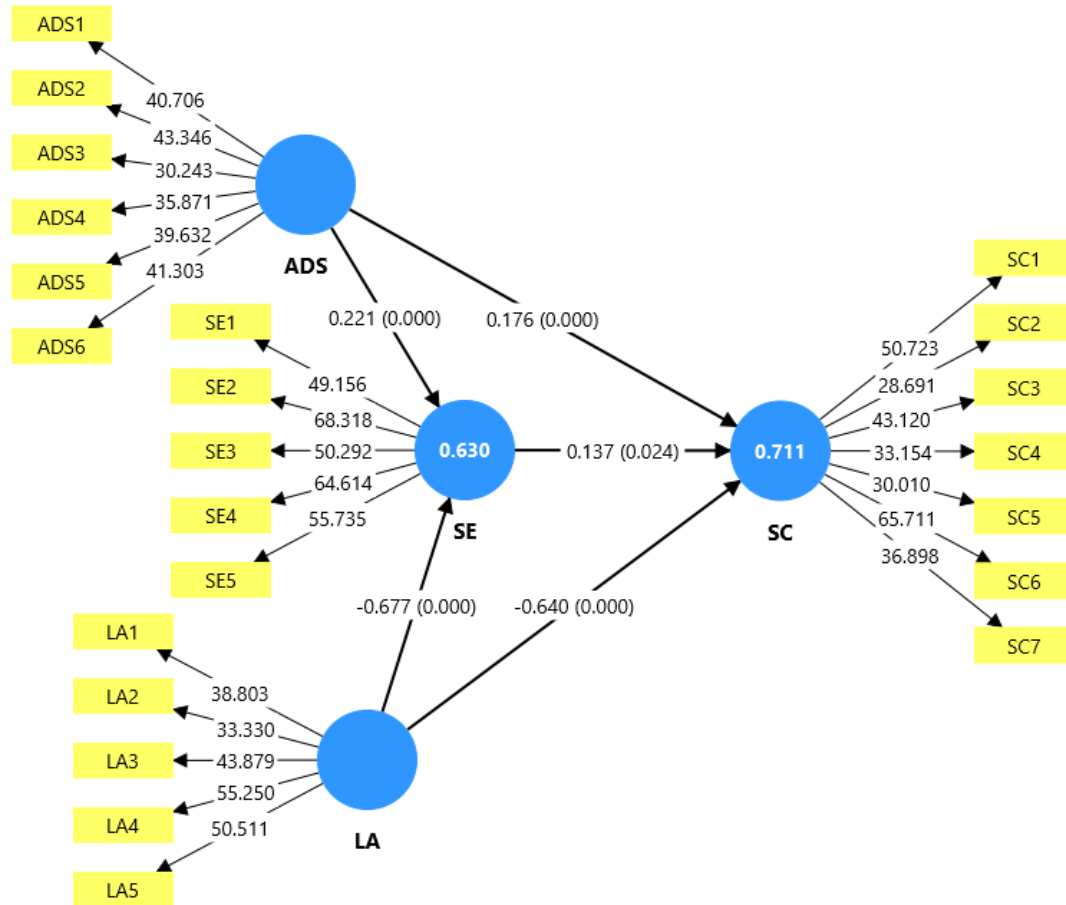


Figure 2: Graphical Results of Path Coefficient

The indirect effect results are shown in Table 7, which tests H6 for the mediating effect of SE. The mediating effect path between ADS and SC via SE was also significant ($\beta = 0.030$, $*p^* = 0.049$), confirming the mediation hypothesis. This result aligns with the social cognitive theory, which holds that environmental assistance, such as AI scaffolding, increases efficacy beliefs, which subsequently breeds confidence in performance (Bandura, 2018). This mediation aligns with Pan's (2025) observation that positive AI experiences increase self-belief. Equally, the indirect route between LA and SC via SE is also significant ($\beta = -0.093$, $*p^* = 0.027$), validating mediation. This is evident in social cognitive theory, which states that anxiety undermines the self-efficacy needed to engage in conversations with confidence, a claim supported by Wang and Wang, (2025) review. Thus, SE serves as a complementary mediator with statistical significance, albeit with a minor impact. The indirect influence of ADS through SE ($\beta = 0.030$) constitutes only 14.6% of ADS's overall effect on SC (total effect = 0.206; direct effect = 0.176). This VAF percentage (14.6%) was derived by dividing the indirect effect by the total effect ($0.030 \div 0.206 \times 100$). For the LA $\rightarrow$ SE $\rightarrow$ SC pathway, the VAF was 12.7% ($0.093 \div 0.733 \times 100$), indicating a statistically significant complementary mediation. Consequently, the majority of AI's impact on confidence is direct, rather than mediated by self-efficacy.

Table 7: Using Mediation Results, Indirect Method

Hypotheses	Path Analysis	Beta	STDEV	T statistics	C.I. Bias			P values
					Bias	2.5%	97.5%	
H6a	ADS $\rightarrow$ SE $\rightarrow$ SC	0.030	0.015	1.967	-0.000	0.005	0.068	0.049
H6b	LA $\rightarrow$ SE $\rightarrow$ SC	-0.093	0.042	2.222	0.001	-0.187	-0.014	0.027

5. Discussion of the Study

This study aimed to test whether self-efficacy mediates the relationship between AI dialogic scaffolding, language anxiety, and speaking confidence among Saudi EFL learners. This study discusses each hypothesis. The

direct effects emphasize the strong and autonomous roles of the instructional environment and the learner's affective condition. The direct relationship between language anxiety and speaking confidence was the most significant negative relationship, which proves the argument that anxiety is a primary and debilitating filter. This is consistent with the affective perspective, which emphasizes that anxiety consumes mental resources and activates negative self-assessments, which in turn directly impairs performance, irrespective of other supports (Gregersen and MacIntyre, 2020). This conclusion supports the results of the Saudi context, where anxiety permeates even during digital transformations (Alharbi, 2025). In contrast, the medium-sized but positive direct impact of AI dialogic scaffolding on confidence indicates that the technology itself is a tangible and beneficial structural advantage. In sociocultural terms, AI is an intermediary device that controls the learning process to facilitate the regular, non-judgmental exchange of speaking skills that is essential for building a sense of agency, as proposed by (Cui, Yang and Xu, 2025). Nevertheless, the absolute novelty of the proposed study's framework lies in the previously tested medium contacts via self-efficacy, which clarifies how and why these external agents shape confidence. The mediation analysis demonstrated that self-efficacy is not an outcome but a process essential to psychology. The assurance that AI dialogic scaffolding extends to is not just about skill development; fundamentally, it drives learners to exercise the skills they develop, which is an underlying tenet of social cognitive theory (Bandura, 2018). Unlike the effective modification of self-efficacy beliefs, simple exposure to technology does not always produce the intended effect on individuals, as the power of technology is not fully manifested, as meta-analytic research on the psychological impact of AI highlights (Ren, Stephens and Lee, 2026). More importantly, the mediation analysis provides an understanding of how anxiety has a deleterious impact. It not only paralyzes performance with anxiety, but the latter actually swallows up the innate faith in abilities. This loss of self-efficacy, as per the social cognitive theory, leads to a secondary passover channel that further suppresses confidence, thereby creating a vicious cycle of avoidance and doubt. This duo-pathway, immediate, and efficacy-mediated impact of anxiety provides a more precise view of its omnipresent impacts, which reinforces the results of the reviews indicating a connection between anxiety and depleted self-belief (Wang and Wen, 2025). Similar patterns are observed in other Asian and Middle Eastern contexts where speaking anxiety is prevalent. For instance, Kantar, Kok and Bakir (2026) found that English-speaking anxiety negatively predicts flourishing among Turkish EFL learners, while Flett, Nepon and Hewitt (2023) reported associations between anxiety and perfectionistic self-presentation. These cross-cultural parallels suggest that self-efficacy mediation may be generalizable beyond Saudi Arabia. In the Saudi EFL context, speaking is often a public face-threatening act. Even with AI, the anticipation of real-world evaluation (e.g., by teachers, peers) may prime anxiety. The strong effect suggests that AI's non-judgmental nature does not automatically override ingrained affective patterns. ADS likely increases self-efficacy not by reducing anxiety directly but by providing repeated successful trials. However, the small indirect effect ($\beta = 0.030$, $p = 0.049$) implies that for AI to work, learners must first notice their own success – a metacognitive step that may be weak in anxious learners. AI systems should explicitly prompt learners to attribute success to effort/ability (e.g., "You improved because you repeated the phrase correctly"). Without this attributional support, the mastery experience may not translate into self-efficacy. This way, the findings eloquently imply that self-efficacy is the critical factor in transforming AI-assisted training into learner empowerment. One should not be satisfied with technological or anxiety-reducing capabilities. Regarding the implications of the results, one may say that to make the process of AI integration effective regularly, especially in the context of a high level of anxiety, as was the case of Saudi EFL, the instructional design should be designed carefully to produce the most advantageous experiences of mastery and persuasive feedback, which has an immediate impact on self-efficacy. This can include adaptive levels of challenge, reflective questions, and feedback that is not based on pointing out errors but on ability. This study thus shifts the paradigm in the technology aspect under discussion, providing instant certainty that it is only a scaffolding for the psychological precondition of self-efficacy in the possibility of guaranteed communication. This mix, coupled with validation of the structural paths, will provide a more comprehensive plan for how one can develop AI technology to address the overall needs of the language learner and bridge the gap between the opportunities of Facebook and the actual psychological empowerment.

6. Research Implications

The current research relies on Social Cognitive Theory because it effectively positions AI dialogic scaffolding as a strong determinant of the environment, playing a proactive role in shaping self-efficacy beliefs, thereby excluding the use of human sources of efficacy information. It also mediates social, cultural, and affective perceptions and forms a compound that explains how digital tools and moods interact in a fundamental psychological process that influences the outcomes of communication. The PLS-SEM methodological approach offers a valid framework for studying technological improvements in learning by accounting for intricate

mediating mechanisms. This may provide a prototype for investigating oblique psychological pathways that cannot be examined using other methods. The central implication for practitioners and instructional designers is that AI tools must be designed purposefully to train learners' skills, foster mastery, and provide efficacy-building feedback. This recommendation aligns with Bandura's (2018) sources of self-efficacy, which entail a shift from generic interactions to scaffolded sequences that gradually build learners' confidence. Empirically, the need to include AI in dealing with the affective department is clear, and teachers and schools should not overlook this option in situations with comparatively high rates of anxiety in Saudi Arabia. One of the strategies aimed at boosting traditional teaching is presented here. Another societal contribution of this study to the field's goals of digital-era education equity and student empowerment is made. It provides them with a way out of language distortions that can decrease their self-confidence and participation in the world's academic and professional communities, where English is the medium of entry.

6.1 Limitations and Future Studies

Although this study is highly insightful, some of its limitations suggest future research directions. Because the design is cross, it does not establish temporal precedence or rule out reverse causality (e.g., confident learners may seek out more AI practice). Future experimental or longitudinal designs are required to test the causal direction. Although Harman's single-behavioral data (e.g., AI interaction logs, observer ratings of speaking confidence) or use a time-lagged design (e.g., measure AI use at T1, self-efficacy at T2, speaking confidence at T3) to reduce same-source bias. The cross-sectional design will be effective in capturing relationships at the current time; however, it will not allow for causal inference or consideration of changes in dynamics in longitudinal relationships between AI and the use of such tools. In addition, the data gathered through self-report, despite its normalcy and validation, can be complemented by behavioral indicators of real speaking behavior or system logs when interacting with an AI, which adds a multi-method perspective. The sample, grounded in Saudi university students, offers contextual relevance. However, its generalizability to other levels of education and cultural contexts with different patterns of technological adoption and affective profiles is also restricted. Future analyses can rely on longitudinal or experimental designs to establish causality and the development of self-efficacy and confidence over time. To validate the model, it can be tested by extending the study to younger students, vocational settings, or other cultures. It would be interesting to include multimodal data, such as the interpretations of the speech output, checking the measurements of the physiological symptoms of anxiety, or eye-tracking during the interaction with AI, which would give a clearer and more comprehensive picture of the learning process and even erase the gap between perceived and actual competence. We acknowledge that our model did not include potentially relevant control variables such as gender, English proficiency level, prior AI experience, year of study, or institution type. While we collected these demographic data (reported in Table 2), we chose not to include them as predictors to preserve parsimony and avoid overfitting, given our sample size ($N = 243$) and the complexity of the mediation model (five direct paths plus two indirect paths). However, this omission means we cannot rule out the possibility that one or more demographic factors explain part of the variance in speaking confidence. Future research with larger samples should test whether the mediation paths remain significant after controlling for these variables. We did not test alternative model specifications (e.g., reversing the mediation: confidence $\rightarrow$ self-efficacy $\rightarrow$ anxiety), multi-group analyses (e.g., by gender or proficiency), or non-linear effects (e.g., quadratic effects of anxiety). These should be addressed in future studies. The reported results assume linearity and correct model specification, which may not hold.

6.2 Conclusion of the Study

This study has covered a crucial issue of unequal performance of learners in AI-based language acquisition by examining the mediating psychological processes. This study acknowledges that self-efficacy is an established mediator between instructional support and oral performance in non-AI contexts. The modest novelty of this study lies in (a) testing the *specific* ADS operationalization defined above, (b) quantifying the *relative* indirect effect of ADS vs. anxiety through self-efficacy in the same model, and (c) providing the first PLS-SEM evidence from Saudi EFL university students. However, the general pattern (scaffolding $\rightarrow$ self-efficacy $\rightarrow$ confidence; anxiety $\rightarrow$ self-efficacy $\rightarrow$ confidence) replicates prior theoretical expectations. Therefore, the contribution is confirmatory and context-specific, not transformative. It revealed that the relationship between technological devices and learners' empowerment was not direct; rather, it ran through learners' belief systems, which were the major mediating variable. The findings confirm the claim that AI-dialogic scaffolding is a vital intervention in the environment, with its direct and indirect positive impact on speaking confidence among Saudi EFL learners, achieved by directly affecting self-efficacy. Meanwhile, language anxiety proved to be a potent adverse factor, having a massive direct impact on a sense of confidence and a considerable indirect impact, as it weakened

learners' self-efficacy beliefs. As a mediator, self-efficacy is a fundamental explanatory variable, and the emphasis shifts from exposure to technology to a psychological variable that results in factual empowerment. The research offers a rational framework that combines technological, cognitive, and affective features, thus contributing to the theoretical knowledge of both applied linguistics and educational technology. It extends beyond the examination of immediate impacts and shapes models of how digital interventions align with the wavy paths through which they exert influence. The studies suggest using an evidence-based approach in the Saudi case, and the researchers should focus on enhancing self-efficacy to break that anxiety and lack of self-confidence. Finally, the study illuminates that the true potential of AI in ed-Tech is not just its ability to copy an interaction; rather, it is its capacity to be organized in a way that instills psychological strength and confidence in the messages it delivers. Thus, the study provides a distinct path for the evolution of a better, more holistic, and learner-focused digital language learning environment.

Declaration of the use of AI: The authors used Grammarly for the proofreading, spelling and checking of the current work. The authors reviewed and edited the content as required, and they take full responsibility for the publication. Authors declared no AI- tools were used to write or generatet text.

Ethical Approval: Ethical approval for this study was obtained from the Institutional Review Board Committee at Prince Sattam bin Abdulaziz University (Approval No. SCBR-642/2026; Date of approval: 31 January, 2026). The study was conducted in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Competing interests: There is no conflict of interest.

Authors' contributions: All have authors have equivalent roles.

Funding: The authors extend their appreciation to Prince Sattam bin Abdulaziz University for Funding this research work through the project number (PSAU/2025/02/37654)

Conflicts of interests: The authors declare that they have no conflicts of interest to report regarding the present study.

Data Availability: Data will be available upon reasonable request from thecorrespondent author.

Consent to Participate: Before collecting data, all participants were informed and provided consent confirmation for participation. They were informed that they could withdraw from the study anytime without any consequences.

References

- Alanazi, A. H. 2024. Achieving global recognition: higher education rankings and the commitment to quality in Saudi Arabia's 2030 Strategic Vision (Doctoral dissertation, University of Glasgow).
- Alharbi, M. S. 2025. Digital transformation and EFL anxiety: A mixed-methods study on the impact of online learning on the language acquisition psychology of Saudi Arabian students. *Social Sciences & Humanities Open*, 12, 101721.
- Almusharraf, A., Bailey, D., Almusharraf, N., & Alotaibi, T. 2025. Students' perceptions of generative AI in EFL writing: Strategies, self-efficacy, satisfaction, and behavioural intention. *Australasian Journal of Educational Technology*, 41(5), 18-36.
- Alshakhi, A. 2025. Speaking test anxiety among adult Saudi EFL learners: Causes, factors, and suggested solutions. *International Journal of Language Testing*, 15(2), 166-180.
- Apriani, E., Cardoso, L., Obaid, A. J., Muthmainnah, M., Wijayanti, E., Esmianti, F., & Supardan, D. (2024). Impact of AI-powered chatbots on EFL students' writing skills, self-efficacy, and self-regulation: a mixed-methods study. *Global Educational Research Review*, 1(2), 57-72.
- Bandura, A. 2018. Toward a psychology of human agency. *Perspectives on Psychological Science*, 13(2), 130-136. <https://doi.org/10.1177/1745691617699284>
- Cancino, M., and Cabello, G. 2024. Foreign Language Speaking Anxiety and Computer Self-Efficacy in an Online Emergency Remote Teaching Environment: Perceptions from EFL Learners. *Mextesol Journal*, 48(4), n4.
- Chang, P., Zhou, L., and Zhang, L. J. 2024. The roles of task-specific self-efficacy and enjoyment in EFL learners' development of speaking CALF: A longitudinal study. *System*, 127, 103539.
- Cui, Z., Yang, H., and Xu, H. 2025. The Effects of Interaction Scenarios on EFL Learners' Technology Acceptance and Willingness to Communicate with AI. *Behavioral Sciences*, 15(10), 1391.
- Dai, X., Wang, Y., Yu, J., Qiu, B., Wang, R., and Na, M. 2025. Motivation and engagement as pathways: how AI-augmented online assessment shapes English-speaking competency in vocational EFL classrooms. *Frontiers in Psychology*, 16, 1730953.
- Flett, G. L., Nepon, T., & Hewitt, P. L. (2023). The Perfectionistic Self-Presentation Failure Scale: Psychometric Properties and Associations with Depression and Social Anxiety. *Journal of Concurrent Disorders*, 12-32.

- Ebadi, S., Velayati, S., Ramezanzadeh, A., and Salman, A. R. 2025. Exploring the impact of AI-powered speaking tasks on EFL learners' speaking performance and anxiety: An activity theory study. *Acta Psychologica*, 259, 105391.
- Ekizer, F. N. 2025. Exploring the impact of artificial intelligence on English language teaching: A meta-analysis. *Acta Psychologica*, 260, 105649.
- Grab, M. Ö. 2025. Integrated AI chatbot practice: A pathway to improved ESL speaking skills. *Social Sciences & Humanities Open*, 12, 101933.
- Gregersen, T., and MacIntyre, P. D. 2020. Acting locally to integrate positive psychology and peace: practical applications for language teaching and learning.
- Griffiee, D. T. 1997. Validating a questionnaire on confidence in speaking English as a foreign language. *JALT Journal*, 19(2), 177-197.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Lux, G., and Troiville, J. 2022. *Modèles d'équations structurelles Partial Least Squares (PLS-SEM)* (No. hal-03782748).
- Hair, J. F., Risher, J. J., Sarstedt, M., and Ringle, C. M. 2019. When to use and how to report the results of PLS-SEM. *European business review*, 31(1), 2-24.
- Heydarnejad, T. 2025. Disclosing motivation and educational self-efficacy in AI-enhanced EFL learning via a lens into reflective thinking, academic mindfulness, and teacher support. *Discover Psychology*, 5(1), 145.
- Horwitz, E. K., Horwitz, M. B., and Cope, J. 1986. Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125-132.
- Kantar, A., Kök, M., and Bakir, A. (2026). Conformity to Masculine Norms and Academic Entitlement: Alexithymia and Empathy as Serial Mediators. *Gender Issues*, 43(1), 21.
- Krejcie, R. V., and Morgan, D. W. 1970. Table for determining sample size from a given population. *Educational and Psychological Measurement*, 30(3), 607-610.
- Lantolf, J. P., Thorne, S. L., and Poehner, M. E. 2014. Sociocultural theory and second language development. In *Theories in second language acquisition* (pp. 221-240). Routledge.
- Li, J., Wang, C., Zhao, Y., and Li, Y. 2024. Boosting Learners' Confidence in Learning English: Can Self-Efficacy-Based Intervention Make a Difference?. *TESOL Quarterly*, 58(4), 1518-1547.
- Lin, Y., Wang, B., Zhang, G., and Su, Y. 2025. The relations among foreign language anxiety, academic buoyancy, and willingness to communicate in EFL classroom. *Frontiers in Psychology*, 16, 1634054.
- Ma, Y. (2022). The triarchy of L2 learners' emotion, cognition, and language performance: Anxiety, self-efficacy, and speaking skill in lights of the emerging theories in SLA—*Frontiers in psychology*, 13, 1002492.
- Ma, Y., and Chen, M. 2025. The Human Touch in AI: Optimizing Language Learning through Self-Determination. *Frontiers in Psychology*, 16, 1568239.
- Pan, Y., and Li, G. 2025. The effects of perceived teacher support and growth language mindset on learner well-being in AI-integrated environment: the mediating role of generative AI attitude. *Frontiers in Psychology*, 16, 1660462.
- Radjuni, M., Sahraeny, S., and Latief, M. R. A. 2025. The Relationship between Self-Efficacy and EFL Students' Speaking Performance: A Case Study of English Department Students. *SHIELD: Journal of Studies on Human Interaction, Education, and Language Development*, 1(1).
- Ren, L., Stephens, J. M., and Lee, K. 2026. The Impact of AI on Learners' Self-Efficacy: A Meta-Analysis. *Behavioral Sciences*, 16(1), 158.
- Septiawan, H., Setiyadi, B., Mahpul, M., Sukirlan, M., and Nisa, K. 2025. The anxiety in English for foreign language speaking class: the case of university students in Lampung context. *Journal of Education and Learning (EduLearn)*, 19(1), 362-370.
- Tian, F., Li, H., Gao, Y., and Li, J. 2026. A GenAI-supported dialogic model to enhance the functions of peer feedback in EFL writing: a mixed-method interventional study: *system*, 136, 103895.
- Tuan, N. D. 2025. Boosting Fluency, Reducing Fear: Utilizing Generative AI as a Scaffolding Tool in EFL Speaking Classes.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wang, L., and Wang, L. 2025. Exploring enjoyment, motivation, self-efficacy, and engagement in AI-assisted English learning: A self-determination theory approach. *Learning and Motivation*, 92, 102197.-
- Wang, Y., and Sun, P. P. 2024. Development and validation of scales for speaking self-efficacy: Constructs, sources, and relations. *Plos one*, 19(1), e0297517.
- Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System*, 39(4), 510-522.
- Xiao, F., Zou, E. W., Lin, J., Li, Z., and Yang, D. 2025. Parent-led vs. AI-guided dialogic reading: Evidence from a randomized controlled trial in children's e-book context. *British Journal of Educational Technology*, 56(5), 1784-1813.
- Yan, J., Wu, C., Tan, X., and Dai, M. 2025. The influence of AI-driven personalized foreign language learning on college students' mental health: a dynamic interaction among pleasure, anxiety, and self-efficacy. *Frontiers in Public Health*, 13, 1642608.
- Zhang, W. 2026. The impact of AI chatbots on EFL learners' oral proficiency and willingness to communicate. *System*, 136, 103919.
- Zhang, X. 2025. Enhancing oral English learning through AI: a case study on the impact of AI-driven speaking applications among Chinese university students. *Frontiers in Psychology*, 16, 1595818.

Generative AI in University Mathematics: Attitudes and Academic-Leisure Use Patterns

Hassan Hossein-Mohand¹, Hossein Hossein-Mohand¹, Manuel García-Alonso² and María del Carmen Olmos-Gómez¹

¹University of Granada, Spain

²UNED (National University of Distance Education), Spain

hassan.h.m@ugr.es

hossein.h.m@ugr.es (corresponding author)

mangarcia@melilla.uned.es

mcolmos@ugr.es

<https://doi.org/10.34190/ejel.24.4.4850>

An open access article under [CC Attribution 4.0](#)

Abstract: The rapid spread of generative artificial intelligence has reshaped university e-learning, yet its incorporation into mathematics learning remains constrained by disciplinary demands for precision, justification, and epistemic scrutiny. This study examined the relationship between socio-demographic variables, access conditions, and attitudes towards AI in mathematics, and analyzed how these factors were associated with reported use of AI tools in two differentiated contexts: academic learning and leisure. A quantitative, observational, non-experimental, cross-sectional study was conducted with 869 students from the University of Granada across the Melilla, Ceuta, and Granada campuses. Data were collected through an online questionnaire that included the IAMAT scale and two dichotomous indicators of AI use. Inferential analyses were estimated with between 834 and 846 complete cases, depending on the procedure. The IAMAT scale showed adequate internal consistency ($\alpha = .796$), high sampling adequacy ($KMO = .888$), and an interpretable two-factor structure that distinguished between usefulness/confidence and uncertainty/errors. AI use was more frequent in mathematics learning (62.6%) than in leisure (28.3%). Logistic regression models indicated that positive attitudes towards AI in mathematics increased the likelihood of use in both contexts. In academic use, older age, lack of Wi-Fi access, and membership of the Melilla campus were associated with a lower probability of use, whereas participation in voluntary work was associated with a higher probability. In leisure use, women showed a lower probability of use than men. In addition, K-means clustering identified six differentiated profiles defined by age, perceived usefulness, and distrust. These profiles discriminated academic use significantly, but not leisure use. The findings suggest that the adoption of AI in mathematics cannot be reduced to technological access alone, because it is also shaped by domain-specific cognitive and affective dispositions. From an e-learning perspective, these findings contribute to AI-supported mathematics education by showing that digital learning environments should move beyond mere access to GenAI and embed critical AI literacy, mathematical verification criteria, and reasoning-oriented tasks in which students explain, check, and revise AI-generated solutions rather than delegate the full intellectual workload to the tool.

Keywords: Generative AI, Higher education, Mathematics learning, AI literacy, Attitudes towards AI, e-Learning

1. Introduction

The rapid diffusion of generative artificial intelligence across higher education has shifted the debate from technical availability to questions of epistemic responsibility, assessment design, and institutional governance (Bond et al., 2024; Jin et al., 2025). In e-learning environments, this shift is visible because large language models promise immediacy, personalisation, and conversational support, whilst also altering how students search for information, construct explanations, and delegate intellectual work to automated systems (Chan and Hu, 2023; Park, 2025; Rapanta et al., 2025). Mathematics learning sharpens these tensions because the value of an answer depends on conceptual precision, procedural validity, and the quality of the justification that supports it (Pepin, Buchholtz and Salinas-Hernández, 2025).

Within this domain, a plausible answer is not necessarily a mathematically valid one. Small symbolic deviations, omitted steps, or formally appealing but invalid inferences can change the status of a solution altogether (Yoon et al., 2024). For that reason, the educational significance of generative AI in mathematics cannot be inferred from generic technology-adoption models alone. What matters is how students judge the tool's usefulness, how they calibrate trust in its outputs, and under which conditions AI is incorporated into tasks such as searching for theory, translating verbal problems into mathematical language, checking solutions, or generating practice (Bewersdorff et al., 2025; Pham, Klaus and Bateh, 2025).

Against this background, the present study addresses two linked empirical gaps. First, large-scale quantitative evidence remains limited on how multidimensional attitudes towards AI in mathematics are associated with reported use among university students, particularly in contexts connected with teacher education. Second, it remains unclear whether AI adoption in mathematics reflects a general orientation towards GenAI across everyday digital life or a more situated academic appropriation. For this reason, the study examines AI use in two contexts, academic learning and leisure. Leisure use is not treated as a direct mathematics-learning outcome, but as a theoretically useful contrast that helps determine whether AI use in mathematics is part of broader recreational engagement with GenAI or is specifically attached to study-related tasks and e-learning conditions. The study also identifies attitudinal profiles in order to determine whether different constellations of usefulness, distrust, and age are linked to differentiated patterns of AI use among university students. The theoretical framework is organised as a progression from the institutional level to the cognitive and pedagogical implications of generative AI in mathematics. It first addresses the move from institutional adoption to critical GenAI literacy, then examines the specific demands of mathematics education, before turning to attitudes, calibrated trust, and the relationship between pedagogical usefulness and cognitive risk. This progression provides the rationale for the study and clarifies why a domain-specific instrument such as IAMAT is analytically preferable to a generic measure of technology acceptance.

1.1 From Institutional Adoption to Critical GenAI Literacy

The arrival of generative AI in higher education can no longer be interpreted as a merely instrumental innovation (Bond et al., 2024; Jin et al., 2025). Comparative analyses of university policies suggest that institutional responses are increasingly structured around three interdependent concerns: academic integrity, redesign of teaching and assessment, and fairness of access. The central question is therefore not simply whether students use GenAI, but under which normative conditions, with what degree of human supervision, and according to which epistemic expectations. Jin et al. (2025), in their analysis of adoption strategies across universities from six world regions, show that institutions are moving from defensive responses to more complex frameworks based on ethical-use guidance, authentic assessment, and targeted training for staff and students. At the same time, they note that substantial gaps remain in privacy protection, transparency, and equitable access.

Within this context, generative AI literacy has become a key concept. Park (2025) frames it as a cluster of interrelated competences that includes understanding how the tool works, using and applying it, evaluating and incorporating its outputs, addressing ethical implications, and adopting an informed attitudinal stance towards it. Rapanta et al. (2025) extend this argument by insisting that critical GenAI literacy must include epistemic, ethical, and relational dimensions. Students therefore need more than prompt-writing skills. They need criteria for judging the credibility of an answer, recognising bias and omission, and deciding when the tool supports learning and when it encourages cognitive delegation. This shift is consequential in teacher education, where the use of AI shapes not only immediate performance but also future professional conceptions of knowledge, validation, and teaching (Bernardi et al., 2025; Busuttil and Calleja, 2025).

1.2 The Specificity of Mathematics Education in Relation to Generative AI

Mathematics education introduces an additional layer of difficulty because the educational value of AI depends on mathematical fidelity rather than on linguistic fluency alone (Yoon et al., 2024). In other domains, a partially approximate answer may still be serviceable; in mathematics, incomplete justification, symbolic imprecision, or plausible but invalid reasoning may undermine the solution itself (Noster, Gerber and Siller, 2024). Pepin, Buchholtz and Salinas-Hernández (2025), in their scoping review of ChatGPT in mathematics education, show that these tools are already being used to support conceptual understanding, lesson design, assessment, personalised assistance, and collaboration. Yet the same review makes clear that such affordances coexist with errors, hallucinations, dependency, and ethical concerns, which means that human oversight is not optional but structural.

From a mathematics education perspective, generative AI should therefore be understood as a resource that reconfigures students' mathematical activity rather than as a neutral shortcut to information. Kock, Salinas-Hernández and Pepin (2025) show that engineering students develop initial schemes of use for ChatGPT when they try to understand concepts, explore alternative modelling strategies, translate mathematical models into code, and optimise it. Contel and Cusi (2025) likewise report that GPT-4 can scaffold metacognitive processes during problem solving, especially planning, monitoring, and revision, provided that interaction with the system is organised in reflective ways. These contributions converge on an important point for the present study: the educational potential of GenAI is activated not by exposure to the tool itself, but by patterns of use that preserve student agency, task regulation, and evaluation of the produced response.

Recent work in mathematics education also suggests that the crucial issue is not only what generative AI can do, but what forms of error it encourages students to normalise. Noster, Gerber and Siller (2024) found that pre-service teachers were exposed to both correct and incorrect ChatGPT outputs across all analysed tasks, and that the likelihood of submitting an incorrect answer increased when the model's initial answer was also incorrect. In the domain of mathematical proof, Yoon et al. (2024) showed that students' uses of GenAI are mediated by their conceptions of proof, their conceptions of AI, and their ethical considerations, leading the authors to argue that university work with proof must explicitly teach critical engagement with AI outputs rather than passive acceptance. In teacher education, Bernardi et al. (2025) and Busuttil and Calleja (2025) further suggest that future teachers may judge the pedagogical appeal of AI-generated proposals more easily than their mathematical correctness, which makes the distinction between usefulness, credibility, and fidelity especially important. Overall, this body of work identifies a common tension that frames the present study. GenAI can support mathematical activity when it scaffolds explanation, monitoring, practice, and revision, but it can also weaken learning when students treat plausible output as already verified knowledge. This tension justifies examining how students' mathematics-specific attitudes towards usefulness, confidence, and possible error are associated differently with academic and non-academic use.

1.3 Attitudes, Calibrated Trust, and Patterns of Use

Research on ChatGPT adoption in higher education shows that attitudes towards AI do not form a simple continuum running from enthusiasm to rejection. Pham, Klaus and Bateh (2025) report that students' intention to use ChatGPT depends chiefly on perceived usefulness and subjective norm, whilst ease of use and trust act more indirectly by reinforcing usefulness. Bewersdorff et al. (2025) complicate the picture by identifying differentiated student profiles, including AI advocates, cautious critics, and pragmatic observers, and by showing that use, interest, attitudes, and AI literacy jointly shape AI self-efficacy. These findings justify treating attitudes not as a decorative background variable, but as a potential organising principle of students' relations with AI.

In mathematics, however, attitudes need to be conceptualised more narrowly than generic acceptance, because AI-supported mathematical activity requires students to evaluate usefulness together with correctness, justification, and possible error (Pepin, Buchholtz and Salinas-Hernández, 2025). The relevant construct includes perceived usefulness, confidence, awareness of error, willingness to verify, and ethical judgement. This is why the notion of calibrated trust is especially productive. A student with stronger AI literacy is not one who always trusts the system, but one who can distinguish when an output may serve as a hint, when it requires external checking, and when it should be rejected because of inconsistency or insufficient mathematical justification. This distinction helps to interpret the dimensions captured by the IAMAT scale, whose items combine perceived benefits with reservations about reliability in mathematically specific tasks. Accordingly, comparing academic and leisure use makes it possible to examine whether these attitudes merely reflect a general orientation towards AI or whether they are more strongly tied to mathematically framed learning tasks.

1.4 Pedagogical Usefulness, Cognitive Risk, and the Rationale for the Study

A theoretically adequate account of generative AI in education must also consider the relation between use and actual learning. Deng et al. (2025), in their meta-analysis of experimental studies, report overall positive effects on academic achievement, affective-motivational outcomes, and the propensity for higher-order thinking, together with a reduction in perceived mental effort. Yet Bauer et al. (2025) warn that better task performance under AI assistance does not necessarily translate into durable learning, and they argue that AI can operate through substitution, augmentation, redefinition, or even inversion of the intended learning process. In mathematics, this warning is particularly salient because an apparently helpful support may displace processes that are constitutive of learning itself, such as justifying a procedure, validating a result, checking symbolic consistency, or sustaining a proof.

Recent empirical evidence supports this cautious stance. Stadler, Bannert and Sailer (2024) found that large language models can reduce perceived cognitive load whilst at the same time producing shallower student reasoning in scientific inquiry tasks. Bastani et al. (2025) show that access to generative AI may improve immediate performance in mathematics and nonetheless harm subsequent learning when the tool is used as an answer machine rather than as support for reasoning. Among pre-service mathematics teachers, Zhang et al. (2025) further report that AI dependency is negatively associated with higher-order competences, whilst AI literacy partly mitigates that risk. Taken together, these studies suggest that the educational issue is not whether AI helps, but under which instructional conditions it helps without eroding the intellectual work it is meant to support.

These considerations return to the empirical gap introduced at the beginning of the article. In mathematics education, a substantial share of the recent evidence remains small-scale, exploratory, qualitative, or task-specific studies, often focused on proof, problem solving, or lesson planning (Contel and Cusi, 2025). Less is known, in large university samples and among future teachers, about how multidimensional attitudes towards AI in mathematics relate to reported use across academic and non-academic contexts. The comparison between academic learning and leisure is therefore not intended to make leisure use a direct mathematics-learning outcome. Rather, it provides a contrast that helps determine whether AI adoption is primarily a general digital habit or a situated e-learning practice linked to study tasks, mathematical verification, and access conditions. The present study addresses that gap by treating generative AI not as a homogeneous block, but as a situated practice tied to specific mathematical tasks and shaped by perceived usefulness, confidence, awareness of error, and access conditions. This perspective makes it possible to move beyond generic questions of adoption towards a more precise understanding of how students relate to AI when learning mathematics.

1.5 Study Objectives

This conceptual framework leads to one general objective and three specific objectives. Because leisure use is included as a comparison context, the objectives distinguish general engagement with GenAI from its academic appropriation in mathematics e-learning. G.O: To analyse how socio-demographic variables, access conditions, and attitudes towards AI in mathematics are related to reported use of AI tools in two contexts, academic learning and leisure.

- S.O.1.- To describe the prevalence of AI-tool use in learning and leisure and to examine the basic psychometric evidence for the IAMAT scale in the analysed sample.
- S.O.2.- To estimate the relative weight of attitudes towards AI in mathematics, together with socio-demographic and access variables, on the probability of AI use in both contexts.
- S.O.3.- To identify differentiated attitudinal profiles and to analyse their association with academic and leisure use of AI.

2. Method

2.1 Study Design

A quantitative, observational, non-experimental, cross-sectional study was conducted using a self-administered questionnaire. The design was intended to examine associations between socio-demographic variables, access conditions, and attitudes towards generative AI in mathematics, and to estimate their relationship with reported AI use in two distinct contexts, academic learning and leisure. Because all variables were observed at a single point in time and no experimental manipulation or random assignment was involved, the study had a descriptive and explanatory-predictive scope. The organisation of the method section followed the reporting logic recommended by STROBE for cross-sectional observational studies (von Elm et al., 2007).

2.2 Participants and Context

The target population comprised students from the University of Granada across the Melilla, Ceuta, and Granada campuses. Participants were recruited through non-probability convenience sampling based on questionnaire dissemination in the ordinary university context and on students' voluntary participation. A total of 869 questionnaires were validly recorded. For inferential analyses, complete-case criteria were applied according to the variables included in each procedure, which resulted in effective sample sizes ranging from 834 to 846 participants. In the sample used for the main analyses ($n = 846$), 67.6% identified as women ($n = 572$) and 32.4% as men ($n = 274$). Mean age was 21.76 years ($SD = 4.56$). Campus distribution was as follows: Melilla ($n = 592$, 70.0%), Ceuta ($n = 159$, 18.8%), and Granada ($n = 95$, 11.2%). With regard to access conditions, 95.5% reported having Wi-Fi connection ($n = 808$). Perceived socio-economic level was predominantly middle ($n = 720$, 85.1%), followed by low ($n = 104$, 12.3%) and high ($n = 22$, 2.6%). In addition, 22.2% reported being in work ($n = 188$) and 12.4% reported involvement in voluntary work ($n = 105$). This composition provided a heterogeneous basis for analysing AI use by age, campus, and contextual conditions, although no claim of probabilistic representativeness is made.

2.3 Instruments and Variables

Data were collected through an online questionnaire administered via Google Forms. The instrument contained 26 items distributed across three blocks. The first gathered socio-demographic and contextual variables, namely gender, age, campus membership, availability of Wi-Fi, perceived socio-economic level, employment status, and

participation in voluntary work. The second included two dichotomous criterion variables, use of AI in learning activities (IA_APR: 0 = no, 1 = yes) and use of AI in leisure or free-time activities (IA_OCIO: 0 = no, 1 = yes). These variables were conceptualised as indicators of reported prevalence of use. Their function was to identify whether AI tools had entered each context of activity, academic learning and leisure, while the IAMAT scale captured students' mathematics-specific attitudes towards possible uses of AI. The third block included the IAMAT scale, which comprised nine five-point Likert items (1 = strongly disagree; 5 = strongly agree) designed to measure attitudes towards the use of AI in mathematics.

The IAMAT scale was developed from the mathematics-related component of a broader questionnaire on university students' use and perceptions of artificial intelligence. In its construction, the items were intended to measure attitudes towards AI-mediated mathematical activity. Item content was therefore organised around mathematically situated tasks and judgements, including searching for theoretical content, solving problems, working with geometry, translating real situations into mathematical language, checking solutions, generating practice exercises, and evaluating reliability.

Before administration, the questionnaire underwent a content-validation process based on expert judgement using a three-round Delphi procedure (Hsu and Sandford, 2007; Lynn, 1986). The expert panel consisted of 15 university lecturers holding a PhD and representing complementary areas related to the construct under study: mathematics and physics (five experts), pedagogy and research methods (six experts), and social sciences and linguistics (four experts). In each round, the experts assessed the appropriateness, relevance, and coherence of the items using a three-point scale ranging from 1 (not appropriate) to 3 (fully appropriate). The items were progressively refined in response to the experts' comments. These refinements involved wording and clarity adjustments, but no item was removed, because all items reached the predefined consensus threshold by the final round. The process ended with a 98% agreement rate, above the 90% criterion established for retaining the items. No separate pilot study with an independent subsample was conducted. Accordingly, the Delphi process functioned as the pre-fieldwork refinement stage, whereas the internal consistency analysis and exploratory factor analysis reported below provide initial empirical evidence of the scale's psychometric behaviour in the final analytical sample.

IAMAT assessed three complementary conceptual cores: perceived usefulness for specific mathematical tasks, confidence in AI-generated responses, and awareness of uncertainty or possible error. Substantively, the items covered searching for theoretical content, solving problems, support in geometry, translating real situations into mathematical language, checking solutions, generating exercises for practice, and judging reliability. Items 7 and 8 were phrased in the negative direction and were treated analytically as indicators of caution about error and reliability. Accordingly, higher values on IAMAT_DESCONF indicate greater distrust or stronger awareness of possible error. It should be noted that calibrated trust was used as an interpretive concept in the discussion to make sense of the coexistence of perceived usefulness, confidence, and awareness of possible error. The full wording of the items and their analytic orientation are provided in the Supplementary Material (Table S1).

For descriptive purposes and profile characterisation, three continuous indicators were derived from the scale: IAMAT_MEAN, computed as the mean of the nine items as a broad summary index; IAMAT_UTIL, computed as the mean of Items 1 to 6 and 9 so as to capture usefulness and confidence; and IAMAT_DESCONF, computed as the mean of Items 7 and 8 so as to represent distrust and concern about error. These composite indices were useful for description and clustering. For the logistic regression models, however, the composite indices were not entered simultaneously because of their conceptual overlap. Instead, standardised factor scores derived from the exploratory factor analysis were used as continuous predictors in order to represent attitudinal dimensions more parsimoniously and reduce redundancy between predictors.

2.4 Procedure

Data collection was conducted online. The questionnaire link was disseminated to the target population within ordinary university activity. The opening screen informed participants about the aim of the study, the voluntary nature of participation, the right to discontinue at any time, and the guarantees of confidentiality and anonymity. No direct personal identifiers were requested, and access to the questionnaire was conditional upon acceptance of informed consent. Once data collection had finished, the database was cleaned prior to analysis through the review of codings, consistency checks, and reverse coding of the negatively worded IAMAT items.

2.5 Data Analysis

Analyses were carried out with IBM SPSS Statistics 25.0. In a first phase, descriptive statistics were calculated for all variables. Absolute and relative frequencies were obtained for categorical variables, and means and standard

deviations for continuous variables. Multivariate analyses were estimated using complete cases for the variables involved in each model.

In a second phase, the psychometric behaviour of the IAMAT scale was examined. Internal consistency was estimated by means of Cronbach's alpha and inspection of corrected item-total correlations. This procedure was chosen because it provides a useful first indicator of internal homogeneity in short survey scales, although its interpretation was made cautiously and always in connection with the dimensional structure of the instrument (Tavakol and Dennick, 2011). An exploratory factor analysis was then conducted using principal axis factoring with Oblimin oblique rotation. Although IAMAT was constructed around theoretically defined conceptual cores, a confirmatory approach was not prioritized at this stage because the scale had not yet been independently validated in this population and no separate validation sample was available. The analysis was therefore conceived as an initial examination of internal structure rather than as a strict test of a previously confirmed measurement model. Principal axis factoring was selected because the focus was on common variance, and Oblimin rotation was used because the underlying attitudinal dimensions were theoretically expected to correlate. Matrix factorability was checked through the Kaiser-Meyer-Olkin index and Bartlett's test of sphericity. Factor retention was guided by eigenvalues above 1, the loading pattern, and the conceptual coherence of the solution, following good-practice recommendations for exploratory validation studies (Watkins, 2018). Regression-based factor scores were saved for subsequent use as continuous predictors. Item-level psychometric diagnostics are reported in the Supplementary Material (Table S2).

In a third phase, two independent binary logistic regression models were estimated, one for IA_APR and one for IA_OCIO. In both models, the dependent variable was coded as 0 = no use and 1 = use. Age and the attitudinal factor scores were entered in a first block, and socio-demographic and contextual variables were entered in a second block, namely gender, availability of Wi-Fi, level of studies, campus, perceived socio-economic level, employment status, and voluntary work. Interpretation was based on B coefficients, Wald statistics, odds ratios with 95% confidence intervals, and overall model fit indicators, including the likelihood-ratio chi-square, Nagelkerke's R^2 , and the Hosmer-Lemeshow test. The number of observed events in both criterion variables was also sufficient relative to the number of predictors considered, which favours estimation stability in logistic models of this complexity (Peduzzi et al., 1996). As an additional sensitivity analysis, and in response to the possibility that usefulness/confidence and uncertainty/errors might operate jointly, the FAC1_1 $\times$ FAC2_1 interaction term was added to both logistic models. This analysis assessed whether the association between usefulness/confidence and AI use varied according to the level of uncertainty/errors. The main-effects models were retained as the primary specification unless the interaction term produced a statistically and substantively meaningful improvement in model fit or interpretation.

In a fourth phase, segmentation techniques were applied in order to identify attitudinal profiles. The clustering variables were selected on both empirical and conceptual grounds. FAC1_1 and FAC2_1 represented the two empirically derived dimensions of IAMAT, namely usefulness/confidence and uncertainty/errors. IAMAT_MEAN was retained as a broad indicator of overall attitudinal intensity, and age was included as a continuous biographical variable because the adoption of AI in academic learning was expected to vary across student trajectories and age groups. The outcome variables, IA_APR and IA_OCIO, were deliberately excluded from the clustering procedure in order to avoid circularity, because the purpose was to examine afterwards whether independently derived profiles were associated with AI use. Other socio-demographic and contextual variables were also not used to define the clusters, since the aim was to obtain attitudinal profiles rather than groups driven primarily by gender, campus, or access conditions.

Preliminary solutions were explored through TwoStep clustering and K-means solutions were then estimated. A six-cluster solution was finally retained because it offered the best balance between empirical separation, stability across exploratory procedures, and substantive interpretability. Solutions with fewer clusters merged theoretically distinct configurations, such as students with high perceived usefulness but different levels of distrust, whereas more fragmented solutions produced very small or substantively redundant groups. The retained solution was interpreted from an explicitly analytical and descriptive logic, with the aim of capturing internal heterogeneity in the sample rather than postulating closed or exhaustive typologies. As an additional quality criterion, the classical literature on the interpretation and validation of non-hierarchical partitions was considered, particularly in relation to compactness and separation between groups (Rousseeuw, 1987).

Finally, the associations between the resulting profiles and AI use in learning and leisure were examined through contingency tables and Pearson's chi-square tests, with Cramér's V reported as an effect-size index. Statistical significance was set at $p < .05$ throughout.

2.6 Ethical Considerations

Participation was voluntary and based on informed consent. Anonymity of responses and confidentiality of information were guaranteed throughout the study. The research received a favourable opinion from the University of Granada Research Ethics Committee (ref. 4518/CEIH/2024, 16 September 2024). The approval confirmed compliance with the applicable ethical principles and with current personal-data protection regulations. Data were used exclusively for scientific purposes and were analysed and reported in aggregated form.

3. Results

3.1 Prevalence of Use and Psychometric Evidence for the IAMAT Scale

In relation to S.O.1, inferential analyses were estimated using complete cases, with $n = 846$ for the psychometric assessment and the model of AI use in learning, and $n = 834$ for the model of AI use in leisure. Overall, 530 students (62.6%) reported using AI tools for learning purposes, whereas 236 (28.3%) reported AI use in leisure or free time. Academic use was therefore more than twice as frequent as recreational use, which locates the phenomenon primarily in practices oriented towards task completion and study support rather than in undifferentiated general consumption.

The evidence gathered for this first objective also showed that the IAMAT scale has a sufficient psychometric basis for the subsequent analyses. Sampling adequacy was high ($KMO = .888$) and Bartlett's test of sphericity was significant, $\chi^2(36) = 4528.13$, $p < .001$. Principal axis factoring with Oblimin rotation identified two correlated dimensions that explained 62.90% of the common variance. As shown in Table 1, the first factor concentrated the applications with the clearest instrumental value for mathematical work, with particularly high loadings for problem solving (.897), geometry (.858), translation of verbal situations into mathematical language (.803), and checking solutions (.798). Generation of practice problems (.661) and general confidence in AI responses (.679) also loaded on this factor, albeit with slightly less weight. The second factor brought together the two indicators of epistemic caution associated with error and uncertainty, with loadings of .818 and .746. The correlation between factors was negative and moderate ($r = -.285$), which indicates that usefulness and caution do not operate as strictly opposite poles. At the descriptive level, item means fell within a moderate range (2.43-2.94). The highest mean was observed for checking solutions, whereas the lowest corresponded to the use of AI for geometry tasks. Internal consistency of the full scale was adequate ($\alpha = .796$). Full item-level diagnostics are provided in the Supplementary Material (Table S2).

Table 1: Rotated factor matrix, communalities, and descriptive statistics for IAMAT items

Abbreviated item	M	SD	h^2	F1	F2
Searching for theory	2.86	1.38	.664	.801	
Problem solving	2.70	1.37	.779	.897	
Geometry	2.43	1.33	.722	.858	
Verbal problems to mathematical language	2.67	1.31	.675	.803	
Checking solutions	2.94	1.36	.663	.798	
Generating practice problems	2.65	1.47	.473	.661	
Potential errors in mathematics	2.74	1.37	.711		.818
Reservations about reliability	2.80	1.31	.539		.746
Confidence in responses	2.67	1.19	.434	.679	

Note. Extraction by principal axis factoring and Oblimin rotation. Blank cells indicate the absence of a salient loading in the reported factor. $KMO = .888$; Bartlett's $\chi^2(36) = 4528.13$, $p < .001$; explained variance = 62.90%; $\alpha = .796$; correlation between factors = $-.285$. F1 = usefulness/confidence; F2 = uncertainty/errors.

3.2 Predictors of AI Use in Learning and Leisure

To address S.O.2, Table 2 reports the final logistic models for both contexts of use. The model for IA_APR was significant, $\chi^2(13) = 82.61$, $p < .001$, showed moderate explanatory power (Nagelkerke $R^2 = .127$), and fitted the data adequately (Hosmer-Lemeshow, $p = .529$). The IA_OCIO model was also significant, $\chi^2(13) = 32.65$, $p = .002$, although with substantially lower explained variance (Nagelkerke $R^2 = .055$; Hosmer-Lemeshow, $p = .507$). Taken together, these values suggest that the considered predictors captured academic use more precisely than leisure use.

The most consistent pattern was the effect of positive attitudes towards AI in mathematics, which increased the probability of use in both domains: OR = 1.37 for learning and OR = 1.38 for leisure. By contrast, the uncertainty/errors factor did not reach statistical significance in either model. This result is theoretically important because it suggests that adoption depends less on eliminating reservations about reliability than on perceiving the tool as functionally valuable for mathematically relevant purposes.

Additional sensitivity models including the FAC1_1 $\times$ FAC2_1 interaction did not provide evidence that the effect of usefulness/confidence on AI use varied significantly according to uncertainty/errors. For academic AI use, the interaction term was not significant, $B = .065$, OR = 1.067, 95% CI [.904, 1.260], $p = .444$, and the increase in model fit was not significant, $\chi^2(1) = .586$, $p = .444$. For leisure AI use, the interaction term was also not significant, $B = .112$, OR = 1.119, 95% CI [.943, 1.327], $p = .198$, with no significant improvement in model fit, $\chi^2(1) = 1.676$, $p = .195$. The main-effects models were therefore retained as the primary specification. This result suggests that the relatively high academic use observed among distrustful moderates should be interpreted as a descriptive profile pattern rather than as evidence of a statistically confirmed interaction between the two attitudinal dimensions.

From that shared pattern onwards, the models diverged. In learning, age showed an inverse effect (OR = 0.93), so that each additional year of age was associated with a lower probability of use. Lack of Wi-Fi reduced the probability of academic use by roughly one half (OR = 0.47), and the contrast Melilla versus Ceuta was also significant (OR = 0.44). Given the uneven distribution of the sample across campuses, this contrast should be read as a sample-specific association rather than as evidence of a general campus-level effect. In addition, participation in voluntary work was associated with a higher probability of academic use (OR = 1.62). In leisure, by contrast, the only socio-demographic variable with a robust effect was gender: women showed a lower probability of recreational use than men (OR = 0.59). The full model coefficients are reported in the Supplementary Material (Tables S3 and S4).

Table 2: Logistic regression models for AI use in learning and leisure

Predictor	IA_APR OR [95% CI]	p	IA_OCIO OR [95% CI]	p
F1. Positive attitudes	1.37 [1.16, 1.62]	< .001	1.38 [1.16, 1.63]	< .001
F2. Uncertainty/errors	0.93 [0.78, 1.11]	.438	0.95 [0.79, 1.15]	.601
Age	0.93 [0.90, 0.96]	< .001	1.00 [0.96, 1.03]	.875
Gender (women vs men)	1.11 [0.80, 1.54]	.532	0.59 [0.42, 0.82]	.002
Wi-Fi (no vs yes)	0.47 [0.23, 0.94]	.032	0.49 [0.20, 1.22]	.125
Level of studies (global test)		.195		.970
Campus (global test)		< .001		.315
Melilla vs Ceuta	0.44 [0.29, 0.68]	< .001		
Socio-economic level (global test)		.014		.731
Employment (yes vs no)	1.35 [0.94, 1.95]	.103	1.02 [0.69, 1.49]	.940
Voluntary work (yes vs no)	1.62 [1.04, 2.53]	.034	0.65 [0.41, 1.04]	.074

Note. Odds ratios (ORs) and 95% confidence intervals are reported. Rows labelled global test correspond to Wald tests for polytomous predictors. For dichotomous predictors, the relevant contrasts are indicated directly in the row labels. Model statistics: IA_APR, $\chi^2(13) = 82.61$, Nagelkerke's $R^2 = .127$, Hosmer-Lemeshow $p = .529$; IA_OCIO, $\chi^2(13) = 32.65$, Nagelkerke's $R^2 = .055$, Hosmer-Lemeshow $p = .507$.

3.3 Attitudinal Profiles and Associations with AI Use

In line with S.O.3, K-means clustering identified six differentiated profiles based on FAC1_1, FAC2_1, IAMAT_MEAN, and age. Table 3 shows that most of the sample was concentrated in three young or early-adult clusters: distrustful moderates (39.0%), young sceptics (24.5%), and young enthusiasts (22.7%). By contrast, distrustful adults (2.4%) and older ambivalent students (1.3%) were very small edge profiles. They are retained in the table for transparency and to describe the full clustering solution, but they should not be used as a basis for substantive generalisation. The main interpretation therefore focuses on the larger and more stable profiles, especially young enthusiasts, young sceptics, distrustful moderates, and moderate adults.

Comparison across the larger profiles revealed a pattern that is not obvious at first sight: greater caution did not necessarily imply lower academic use. The distrustful moderates profile, which comprised 39.0% of the sample, showed a relatively high score in IAMAT_DESCONF (2.91) and still reported academic use above the sample average. This descriptive pattern reinforces the finding from the logistic models, namely that perceived usefulness discriminates adoption more clearly than the mere reduction of reservations about reliability. Patterns observed in the two smallest profiles are reported descriptively only and should be treated as exploratory.

Associations between cluster membership and AI use were not equivalent across the two contexts. The relationship with academic use was significant, $\chi^2(5) = 40.19$, $p < .001$, with a moderate effect size (Cramér's $V = .22$). The highest percentage of IA_APR was observed in young enthusiasts (76.6%), followed by distrustful moderates (63.6%) and young sceptics (59.4%); the lowest value corresponded to older ambivalent students (18.2%). By contrast, the association with leisure use did not reach significance, $\chi^2(5) = 7.17$, $p = .209$, and the effect size was small ($V = .09$). Variation in IA_OCIO was also narrower overall, ranging from 20.0% to 36.4%.

Table 3: Characterisation of cluster profiles in attitudes towards AI and age

Cluster	n (%)	IAMAT_MEAN M (SD)	IAMAT_UTIL M (SD)	IAMAT_DESCONF M (SD)	Age M (SD)
Moderate adults	86 (10.2)	2.66 (0.80)	2.67 (1.12)	2.60 (1.19)	26.66 (1.59)
Distrustful moderates	330 (39.0)	2.57 (0.83)	2.48 (1.10)	2.91 (1.26)	22.32 (1.04)
Young enthusiasts	192 (22.7)	3.51 (0.51)	3.80 (0.63)	2.52 (1.07)	19.31 (0.99)
Young sceptics	207 (24.5)	2.28 (0.54)	2.12 (0.64)	2.84 (1.16)	18.46 (0.96)
Distrustful adults	20 (2.4)	2.33 (0.56)	2.13 (0.84)	3.03 (1.32)	34.95 (3.32)
Older ambivalent students	11 (1.3)	2.53 (1.06)	2.56 (1.53)	2.41 (1.55)	47.36 (3.64)

Note. Clusters were obtained through K-means with $k = 6$ on FAC1_1, FAC2_1, IAMAT_MEAN, and age. Valid $N = 846$.

Figure 1 synthesises the gap between academic and leisure use across clusters. In nearly all profiles, academic use clearly exceeded recreational use, with especially wide gaps among young enthusiasts (41.3 percentage points) and distrustful moderates (36.5 points). The only apparent exception was the profile of older ambivalent students, in which leisure use (36.4%) exceeded academic use (18.2%). However, because this profile included only 11 participants, this reversal is reported for transparency but is not interpreted substantively. The more robust conclusion is that, in the larger profiles, attitudinal segmentation discriminated academic incorporation of AI more effectively than recreational use. Observed counts and row percentages are reported in the Supplementary Material (Table S5).

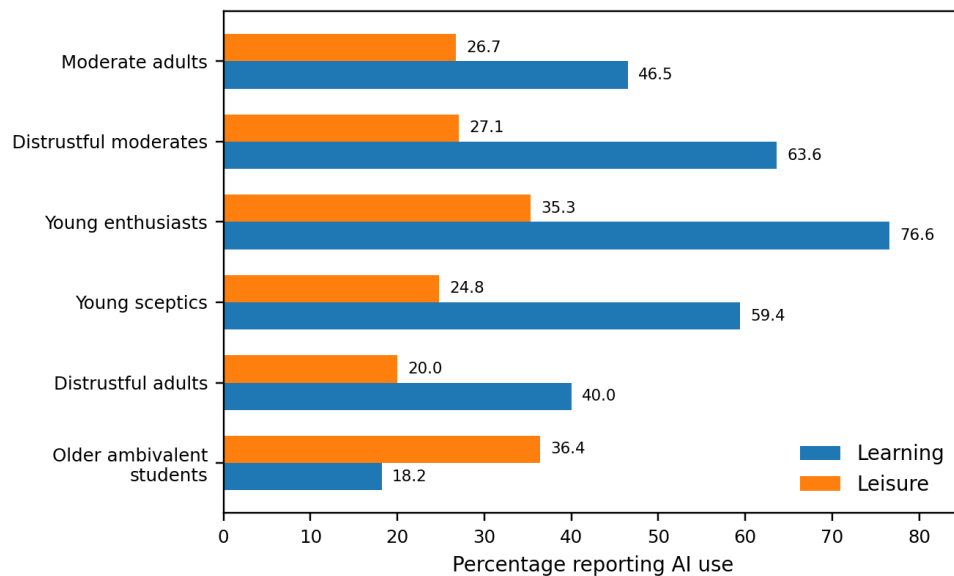


Figure 1: Reported AI use in learning and leisure by cluster

Note. Percentages represent students reporting AI use within each cluster. The association between cluster membership and IA_APR was significant, $\chi^2(5) = 40.19$, $p < .001$, $V = .22$, whereas the association between cluster membership and IA_OCIO did not reach significance, $\chi^2(5) = 7.17$, $p = .209$, $V = .09$.

4. Discussion

4.1 Meaning of the Main Findings

Taken together, the findings that respond to G.O. and, more specifically, to S.O.1 and S.O.2 show that the incorporation of generative AI into university mathematics learning does not follow a binary logic of acceptance versus rejection. What emerges instead is a pattern of selective adoption shaped by the value students attribute to the tool for specific academic purposes. The fact that use in learning substantially exceeds use in leisure suggests that, in this sample, AI is not functioning merely as a general-purpose technological novelty. It is being mobilised above all as an instrumental resource tied to concrete study needs, particularly searching for theory, checking solutions, and generating practice. This reading is consistent with recent work in higher education showing that student adoption of GenAI tends to combine pragmatic openness with persistent concerns about accuracy, privacy, integrity, and learning quality (Bond et al., 2024; Chan and Hu, 2023; Aldreabi et al., 2025).

Within that broader pattern, the most relevant result of the explanatory models is that positive attitudes towards AI in mathematics increased the probability of use in both learning and leisure, whereas the uncertainty/errors dimension was not statistically significant. This absence of effect should not be interpreted as evidence that critical judgement is irrelevant, but rather as a sign that recognising that AI can be wrong does not necessarily prevent students from using it. In a discipline such as mathematics, where checking is itself part of the learning activity, it is plausible that some students combine frequent use with active monitoring of error. The pattern is compatible with, but does not directly measure, a form of calibrated trust. In other words, some students may use AI while remaining aware of possible error and the need for verification. This interpretation should be treated cautiously because calibrated trust was inferred from the coexistence of perceived usefulness, reported use, and awareness of uncertainty, rather than measured as an independent construct. Even so, the pattern is consistent with recent arguments on critical GenAI literacy and with reviews specific to mathematics education that call for reflective, verifiable, and pedagogically purposeful uses of AI (Pepin et al., 2025; Rapanta et al., 2025).

4.2 Academic Versus Leisure Use: Interpretive Scope of the Contrast

The contrast between IA_APR and IA_OCIO should be understood as an analytical boundary test rather than as an attempt to treat leisure use as a direct mathematics-learning outcome. Its contribution lies in helping to distinguish general engagement with GenAI from its situated academic appropriation in mathematics e-learning. Although positive attitudes retained a stable effect in both models, academic use was much more sensitive to contextual and biographical variables than leisure use. The learning model retained age, availability of Wi-Fi,

campus, and voluntary work, whereas the leisure model retained gender as its only robust socio-demographic predictor and showed much lower explanatory power overall. This asymmetry is informative: if the same predictors had explained both contexts with similar strength, the results would have pointed mainly to general digital engagement with AI. Instead, the weaker leisure model helps delimit the specifically academic pattern, suggesting that AI use in mathematics is more closely linked to study tasks, access conditions, and e-learning opportunity structures.

These contextual associations should be interpreted cautiously. The negative effect of age on academic use points to greater incorporation of AI into study routines among younger students, perhaps because of more recent exposure to these tools or because conversational interfaces feel closer to their ordinary digital practices. At the same time, the lower probability of use among students without Wi-Fi reminds us that the integration of AI continues to depend on basic material infrastructures and that the so-called GenAI divide cannot be detached from access conditions (Beckman et al., 2025). The campus contrast requires particular caution because the sample was unevenly distributed across sites, with a clear predominance of Melilla. It should therefore be interpreted as a contextual association within this sample, not as evidence that one campus is intrinsically more or less favourable to AI use. Differences in degree composition, teaching exposure, local digital cultures, or unmeasured institutional conditions may partly account for the observed pattern and should be examined in future stratified studies.

The positive association between voluntary work and academic use may reflect broader dispositions towards agency, social participation, or openness to innovation, although that interpretation requires further testing. In leisure, the lower probability of use among women is compatible with studies reporting demographic differences in perceived readiness, comfort, and benefit when students engage with generative AI (Maxwell et al., 2025). Yet the fact that this gender effect did not appear in the academic model is equally revealing, because it suggests that study-related uses may be less shaped by that gap than more general recreational engagement.

4.3 Implications for Mathematics Education and e-Learning

The findings related to S.O.3 reinforce the view that there is no single way of relating to AI in mathematics. The analytical value of the cluster profiles is diagnostic rather than taxonomic: they should not be read as fixed student types, but as configurations that help identify how perceived usefulness, distrust, and age combine in relation to academic AI use. In this respect, the distrustful moderates profile is especially informative because it shows that awareness of error can coexist with above-average academic use. This does not prove calibrated trust, but it suggests that distrust may sometimes operate as a verification-oriented stance rather than as a simple barrier to adoption. The small edge profiles are reported for completeness but are not used as a basis for substantive interpretation. Overall, the fact that profiles discriminated academic use more clearly than leisure use supports the domain relevance of IAMAT, attitudes towards AI in mathematics appear more informative when the activity is mathematically and academically framed. From the perspective of mathematics education and e-learning, the central implication is not that universities should increase AI use indiscriminately, but that they should improve its epistemic and pedagogical quality. A substantial share of students already use these tools for searching for theory, checking solutions, and practising, yet that does not by itself guarantee mathematically robust learning. Recent literature repeatedly warns that AI can support understanding, immediate feedback, and self-regulation whilst also fostering dependency, premature closure of reasoning, and weakened justification when it is used as a supplier of finished answers (Pepin et al., 2025; Bernardi et al., 2025).

For that reason, teacher education and mathematics e-learning should move the focus away from mere technical prompt proficiency and towards practices of verification and argumentation. Productive task designs might require students to compare human and AI-generated solutions, identify conceptual errors, request alternative explanations, reconstruct omitted steps, justify why an answer is acceptable, or discuss the limits of the tool according to the nature of the task. Such an orientation aligns closely with critical GenAI literacy (Park, 2025; Rapanta et al., 2025), because it treats AI not as an autonomous problem solver but as an object of disciplinary scrutiny.

In practical terms, the findings suggest three complementary lines of action for AI-supported mathematics e-learning. First, educators should guide students towards verification-oriented use. Instead of asking only whether AI was used, learning tasks can require students to document the AI-generated response, identify possible errors or omissions, justify which parts are mathematically valid, and produce a final revised solution. Secondly, AI can be integrated into task design as an object of critique rather than as a shortcut to the answer. For example, students may compare a human solution with an AI-generated solution, detect invalid inferences, reconstruct missing steps, or explain why two apparently plausible solutions are not equivalent. Thirdly,

institutions should support responsible use through access, training, and policy. The relevance of Wi-Fi access in the academic model suggests that AI integration cannot be separated from digital infrastructure, while the role of attitudes indicates that training should address not only technical prompting but also mathematical verification, ethical use, and limits of automation.

In initial teacher education this is especially important: future teachers will not only learn with AI, but will also have to decide when, how, and for what purposes to integrate it into their own teaching.

This requirement becomes even more pressing in the light of evidence on the ambivalent effects of generative AI on learning. Bastani et al. (2025) show that access to GenAI may improve immediate mathematics performance and simultaneously weaken subsequent learning when the tool substitutes for reasoning. Stadler, Bannert and Sailer (2024) likewise found that large language models reduce mental effort but can compromise depth of inquiry. In pre-service mathematics teachers, Zhang et al. (2025) reported that AI dependency is negatively associated with higher-order competences, whereas AI literacy partly buffers that risk. Read against that evidence, the present findings point to both an opportunity and a warning. The opportunity lies in the fact that students already recognise concrete, potentially productive uses of AI. The warning is that more academic use does not automatically amount to better educational integration.

Finally, although the questionnaire underwent expert-based content validation through a Delphi procedure, no separate pilot administration with an independent subsample was conducted. The psychometric results should therefore be understood as initial evidence obtained in the final analytical sample, and future research should replicate the IAMAT structure through confirmatory factor analysis in independent samples.

4.4 Limitations and Research Agenda

The study has several limitations that should shape interpretation of the findings. First, it is a cross-sectional study based on self-report, so it cannot establish causal direction and cannot show how attitudes and uses evolve over time. Secondly, the sample comes from a single university and was recruited by convenience, which restricts external generalisation. In addition, campus distribution was uneven, with a clear predominance of Melilla; therefore, campus contrasts should be interpreted as sample-specific contextual associations rather than as general campus-level effects. Thirdly, the dependent variables register only presence or absence of use, not its frequency, intensity, quality, specific purpose, or the concrete AI tool employed. They should therefore be interpreted as prevalence indicators rather than as measures of how students use AI. From a mathematics-education standpoint, using AI occasionally to check a procedure is not equivalent to using it to delegate the full solution of a task, and this distinction matters substantively. Future studies should include graded indicators of frequency, identify the specific AI tools used, and distinguish task-level purposes such as searching for theory, checking solutions, generating practice exercises, translating verbal problems into mathematical language, or obtaining complete answers. A more specific methodological limitation should also be noted. IAMAT includes usefulness and confidence items framed in relation to concrete mathematical tasks, which gives the instrument high domain relevance but also places it close to the very phenomenon of academic use that it later helps to explain. Part of the association between positive attitudes and AI use may therefore reflect conceptual proximity between predictor and criterion rather than a relation between fully independent constructs. This does not invalidate the findings, but it does suggest that future studies should separate more clearly prior beliefs about AI, accumulated experience of use, actual frequency of use, and effects on mathematical performance. Further work would also benefit from longitudinal designs, measures of solution quality or performance, logged interaction data from specific tools, and comparisons across degree programmes and universities.

5. Conclusions

In relation to G.O, the study shows that generative AI use in university mathematics learning follows a selective and situated logic that is more complex than a reading centred solely on technical availability would suggest. Positive attitudes towards AI in mathematics were associated with a higher probability of use in both learning and leisure, but academic use was clearly more frequent and more sensitive to age, access, and institutional context. The leisure-use findings contribute to this argument by functioning as a contrastive boundary: they show that the academic appropriation of AI in mathematics cannot be reduced to a general recreational orientation towards GenAI. Instead, it appears more closely linked to study-related tasks, e-learning conditions, and domain-specific attitudes. Educationally, the results support a route that avoids both uncritical permissiveness and defensive restriction. AI-supported mathematics learning should be designed around verification, explanation, justification, and revision. In practice, this means asking students not only to obtain AI-generated answers, but to examine their validity, detect errors, reconstruct missing reasoning, and decide when

the tool supports learning and when it replaces the intellectual work that mathematics education seeks to cultivate. In initial teacher education this orientation is especially urgent, because future teachers will need to judge not only whether AI can be used, but when its use strengthens mathematical reasoning and when it undermines it. The central contribution of the study is therefore both theoretical and practical. Theoretically, it shows that AI adoption in mathematics is patterned by domain-specific attitudes, access conditions, and differentiated student profiles rather than by access alone or by generic technology acceptance. Practically, it indicates that e-learning policies and mathematics-education practices should integrate AI through guided, accountable, and verification-oriented tasks, supported by institutional training and equitable access. Recognising this heterogeneity is a necessary step if AI is to strengthen, rather than erode, mathematical reasoning in digital learning environments.

6. Supplementary Material

A separate supplementary file accompanies the article to increase reporting transparency without overloading the main manuscript. It includes the full wording and analytic orientation of the IAMAT items (Table S1), item-level psychometric diagnostics (Table S2), the complete coefficients of the logistic models for IA_APR and IA_OCIO (Tables S3 and S4), and the observed counts and percentages of AI use in learning and leisure by cluster (Table S5).

7. Declarations

AI Statement: Generative AI tools were used during manuscript preparation to support linguistic reorganisation and formal refinement of some draft sections. The selection of evidence, interpretation of results, argumentative decisions, and final revision of the manuscript remained the sole responsibility of the authors.

Ethics Statement: The study received a favourable opinion from the University of Granada Research Ethics Committee (ref. 4518/CEIH/2024, 16 September 2024). All participants gave informed consent before taking part, participation was voluntary, and data were processed anonymously and in aggregated form.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any personal, academic, or financial relationships that could be construed as a potential conflict of interest.

Author Contributions: Conceptualization, Hassan H.M. and Hossein H.M.; software, Hossein H.M.; formal analysis, Hossein H.M. and María del Carmen O.G.; document analysis/investigation, Hassan H.M.; data curation, Hassan H.M. and Hossein H.M.; writing – original draft, Hossein H.M.; writing – review and editing, Hassan H.M. and Manuel G.A.; supervision, Manuel G.A.; project coordination and funding acquisition, María del Carmen O.G.

Funding: This work was supported by the research project MEL-14-UGR24, “*Iluminando oportunidades interseccionales: aprendizaje para la mejora educativa y laboral del uso de la inteligencia artificial (IA) en jóvenes*” [Illuminating intersectional opportunities: learning for educational and professional improvement through the use of artificial intelligence (AI) among young people], funded under the 2024 UGR/Autonomous City of Melilla Research Projects programme.

References

- Aldreabi, H., Dahdoul, N.K.S., Alhur, M., Alzboun, N. and Alsalmi, N.R., 2025. Determinants of student adoption of generative AI in higher education. *Electronic Journal of e-Learning*, 23(1), pp. 15-33. <https://doi.org/10.34190/ejel.23.1.3599>
- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakci, Ö. and Mariman, R., 2025. Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26), e2422633122. <https://doi.org/10.1073/pnas.2422633122>
- Bauer, E., Greiff, S., Graesser, A.C., Holstein, K., Rummel, N., Hmelo-Silver, C.E., Aleven, V. and McNamara, D.S., 2025. Looking beyond the hype: Understanding the effects of AI on learning. *Educational Psychology Review*, 37, 45. <https://doi.org/10.1007/s10648-025-10020-8>
- Beckman, K., Apps, T., Howard, S.K., Rogerson, C., Rogerson, A. and Tondeur, J., 2025. The GenAI divide among university students: A call for action. *The Internet and Higher Education*, 67, 101036. <https://doi.org/10.1016/j.iheduc.2025.101036>
- Bernardi, M.L., Capone, R., Faggiano, E. and Rocha, H., 2025. Generative AI in mathematics education: Pre-service teachers' knowledge and implications for their professional development. *International Journal of Mathematical Education in Science and Technology*, 56(8), pp. 1513-1530. <https://doi.org/10.1080/0020739X.2025.2490104>
- Bewersdorff, A., Hornberger, M., Nerdel, C. and Schiff, D.S., 2025. AI advocates and cautious critics: How AI attitudes, AI interest, use of AI, and AI literacy build university students' AI self-efficacy. *Computers and Education: Artificial Intelligence*, 8, 100340. <https://doi.org/10.1016/j.caeai.2024.100340>

- Bond, M., Khosravi, H., de Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S.W. and Siemens, G., 2024. A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, 21, 4. <https://doi.org/10.1186/s41239-023-00436-z>
- Busuttil, L. and Calleja, J., 2025. Teachers' beliefs and practices about the potential of ChatGPT in teaching mathematics in secondary schools. *Digital Experiences in Mathematics Education*, 11, pp. 140-166. <https://doi.org/10.1007/s40751-024-00168-3>
- Chan, C.K.Y. and Hu, W., 2023. Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20, 43. <https://doi.org/10.1186/s41239-023-00411-8>
- Contel, F. and Cusi, A., 2025. Investigating the role of ChatGPT in supporting metacognitive processes during problem-solving activities. *Digital Experiences in Mathematics Education*, 11, pp. 167-191. <https://doi.org/10.1007/s40751-024-00164-7>
- Deng, R., Jiang, M., Yu, X., Lu, Y. and Liu, S., 2025. Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, 105224. <https://doi.org/10.1016/j.compedu.2024.105224>
- Hsu, C.-C. and Sandford, B.A., 2007. The Delphi technique: Making sense of consensus. *Practical Assessment, Research and Evaluation*, 12(10), pp. 1-8. <https://doi.org/10.7275/pdz9-th90>
- Jin, Y., Yan, L., Echeverria, V., Gašević, D. and Martinez-Maldonado, R., 2025. Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, 8, 100348. <https://doi.org/10.1016/j.caeai.2024.100348>
- Kock, Z.-J., Salinas-Hernández, U. and Pepin, B., 2025. Engineering students' initial use schemes of ChatGPT as an instrument for learning. *Digital Experiences in Mathematics Education*, 11, pp. 192-218. <https://doi.org/10.1007/s40751-025-00169-w>
- Lynn, M.R., 1986. Determination and quantification of content validity. *Nursing Research*, 35(6), pp. 382-386. <https://doi.org/10.1097/00006199-198611000-00017>
- Maxwell, D., Oyarzun, B., Kim, S. and Bong, J.Y., 2025. Generative AI in higher education: Demographic differences in student perceived readiness, benefits, and challenges. *TechTrends*, 69, pp. 1248-1259. <https://doi.org/10.1007/s11528-025-01109-6>
- Noster, N., Gerber, S. and Siller, H.-S., 2024. Pre-service teachers' approaches in solving mathematics tasks with ChatGPT. *Digital Experiences in Mathematics Education*, 10, pp. 543-567. <https://doi.org/10.1007/s40751-024-00155-8>
- Park, J., 2025. A systematic literature review of generative artificial intelligence literacy in schools. *Computers and Education: Artificial Intelligence*, 9, 100487. <https://doi.org/10.1016/j.caeai.2025.100487>
- Peduzzi, P., Concato, J., Kemper, E., Holford, T.R. and Feinstein, A.R., 1996. A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, 49(12), pp. 1373-1379. [https://doi.org/10.1016/S0895-4356\(96\)00236-3](https://doi.org/10.1016/S0895-4356(96)00236-3)
- Pepin, B., Buchholtz, N. and Salinas-Hernández, U., 2025. A scoping survey of ChatGPT in mathematics education. *Digital Experiences in Mathematics Education*, 11, pp.9-41. <https://doi.org/10.1007/s40751-025-00172-1>
- Pham, L., Klaus, T. and Bateh, J., 2025. Key factors influencing intention to use ChatGPT: An empirical study of U.S. students. *Acta Psychologica*, 260, 105516. <https://doi.org/10.1016/j.actpsy.2025.105516>
- Rapanta, C., Bhatt, I., Bozkurt, A., Chubb, L.A., Erb, C., Forsler, I., Gravett, K., Koole, M., Lintner, T., Örtengren, A., Petricini, T., Rodgers, B., Webster, J., Xu, X., Christensen, I.-M.F., Dohn, N.B., Christensen, L.L.W., Zeivots, S. and Jandrić, P., 2025. Critical GenAI literacy: Postdigital configurations. *Postdigital Science and Education*, 7, pp. 1296-1333. <https://doi.org/10.1007/s42438-025-00573-w>
- Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, pp. 53-65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Stadler, M., Bannert, M. and Sailer, M., 2024. Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, 160, 108386. <https://doi.org/10.1016/j.chb.2024.108386>
- Tavakol, M. and Dennick, R., 2011. Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, pp. 53-55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C. and Vandenbroucke, J.P., 2007. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *PLoS Medicine*, 4(10), e296. <https://doi.org/10.1371/journal.pmed.0040296>
- Watkins, M.W., 2018. Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, 44(3), pp. 219-246. <https://doi.org/10.1177/0095798418771807>
- Yoon, H., Hwang, J., Lee, K., Roh, K.H. and Kwon, O.N., 2024. Students' use of generative artificial intelligence for proving mathematical statements. *ZDM - Mathematics Education*, 56, pp. 1531-1551. <https://doi.org/10.1007/s11858-024-01629-0>
- Zhang, D., Wijaya, T.T., Wang, Y., Su, M., Li, X. and Damayanti, N.W., 2025. Exploring the relationship between AI literacy, AI trust, AI dependency, and 21st century skills in preservice mathematics teachers. *Scientific Reports*, 15, 14281. <https://doi.org/10.1038/s41598-025-99127-0>

Pedagogical Alignment, Adaptivity, and Analytics in Intelligent Tutoring Systems: A Systematic Review

Ilyass Houssam, Zouhair Chiba and Mounia Miyara

LIS Labs, Faculty of Sciences Aïn Chock, Hassan II University of Casablanca, Morocco

ilyasshoussam9@gmail.com (corresponding author)

<https://doi.org/10.34190/ejel.24.4.4779>

An open access article under [CC Attribution 4.0](#)

Abstract: Intelligent tutoring systems have changed quickly since large language models became widely available in 2023, raising a practical question for e-learning research: when a tutor is built on a general-purpose language model rather than on hand-encoded rules, what happens to its pedagogical grounding, learner adaptivity, and reproducibility infrastructure? Earlier reviews have examined these matters separately, but none has considered how they come together in the systems now being built. This systematic review addresses that gap through three analytical axes: pedagogical alignment with an instructional theory, adaptivity and learner modelling, and analytics standards supporting comparison and reproducibility. The review followed the SPAR-4-SLR protocol and used PRISMA 2020 as a reporting framework where the recoverable record allowed. Candidate records were screened, verified against their source documents, and coded conservatively against pre-specified criteria. Twelve primary studies were retained after source verification. Pedagogical alignment and adaptivity were each addressed by eleven of the twelve studies (92 percent), although implementation varied widely, from explicit frameworks such as Cognitive Apprenticeship, Productive Failure, and Socratic questioning to looser prompt-based behaviour. Analytics standards and reproducibility were the weakest axis, addressed by three studies (25 percent). No study combined all three axes with explicit uncertainty quantification or calibration of tutoring decisions. Controlled learning-outcome evidence was also scarce: only one study reported a controlled comparison, and its findings were preliminary. For e-learning practice, the review gives teachers, platform designers, and institutional decision-makers a realistic basis for expectation: current language-model tutoring systems are promising and often pedagogically plausible, but their learning benefits remain under-demonstrated and should be evaluated locally before large-scale adoption. For e-learning knowledge, the review contributes a source-verified synthesis of recent intelligent tutoring research and identifies a clear methodological gap: the field is progressing faster in system design than in shared evaluation, interoperable analytics, and calibrated decision-making. The paper concludes that future work should combine instructional theory, inspectable learner models, standardised logging, open benchmarks, and uncertainty-aware evaluation within the same tutoring systems.

Keywords: Intelligent tutoring systems, Large language models, Learner modelling, Learning analytics, Pedagogical alignment, Systematic review

1. Introduction

Intelligent tutoring systems have occupied educational technology research for close to four decades. The early motivation was Bloom's (1984) finding that one-to-one human tutoring produced gains of around two standard deviations over conventional classroom teaching, a result that set a target successive generations of computer-based instruction have tried to reach. Meta-analyses since then have found real but more modest effects. VanLehn (2011) reports a Cohen's d of about 0.76 for step-based tutoring against no tutoring but only 0.31 for answer-based systems, Steenbergen-Hu and Cooper (2014) report between 0.32 and 0.37 for college students, and Ma et al. (2014) and Kulik and Fletcher (2016) report comparable ranges across domains. These are useful effects, but they sit well short of two standard deviations.

The situation has shifted quickly since 2023. Large language models that hold a natural conversation have entered tutoring research at speed, and a system can now cover a wide range of subject matter and produce pedagogically plausible responses without the extensive rule encoding earlier systems required. The change is real, but it has not settled into a clear picture. It is not yet known whether a tutor built on a language model can match the pedagogical fidelity of one built on instructional theory, nor how a coherent account of what a learner knows can be maintained when the tutoring agent does not itself retain state. Comparison is harder still, because systems built in different research traditions rarely share outcome measures or data formats.

Several earlier reviews have examined parts of this territory. Zawacki-Richter et al. (2019) surveyed artificial intelligence in higher education and noted how rarely educators were involved, Roll and Wylie (2016) traced the field's longer trajectory, and Luckin et al. (2016) set out an early agenda for the area. None of these, however, has examined how pedagogical grounding, adaptive mechanisms, and reproducibility infrastructure come together within the language-model systems now being built. Part of the reason is disciplinary: computer

scientists tend to focus on algorithms, learning scientists on instructional design, and educational data miners on the analysis of log data, and these communities seldom work in concert.

This review addresses that gap through three analytical axes. Pedagogical alignment concerns whether a system implements an instructional theory, such as scaffolding, mastery learning, Cognitive Apprenticeship, or Productive Failure, rather than merely naming one. Adaptivity and learner modelling concern how personalisation is achieved, whether through a classical probabilistic model such as Bayesian Knowledge Tracing, through retrieval-augmented generation, or through the emergent behaviour of a language model. Analytics standards and reproducibility concern a system's engagement with the infrastructure that allows replication and comparison, including interoperability standards such as the Experience API (xAPI) and a Learning Record Store, the release of source code, and the use of shared benchmarks. Three research questions follow. RQ1: to what extent, and with what fidelity, do these systems realise explicit pedagogical theories? RQ2: what forms of adaptivity and learner modelling are used, and how are they implemented? RQ3: which analytics and interoperability standards are adopted, and what do they allow in terms of comparison and replication?

A distinguishing feature of this review is that every included study was checked against its actual source document. As Section 2 explains, this step proved necessary: a first pass found that several frequently cited bibliographic records did not match the papers they named. The verified corpus is therefore smaller than the screening totals might suggest, but each study in it has been confirmed to exist and to report what is attributed to it. Section 2 describes the protocol, the verification procedure, and the coding framework. Section 3 presents the findings axis by axis and how the axes occur together. Section 4 interprets them against the meta-analytic record and for e-learning practice. Sections 5 and 6 set out the limitations and the priorities for future work.

2. Materials and Methods

The review followed the SPAR-4-SLR protocol (Paul et al., 2021), which divides the process into assembling, arranging, and assessing, and requires each decision to be documented. The protocol was chosen because its emphasis on logging decisions suited a corpus expected to mix classical tutoring systems with recent language-model systems that report their work in very different ways, and because its audit trail partly offsets the limitations of a review conducted by a single reviewer.

2.1 Reviewer Design, Verification, and Audit Trail

Screening, source verification, extraction, and coding were conducted by the lead reviewer and logged in an audit trail. To reduce classification error, the review used pre-specified inclusion and exclusion criteria, conservative binary coding, full-text source verification, and an evidence note for every coding decision. Ambiguous cases were coded as not addressing an axis unless direct textual evidence supported inclusion. The complete coding sheet is provided as Supplementary Table S1 (Appendix A). Although the design does not include an independent second-coder reliability estimate, and Cohen's kappa is therefore not reported, the audit trail allows readers to inspect the basis for each inclusion and coding decision, and independent double coding is identified in Section 5 as the principal extension for future work.

Source verification was central to the design. To strengthen source reliability, every candidate study was opened and checked against its bibliographic record before coding, so that the synthesis rests on verified full-text documents rather than on citation metadata. This check identified a set of records whose details did not match the documents they named: five of the files initially gathered were different papers from the ones cited, one described artificial-agent learning rather than human tutoring, and one had been posted after the search window. These records were not used as primary evidence, and their disposition is reported in Supplementary Table S2 (Appendix B). The synthesis therefore rests on the twelve source-verified studies.

2.2 Assembling: Scope, Search, and Screening

The scope was defined around the three axes above. A study was included if it was a peer-reviewed article, a full conference paper, or a substantial preprint; if it described or evaluated a working tutoring system, or analysed one in a way that bore directly on an axis; if it was in English and within the search window; and if it engaged with at least one axis in a substantive rather than a passing way. A study was excluded if it addressed an axis only superficially, dealt with purely technical matters unrelated to pedagogy or learning, offered only speculation, was a duplicate, fell outside the window, or, after verification, could not be confirmed to be the work that had been cited.

2.2.1 Search strategy

Six databases were searched, chosen to balance disciplinary coverage: Scopus and Web of Science Core Collection for indexed multidisciplinary and high-impact work, IEEE Xplore and the ACM Digital Library for computer science and human-computer interaction, SpringerLink for education and the learning sciences, and arXiv for the recent preprints in which much language-model tutoring first appears. The master query combined five groups of terms, joined with the Boolean operator AND: terms for intelligent tutoring and pedagogical agents; terms for pedagogical alignment, including named theories such as scaffolding, mastery learning, Cognitive Apprenticeship, and Productive Failure; terms for adaptivity and learner modelling, including knowledge tracing and retrieval-augmented generation; terms for learning analytics and interoperability standards, including xAPI and the Learning Record Store; and terms for the educational context. Each group was a set of synonyms joined with OR and wildcarded where appropriate. The query was then adapted to the field-tag conventions of each database. The full master query and its per-database adaptations are provided as supplementary material (Appendix C, Supplementary Table S3). After retrieval, records were deduplicated by matching on title and DOI, and backward and forward citation searches were run on the studies that passed full-text screening. Because the original platform-level search export was not recoverable at the revision stage, the upstream identification and deduplication counts are not reported; to preserve transparency, the review instead documents all full-text inclusion decisions and all source-verification exclusions in the supplementary material, and restricts the synthesis to the twelve studies whose source documents were opened and verified directly (Figure 1).

The main searches covered records available up to 30 June 2025, with backward and forward citation checks conducted during revision. Searches were restricted to English-language records within the review window. Each candidate source was opened and checked against its bibliographic record before coding, and a record was retained only where the source document matched the cited study and reported evidence relevant to at least one review axis. Because the original platform-level exports were unavailable at revision, the review does not reconstruct upstream identification or deduplication counts retrospectively. The supplementary materials instead document the recoverable audit trail (Appendices A to D): the search strategy, the source-verification decisions and exclusion reasons, the coding evidence, the quality assessment, and the final verified corpus.

2.2.2 Note on the uncertainty dimension

An earlier version of this review also examined whether systems quantified the uncertainty of their own decisions and calibrated their confidence against learner behaviour. Because no included study did so, that dimension is not reported as a separate results axis; it is taken up in the conclusion, where its absence is the clearest single gap the review identifies.

2.3 Arranging and Assessing

A structured form recorded each study's bibliographic details, design, domain, and participants, together with a coding of each axis as addressed (1) or not (0) and a short evidence note quoting or locating the supporting passage. Where it was unclear whether a study addressed an axis, it was coded as not addressing it, so the figures reported below are conservative. A study was coded as addressing pedagogical alignment only where it implemented or substantively operationalised an instructional theory, not where it merely named one; as addressing adaptivity where it personalised instruction on the basis of an inferred or retrieved learner state; and as addressing analytics where it engaged with interoperability standards, released usable code or data, or used a shared benchmark.

Each study was also scored against five quality criteria, each from 0 to 2 for a maximum of ten, set out in Table 1: clarity of the research questions, appropriateness of the methods, sufficiency of implementation detail, rigour of the evaluation, and clarity of reporting. Scores across the twelve studies ranged from 6 to 8, with a median of 7 (Appendix D, Supplementary Table S4). The two lowest-scored studies were the authoring platform and the tutor-assessment study, both of which evaluate something other than learning with students; they are retained in the synthesis but weighted lightly wherever a claim rests on several sources. Because the studies differed so much in outcomes, designs, and domains, they were synthesised narratively rather than pooled.

Table 1: Quality assessment rubric (five criteria, maximum score 10)

Criterion	0	1	2
Research questions clearly stated	No questions	Implicit or vague	Clear and specific

Criterion	0	1	2
Methods appropriate and described	Not described	Partly described	Fully described
Implementation detail sufficient	No detail	Some detail	Reproducible
Evaluation rigorous	No evaluation	Informal or small	Formal study
Findings clearly reported	Unclear	Partly reported	Fully reported

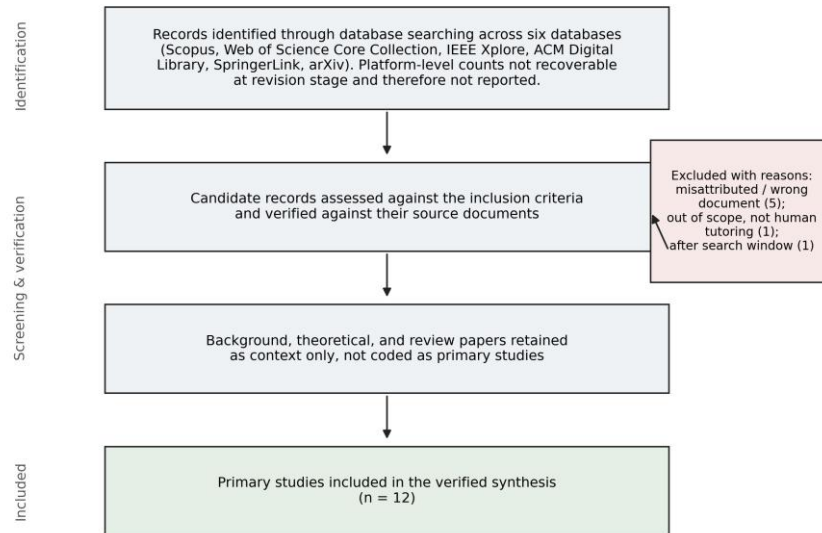


Figure 1: Flow of study selection and source verification. The diagram reports the recoverable source-verification pathway and the final verified corpus; upstream platform-level identification and deduplication counts were not reconstructed retrospectively, and the detailed search strategies, exclusion reasons, and coding decisions are provided in the supplementary materials (Appendices A to D)

3. Results

3.1 Characteristics of the Included Studies

The twelve verified studies span 2018 to 2025, and ten of them appeared in 2024 or 2025, which reflects the recent surge of language-model tutoring. Most are recent preprints rather than journal articles, a point returned to in the limitations. The domains were varied: language learning, programming, mathematics, electrical engineering, and multi-domain platforms. Most studies presented a system together with a technical or usability evaluation; only one reported a controlled comparison with a learning-outcome measure. Two of the twelve are included with the reservations noted in Section 2.3: one is an authoring platform rather than a tutor in use, and one assesses tutors' responses rather than tutoring learners directly. Both are retained for completeness and weighted lightly; a reader who excluded them would see ten studies and a slightly higher analytics proportion. They were retained because they directly evaluate tutor behaviour, authoring, or tutoring infrastructure relevant to language-model intelligent tutoring, but they are interpreted cautiously throughout, since they do not provide full learning-outcome evidence from a deployed tutoring system.

Table 2 summarises the included studies and their coding. The full study-by-study coding, with an evidence note for every judgment, is in Supplementary Table S1 (Appendix A).

By domain, three studies addressed language learning (Kwon et al., 2024; Liu et al., 2024; Liu et al., 2025), two mathematics (Puech et al., 2024; Kakarla et al., 2024), two electronics or electrical engineering (Graesser et al., 2018; Knievel, Bernhardt and Bernhardt, 2025), one programming (Li et al., 2025), and one the development of research skills (Degen and Asanov, 2025), with the remaining three operating across domains as platforms or testbeds (Smith, Gupta and MacLellan, 2024; Weitekamp, Siddiqui and MacLellan, 2025; Borchers and Shou, 2025). By design, the corpus is dominated by system descriptions paired with a technical evaluation, such as the accuracy of a component or an ablation of the architecture. One study reported a controlled comparison with a

learning measure (Degen and Asanov, 2025), one a usability study with instructors (Smith, Gupta and MacLellan, 2024), and one a systematic benchmarking of model behaviour against an existing tutoring system (Borchers and Shou, 2025). Only the established ElectronixTutor (Graesser et al., 2018) predates the recent wave; the other eleven are concentrated in 2024 and 2025. This distribution shapes what the review can conclude: it describes a young, design-led literature in which working systems are plentiful but controlled evidence of their effect on learning is still scarce.

Table 2: The twelve included studies and their coding on the three axes

Study	Year	Domain	Ped.	Adapt.	Analyt.
BIPED (Kwon et al.)	2024	Language (ESL)	Yes	Yes	No
CogGen (Li et al.)	2025	Programming	Yes	Yes	No
AITEE (Knievel, Bernhardt and Bernhardt)	2025	Electrical eng.	Yes	Yes	No
Socratic AI (Degen and Asanov)	2025	Research skills	Yes	Yes	No
Scaffolding tutor (Liu et al.)	2024	Language	Yes	Yes	No
SingaKids (Liu et al.)	2025	Language	Yes	Yes	No
Productive Failure tutor (Puech et al.)	2024	Mathematics	Yes	Yes	No
Apprentice Tutor Builder (Smith, Gupta and MacLellan)	2024	Multi-domain	No	Yes	No
Tutor assessment (Kakarla et al.)	2024	Mathematics	Yes	No	No
TutorGym (Weitekamp, Siddiqui and MacLellan)	2025	Multi-domain	Yes	Yes	Yes
ElectronixTutor (Graesser et al.)	2018	Electronics	Yes	Yes	Yes
Adaptivity benchmark (Borchers and Shou)	2025	Multi-domain	Yes	Yes	Yes

Note: Ped. = pedagogical alignment; Adapt. = adaptivity and learner modelling; Analyt. = analytics standards and reproducibility.

3.2 Distribution Across the Three Axes

Pedagogical alignment and adaptivity were each addressed by eleven of the twelve studies (92 percent). Analytics standards were addressed by three (25 percent). Figure 2 shows the distribution. The imbalance is the central descriptive finding: designing a system that is pedagogically informed and adaptive is now ordinary practice, but engaging with the shared infrastructure that would let one system be compared with another is not.

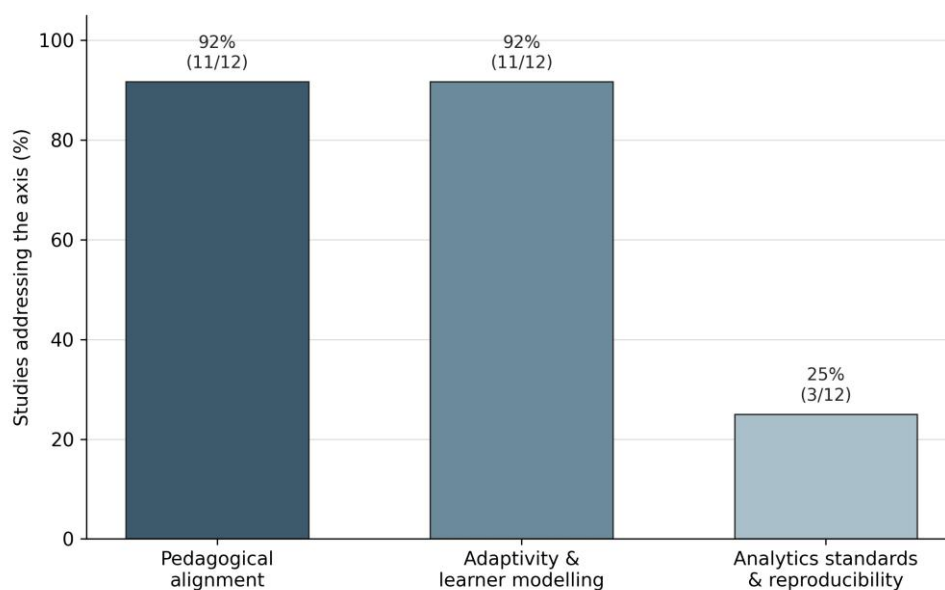


Figure 2: Proportion of the twelve included studies addressing each axis

3.3 Pedagogical Alignment (RQ1)

Most studies engaged with pedagogy, but the depth varied. The strongest cases implemented a named instructional theory in a way that shaped the system's behaviour. CogGen (Li et al., 2025) is built on the Cognitive Apprenticeship framework, mapping its six teaching moves, from modelling through coaching, scaffolding, articulation, and reflection to exploration, onto a domain-specific language that structures each generated turn around programming videos. The Productive Failure tutor of Puech et al. (2024) operationalises Kapur's account of productive failure as an explicit steering strategy, in which the tutor deliberately withholds direct instruction so that learners first generate and explore their own, often flawed, solutions. AITEE (Knievel, Bernhardt and Bernhardt, 2025) and the Socratic AI tutor of Degen and Asanov (2025) implement Socratic questioning to keep the initiative with the learner, the former in electrical engineering and the latter in the development of research questions, where it scaffolds the learner through an adapted PICOT structure. The two scaffolding-based systems (Liu et al., 2024; Liu et al., 2025) make graduated support explicit, fading prompts as competence grows; the more recent of the two, a multilingual dialogic tutor for young language learners, treats the calibration of scaffolding to proficiency as a design target in its own right.

BIPED (Kwon et al., 2024) takes a distinctive empirical route to pedagogical grounding. Rather than encoding a single theory, its authors annotated recordings of human tutoring to derive a lexicon of thirty-four tutor dialogue acts and nine student acts, then built a two-step system that first predicts an appropriate tutor act for the current turn and only then generates an utterance to realise it. This grounds the system's moves in observed tutoring practice and makes its pedagogical intent inspectable turn by turn, although the reported accuracy of act prediction is modest, which the authors acknowledge as a limit on how reliably the intended move is carried out.

Two patterns qualify the picture. First, pedagogical grounding was often asserted more firmly than it was tested: few studies checked whether a pedagogically structured design produced better learning than an unstructured one, so the link between the theory invoked and any measured benefit usually remains an assumption. Second, the systems built on language models tend to realise pedagogy through prompting rather than through encoded rules, so the intended strategy is encouraged rather than guaranteed, and several authors note that behaviour varies from one run to the next. The authoring platform of Smith, Gupta and MacLellan (2024) was coded as not implementing a pedagogical theory itself, since the pedagogy resides in the tutors a user builds with it rather than in the platform; its contribution is to lower the expertise needed to author a tutor at all, by letting a non-programmer demonstrate worked examples from which an apprentice-learning agent induces an expert model.

3.4 Adaptivity and Learner Modelling (RQ2)

Adaptivity was the most developed axis and took three broad forms. A first group keeps an explicit learner model. CogGen (Li et al., 2025) uses Bayesian Knowledge Tracing to estimate mastery and decide when to advance, so that the Cognitive Apprenticeship sequence is driven by an inferred learner state rather than by a fixed script. The testbed of Weitekamp, Siddiqui and MacLellan (2025) and the established ElectronixTutor (Graesser et al., 2018) both track mastery over knowledge components, the latter through a recommender that draws on a persistent student model combining cognitive and non-cognitive attributes. These systems inherit the interpretability of the classical tradition: the basis for each adaptive decision can be read off the model. A second group adapts through retrieval. AITEE (Knievel, Bernhardt and Bernhardt, 2025) uses a graph-based similarity measure to pull the most relevant lecture material into the prompt and pairs this with a circuit simulator, so that adaptation is a matter of supplying the model with the right context rather than of maintaining a model of the learner.

A third and now common group relies on the language model itself, with adaptivity emerging from the prompt context and the conversation history rather than from a maintained model of the learner (Degen and Asanov, 2025; Kwon et al., 2024; Liu et al., 2024; Liu et al., 2025; Puech et al., 2024). These systems adjust the depth of an explanation, the difficulty of a question, or the wording of a hint in response to what a learner has just said, and they do so fluently. What they generally do not do is represent the learner's state in a form that persists across a session or that can be inspected, which makes their adaptation harder to audit and, as several authors note, less stable from one interaction to the next.

Two limits run through the corpus. The systems that adapt through a language model rarely report a quantitative measure of how precisely they adapt, and almost none compares adaptation driven by a language model against a model-based control. The clearest evidence on this point comes from Borchers and Shou (2025), whose benchmarking study varied the contextual information given to several models across seventy-five real tutoring

scenarios and assessed the resulting moves for both adaptivity and pedagogical soundness. Even the best-performing model only marginally matched the adaptivity of an established tutoring system, and the model judged most pedagogically sound was also the least reliable at following instructions. That result suggests the gap between emergent and modelled adaptivity is real and measurable, and that fluency in dialogue should not be mistaken for precision in adaptation.

3.5 Analytics Standards and Reproducibility (RQ3)

Analytics and reproducibility infrastructure formed the weakest axis, addressed by three of the twelve studies. ElectronixTutor (Graesser et al., 2018) records activity to a Learning Record Store, the interoperability mechanism associated with the Experience API. TutorGym (Weitekamp, Siddiqui and MacLellan, 2025) is an open, shared testbed with a public code repository that lets different agents be evaluated against the same tasks. Borchers and Shou (2025) release open benchmarking code and a reproducible procedure. The remaining nine studies released, at most, a dataset or a dependency, and none adopted a community xAPI profile or a shared outcome measure that would let its results be placed alongside another system's.

The consequences run through the synthesis. No two of the language-model studies used the same outcome measure on a comparable population, which rules out any quantitative pooling, and the scarcity of open implementations makes external replication difficult. Where code or data were released at all, it was typically a model dependency or a one-off dataset rather than anything designed for reuse, and outcome reporting ranged from automatic accuracy metrics to small usability studies, with no shared definition of what a successful tutoring interaction is. The two pieces of genuine shared infrastructure in the corpus, TutorGym and the benchmark of Borchers and Shou, are recent and point to what a more cumulative version of the field might look like; notably, both come from research groups working on the evaluation problem itself rather than on a particular tutor, which may be why analytics is so much better served in their work than in the system-building studies that make up the rest of the corpus.

3.6 Integration Across Axes

Pedagogical alignment and adaptivity occurred together in ten of the twelve studies, the dominant pairing and an unsurprising one. Analytics co-occurred with each of the other two axes in only three studies. Figure 3 shows the full pattern. Three studies addressed all three axes at once: ElectronixTutor (Graesser et al., 2018), TutorGym (Weitekamp, Siddiqui and MacLellan, 2025), and the benchmarking study of Borchers and Shou (2025). It is worth noting that the most integrated work is either an established pre-language-model system or recent infrastructure built deliberately for comparison, rather than one of the new conversational tutors, which cluster on pedagogy and adaptivity and leave analytics aside.

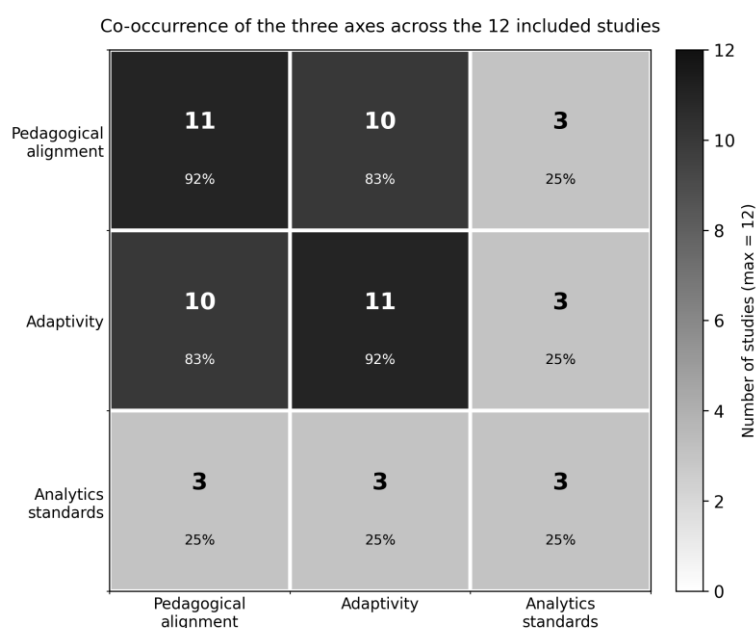


Figure 3: Co-occurrence of the three axes across the twelve included studies (diagonal cells give the total per axis)

This pattern is the review's clearest structural result. The newest conversational tutors are strong on pedagogy and adaptivity but treat analytics as an afterthought, while the studies that take analytics seriously are not themselves novel tutors. No single study brings the three strengths together, and, as the conclusion discusses, none adds a treatment of decision uncertainty on top. The field is therefore advancing along separate tracks: a system-building track that produces fluent, pedagogically framed tutors whose behaviour is hard to verify, and an evaluation track that produces the infrastructure to verify systems but few new tutors to run through it. The value of the most integrated work in the corpus lies in showing that the tracks can meet, but they do so rarely, and never yet in a single recent language-model tutor.

4. Discussion

4.1 Main Findings

The review describes a field that is technically sophisticated on the two axes its researchers attend to most, pedagogical alignment and adaptivity, and underdeveloped on the one that depends on cross-disciplinary infrastructure. Pedagogy is widely claimed but unevenly implemented, and in the language-model systems it is encouraged through prompting rather than guaranteed by design. Adaptivity is the most mature axis, yet the shift from modelled adaptivity to emergent, prompt-driven adaptation is exactly where the evidence is thinnest, and the one benchmarking study that bears on it suggests that language models still trail established systems in matching a learner's needs.

4.2 The Analytics gap

The clearest structural finding is that analytics standards are underdeveloped, and that this holds back the accumulation of knowledge. Most of the recent systems cannot be compared with one another, because they define outcomes differently and rarely share code or data. The obstacle is one of coordination rather than technology: the standards exist, and the two shared resources in the corpus show they can be used, but adopting them requires the kind of cooperation the field's fragmentation makes difficult. Even a small set of agreed benchmarks and a shared logging profile would change what comparison and synthesis are possible.

It is worth dwelling on why this matters more than it might first appear. A field that cannot compare its systems cannot tell which design choices are responsible for which outcomes, and so it cannot improve cumulatively; each new system starts an argument rather than settling one. The neighbouring field of educational data mining made progress in part by building shared datasets and repositories that allowed the same question to be asked of many systems, and the analytics tradition has long argued that interoperability is a precondition for that kind of advance (Baker and Siemens, 2014). The tutoring subfield has not yet followed. The two exceptions in this corpus are instructive: TutorGym is valuable precisely because it lets different agents be run against the same tasks, and the contribution of Borchers and Shou (2025) is as much the reusable benchmarking procedure as the particular result. Both point to a version of the field in which claims about adaptivity or pedagogy could be checked against a common yardstick rather than asserted system by system.

4.3 Comparison with the Meta-analytic Record

Where causal evidence exists, it is consistent with the wider record. VanLehn (2011), Steenbergen-Hu and Cooper (2014), Ma et al. (2014), and Kulik and Fletcher (2016) place intelligent tutoring in the moderate range, well below the two-sigma mark of expert human tutoring. The verified corpus adds little new causal evidence of its own. Only the Socratic AI study (Degen and Asanov, 2025) reports a controlled comparison, contrasting a Socratic tutor with an uninstructed chatbot on a pre-test and post-test design, and the authors describe their findings as preliminary and plan to extend the sample. The Productive Failure tutor of Puech et al. (2024) reports an encouraging field test, but the relevant comparison concerns the fidelity of the tutoring strategy rather than a controlled learning-outcome contrast, so it is not read here as an effect-size estimate. On present evidence, language-model tutors look likely to land in the same moderate range as their predecessors rather than to transform it, and the field still lacks the controlled comparisons that would settle the question.

4.4 Implications for e-Learning Practice

For researchers, the analytics axis is where the greatest leverage lies, since the alternative to shared evaluation is a continuing accumulation of case studies that cannot be compared.

For those who design systems, language-model tutors call for particular attention to pedagogical fidelity, whether through hybrid architectures that pair a language model with a persistent learner model or through carefully constrained prompting. The corpus hints at what such hybrids might look like: CogGen's use of Bayesian

Knowledge Tracing alongside a generative front end, and the way TutorGym and ElectronixTutor anchor adaptation in an explicit model of knowledge components, suggest that the interpretability of the classical tradition and the fluency of the generative one are not mutually exclusive. A tutor that generates dialogue freely but defers decisions about what to teach next to an inspectable learner model would address the main weakness of the purely prompt-based systems, namely that their adaptation cannot be audited or guaranteed.

For teachers and other practitioners introducing these tools, the review offers a realistic basis for expectation: the benefits are plausible but largely unproven, the systems work best as a supplement to classroom instruction rather than a replacement for it, and material aimed at practitioners should reflect those limits rather than present language-model tutoring as more transformative than the evidence supports. The honest message for a school or university weighing adoption is that these systems are promising and worth piloting, but that their effectiveness for a particular cohort is an open question that local evaluation, rather than vendor demonstration, should answer.

4.5 Beyond the Three Axes

Two issues fall outside the coding framework but bear on the field's overall effect. The first is privacy: conversational tutors process detailed records of interaction and material that can reveal sensitive cognitive or emotional states, and the privacy-protecting architectures and governance well developed elsewhere in artificial intelligence are little used in education. The second is equity of access: the promise of intelligent tutoring has been to widen access to personalised instruction, yet subscription platforms and the infrastructure they need risk directing the best tools towards the learners who need them least. The studies reviewed here do not address either question systematically, and a future review that treated them as analytical dimensions in their own right would add something the current literature lacks.

5. Limitations

Several limitations define the scope of the conclusions. First, the review was conducted by a single reviewer; although conservative coding, source verification, and an auditable extraction sheet reduced the risk of classification error, future work should extend the design through independent double coding and inter-rater reliability reporting. Second, the verified corpus is small and includes several recent preprints, reflecting the early stage of language-model tutoring research, so some findings may settle as that work reaches final publication. Third, controlled learning-outcome evidence remains scarce, so the review is best interpreted as a source-verified map of current design practice rather than as a pooled estimate of effectiveness. The three-axis framework is also one reasonable lens among several, and a framework built around affective computing, multimodal interaction, or accessibility would bring different features into focus.

Two reporting considerations follow from this. The certainty of the evidence is bounded by the corpus size, the predominance of preprints, and the scarcity of controlled studies, and some reporting bias is plausible, since novel or successful systems are more likely to be disseminated than unsuccessful classroom deployments. In addition, because the original platform-level exports were unavailable at revision, conventional identification and deduplication counts are not reported; the review therefore prioritises recoverable full-text verification, documented exclusions, and transparent study-level coding, which together support an auditable reading of the corpus.

6. Conclusion

On the evidence of this review, recent intelligent tutoring research is technically accomplished in grounding systems pedagogically and in making them adaptive, and underdeveloped in the shared infrastructure that lets results be compared and reproduced. The arrival of language-model tutors has widened what is possible in dialogue and subject coverage, but it has not yet been shown to lift learning outcomes beyond the moderate range established for earlier systems, and the most integrated work in the corpus is either an older system or new infrastructure built for comparison rather than one of the new conversational tutors.

One absence deserves to be named directly, because it cuts across all three axes. None of the twelve systems quantifies the uncertainty of its own decisions or calibrates its confidence against what learners actually do. Elsewhere in artificial intelligence, wherever decisions carry consequences, calibration metrics such as Expected Calibration Error and Brier scores are reported as a matter of course; in tutoring, where a confidently stated wrong answer can mislead a learner, they are absent. The methods are well established and the difficulty is one of attention rather than technique, which makes this the clearest single priority for the field.

The most productive direction, then, is to make the field more cumulative and more self-aware: to run controlled comparisons of language-model tutoring against tutoring grounded in instructional theory, to report calibration using the metrics already standard elsewhere, to develop shared benchmarks and logging profiles as a community, and to build systems that bring pedagogical theory, adaptive mechanisms, and standardised analytics together in one place. None of these alone will reach Bloom's two-sigma benchmark, but together they would let the field find out, for the first time, how close it can come. The central gap is therefore not simply that language-model tutors need better prompts, but that the next generation of intelligent tutoring systems must combine instructional theory, inspectable learner models, interoperable analytics, open evaluation protocols, and explicit uncertainty calibration within the same system.

AI Statement: The authors used an AI-assisted tool during two stages of this review. First, Claude was used to support the application of the SPAR-4-SLR protocol, specifically in organising and analysing candidate studies against the defined inclusion and exclusion criteria and in coding primary studies across the analytical axes. All screening decisions and coding judgements were verified and finalised by the authors. Second, Claude was used to assist with linguistic editing and grammatical enhancement of the manuscript. The intellectual content, research design, data interpretation, and conclusions are entirely the authors' own.

Ethics Statement: This systematic review analysed published or publicly available scholarly documents and did not involve human participants, intervention, or access to private learner records. Formal ethics approval was therefore not required.

Protocol and registration: No external review protocol was registered. The review followed the SPAR-4-SLR procedure, and the screening, verification, extraction, and coding decisions were logged in an internal audit trail. The audit trail, search strategy, coding sheet, quality assessment, and source-verification exclusion table are provided as supplementary materials (Appendices A to D).

Data availability: The data supporting this review consist of the search strategy, the study-selection audit trail, the source-verification exclusion table, the study extraction and coding sheet, and the quality assessment scores, all provided as supplementary files (Appendices A to D). No new learner-level dataset was created or analysed.

Competing interests: The authors declare no competing interests.

Author contributions: The authors jointly designed the review, contributed to manuscript revision, and approved the submitted version. The search, screening, source verification, extraction, coding, figures, and supplementary materials were prepared and checked through the documented audit trail.

References

- Baker, R.S. and Siemens, G., 2014. Educational data mining and learning analytics. In: R.K. Sawyer, ed. *The Cambridge Handbook of the Learning Sciences*. 2nd ed. Cambridge: Cambridge University Press, pp. 253-272.
- Bloom, B.S., 1984. The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), pp. 4-16.
- Borchers, C. and Shou, T., 2025. *Can large language models match tutoring system adaptivity? A benchmarking study*. arXiv:2504.05570.
- Degen, P.-B. and Asanov, I., 2025. *Beyond automation: Socratic AI, epistemic agency, and the implications of the emergence of orchestrated multi-agent learning architectures*. Preprint, University of Kassel.
- Dolata, M., Katsiuba, D., Wellnhammer, N. and Schwabe, G., 2023. Learning with digital agents: an analysis based on the activity theory. *Journal of Management Information Systems*, 40(1), pp. 56-95.
- Graesser, A.C., Hu, X., Nye, B.D., VanLehn, K., et al., 2018. ElectronixTutor: an intelligent tutoring system with multiple learning resources for electronics. *International Journal of STEM Education*, 5, 15.
- Holstein, K., Alevan, V. and Rummel, N., 2020. A conceptual framework for human-AI hybrid adaptivity in education. In: *Artificial Intelligence in Education*. Lecture Notes in Computer Science. Cham: Springer, pp. 240-254.
- Kakarla, S., Thomas, D.R., Lin, J., Gupta, S. and Koedinger, K.R., 2024. *Using large language models to assess tutors' performance in reacting to students making math errors*. arXiv:2401.03238.
- Knievel, C., Bernhardt, A. and Bernhardt, C., 2025. *AITEE - agentic tutor for electrical engineering*. arXiv:2505.21582.
- Kulik, J.A. and Fletcher, J.D., 2016. Effectiveness of intelligent tutoring systems: a meta-analytic review. *Review of Educational Research*, 86(1), pp. 42-78.
- Kwon, S., Kim, S., Park, M., Lee, S. and Kim, K., 2024. *BIPED: pedagogically informed tutoring system for ESL education*. arXiv:2406.03486.
- Li, W., Pea, R., Haber, N. and Subramonyam, H., 2025. *CogGen: a learner-centered generative AI architecture for intelligent tutoring with programming videos*. arXiv:2506.20600.
- Liu, Z., Yin, S.X., Lee, C. and Chen, N.F., 2024. *Scaffolding language learning via multi-modal tutoring systems with pedagogical instructions*. arXiv:2404.03429.

- Liu, Z., Lin, G., Tan, H.L., Zhang, H., et al., 2025. *SingaKids: a multilingual multimodal dialogic tutor for language learning*. arXiv:2506.02412.
- Luckin, R., Holmes, W., Griffiths, M. and Forcier, L.B., 2016. *Intelligence Unleashed: An Argument for AI in Education*. London: Pearson.
- Ma, W., Adesope, O.O., Nesbit, J.C. and Liu, Q., 2014. Intelligent tutoring systems and learning outcomes: a meta-analysis. *Journal of Educational Psychology*, 106(4), pp. 901-918.
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., et al., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372, n71.
- Paul, J., Lim, W.M., O'Cass, A., Hao, A.W. and Bresciani, S., 2021. Scientific procedures and rationales for systematic literature reviews (SPAR-4-SLR). *International Journal of Consumer Studies*, 45(4), pp. 1-16.
- Puech, R., Macina, J., Chatain, J., Sachan, M. and Kapur, M., 2024. *Towards the pedagogical steering of large language models for tutoring: a case study with modeling productive failure*. arXiv:2410.03781.
- Roll, I. and Wylie, R., 2016. Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 26, pp. 582-599.
- Shute, V.J. and Ventura, M., 2013. *Stealth Assessment: Measuring and Supporting Learning in Video Games*. Cambridge, MA: MIT Press.
- Smith, G., Gupta, A. and MacLellan, C., 2024. *Apprentice tutor builder: a platform for users to create and personalize intelligent tutors*. arXiv:2404.07883.
- Steenbergen-Hu, S. and Cooper, H., 2014. A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology*, 106(2), pp. 331-347.
- VanLehn, K., 2011. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), pp. 197-221.
- Weitekamp, D., Siddiqui, M.N. and MacLellan, C.J., 2025. *TutorGym: a testbed for evaluating AI agents as tutors and students*. arXiv:2505.01563.
- Zawacki-Richter, O., Marín, V.I., Bond, M. and 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, 39.

Appendix A. Supplementary Table S1: Study-by-study Coding Sheet

Study	Pedagogical alignment	Adaptivity	Analytics
BIPED (Kwon et al., 2024)	Yes	Yes	No
CogGen (Li et al., 2025)	Yes	Yes	No
AITEE (Knievel, Bernhardt and Bernhardt, 2025)	Yes	Yes	No
Socratic AI (Degen and Asanov, 2025)	Yes	Yes	No
Scaffolding tutor (Liu et al., 2024)	Yes	Yes	No
SingaKids (Liu et al., 2025)	Yes	Yes	No
Productive Failure tutor (Puech et al., 2024)	Yes	Yes	No
Apprentice Tutor Builder (Smith, Gupta and MacLellan, 2024)	No	Yes	No
Tutor assessment (Kakarla et al., 2024)	Yes	No	No
TutorGym (Weitekamp, Siddiqui and MacLellan, 2025)	Yes	Yes	Yes
ElectronixTutor (Graesser et al., 2018)	Yes	Yes	Yes
Adaptivity benchmark (Borchers and Shou, 2025)	Yes	Yes	Yes

Appendix B. Supplementary Table S2: Source-verification Exclusions

Exclusion category	n	Reason	Treatment
Mismatched / wrong document	5	Supplied source document did not match the cited bibliographic record	Excluded from primary synthesis
Not human tutoring	1	Concerned artificial-agent (multi-agent RL) learning, not human tutoring	Excluded from primary synthesis

Exclusion category	n	Reason	Treatment
Posted after search cut-off	1	Record fell outside the review search window	Excluded from primary synthesis
Total	7		Reported in the flow diagram

Appendix C. Supplementary Table S3: Search Strategy

Master query (five concept groups joined by AND; synonyms within each group joined by OR), adapted to each database's field syntax:

("intelligent tutoring system*" OR "AI tutor*" OR "pedagogical agent*" OR "educational chatbot*" OR "LLM tutor*" OR "dialogue-based tutor*")

AND (pedagog* OR "instructional design" OR "learning theory" OR scaffolding OR "mastery learning" OR "cognitive apprenticeship" OR "productive failure" OR Socratic)

AND (adaptiv* OR "learner model*" OR "student model*" OR "knowledge tracing" OR BKT OR IRT OR "retrieval-augmented" OR RAG OR personali*)

AND ("learning analytics" OR analytics OR xAPI OR SCORM OR "learning record store" OR LRS OR reproducib* OR benchmark* OR "open source")

AND (education OR learning OR student* OR "higher education" OR "K-12")

Scopus: TITLE-ABS-KEY(master query)

Web of Science Core Collection: TS=(master query)

IEEE Xplore: each group wrapped in "All Metadata": (...), joined with AND

ACM Digital Library: each group as [[All: term] OR ...], joined with AND

SpringerLink: master query in title/abstract/full-text advanced search

arXiv: all: terms per group, joined with AND, restricted to cat:cs.AI OR cs.CL OR cs.CY OR cs.LG OR cs.HC

Appendix D. Supplementary Table S4: Quality Assessment

Study	Total /10	Note
BIPED (Kwon et al., 2024)	7	Dialogue-act grounding; no learning-outcome trial
CogGen (Li et al., 2025)	7	Cognitive Apprenticeship + BKT; technical evaluation
AITEE (Knievel, Bernhardt & Bernhardt, 2025)	7	Socratic + RAG; knowledge-application benchmark
Socratic AI (Degen & Asanov, 2025)	7	Controlled comparison; preliminary
Scaffolding tutor (Liu et al., 2024)	7	Empirical case study
SingaKids (Liu et al., 2025)	7	Child-user evaluation
Productive Failure tutor (Puech et al., 2024)	7	Field test; not a controlled outcome effect
Apprentice Tutor Builder (Smith, Gupta & MacLellan, 2024)	6	Authoring platform; usability study (n=14)
Tutor assessment (Kakarla et al., 2024)	6	Assesses tutor responses, not tutoring learners
TutorGym (Weitekamp, Siddiqui & MacLellan, 2025)	8	Shared open testbed
ElectronixTutor (Graesser et al., 2018)	8	Established ITS; LRS logging
Adaptivity benchmark (Borchers & Shou, 2025)	8	Open, reproducible benchmarking method

Modeling AI Tool Adoption in Higher Education: The Role of Authentic Learning and Prompting Competence

Sultan Hammad Alshammari and Mohammed Habib Alshammari

Department of Educational Technology, College of Education, University of Ha'il, Saudi Arabia

sh.alshammari@uoh.edu.sa

m.habib@uoh.edu.sa

<https://doi.org/10.34190/ejel.24.4.4784>

An open access article under [CC Attribution 4.0](#)

Abstract: As AI tools are increasingly used in higher education, understanding the factors affecting students' adoption intentions has become theoretically and practically important. Previous studies have mainly relied on traditional technology acceptance constructs, while comparatively little attention has been given to pedagogical and competence-based conditions shaping students' cognitive evaluations of AI systems. To address this gap, the current study builds on the technology acceptance model (TAM) by proposing authentic learning (AL) and prompt engineering competence (PEC) as precursors of perceived usefulness (PU), perceived ease of use (PEU), and behavioral intention (BI) to use AI tools. The study was based on data collected from 309 undergraduate students at the University of Ha'il. A two-step structural equation modeling (SEM) approach was employed using AMOS software. Confirmatory factor analysis confirmed construct reliability, convergent validity, and discriminant validity. SEM was then conducted to test the hypotheses. The findings show that AL significantly predicts both PU and PEU, whereas PEC significantly predicts PEU but not PU. Both PU and PEU were found to be important predictors of BI. Bootstrapping results reveal that AL affects BI through PU and PEU, while PEC affects BI entirely through PEU. The results also confirm considerable explanatory power, with an R^2 of .71 for BI. These findings extend TAM by reconceptualizing AL as a foundational pedagogical precursor influencing AI adoption and by clarifying the unique role of PEC in improving PEU. Integrating pedagogical and competence-based determinants into AI-enabled higher education advances technology acceptance theory and explains the AI adoption mechanism more precisely. The findings provide practical guidance for educators and instructional designers by emphasizing the importance of integrating authentic learning tasks and developing students' prompt engineering skills to enhance meaningful AI-supported learning.

Keywords: Artificial intelligence in higher education, Technology acceptance model, Authentic learning, Prompt engineering competence, Behavioral intention

1. Introduction

The swift emergence of AI tools, especially generative AI, has already changed higher education (Kurtz et al., 2024). There is a growing trend of universities utilizing AI for teaching, assessment, administration, and research (Flores-Velásquez et al., 2024). AI-powered systems are being developed to support personalized learning pathways, adaptive feedback, and data-driven decision-making, offering unprecedented opportunities for engagement and improved learning outcomes (Vafadar and Amani, 2024). This potential is enhanced through generative AI allowing students and teachers to create text, code, and visual outputs to enhance creativity, productivity, and problem-solving abilities (Hughes et al., 2025; Özbay, Özbay and Tutkunca, 2025; Alshammari et al., 2026).

The implementation of AI presents numerous opportunities but also poses several teaching and ethical issues. Issues of academic integrity, bias, privacy, and intellectual property have already sparked considerable debate within the academic community (Vafadar and Amani, 2024; Hughes et al., 2025). The rapid creation of complex content by AI poses a threat to the traditional approach to assessment. The use of AI may lead to cognitive dependency and the weakening of critical thinking skills among students (Hughes et al., 2025). Furthermore, not all students have easy access to sophisticated AI programs, thereby creating a digital divide (Özbay, Özbay and Tutkunca, 2025). Finding solutions to these issues will require more than the implementation of regulations governing the use of AI in education. It will also be necessary for educators to gain a deeper understanding of how students intend to use AI (Rughiniş et al., 2025).

It is important, theoretically as well as practically, to discover students' intentions for adopting AI tools, owing to the potential two-edged sword AI technologies represent (Alenezi, 2023; Alshammari and Alshammari, 2024; Alshammari et al., 2025).

Understanding how students judge AI systems and under what conditions constructive and responsible use of AI occurs is key to effective integration. In recent decades, the technology acceptance model (TAM) has provided

a useful lens for understanding technology adoption (Flores-Velásquez et al., 2024). The TAM states that the main determinants of behavioral intention are perceived usefulness (PU), or how much a system enhances performance, and perceived ease of use (PEU), or how effortless a system is perceived to be (Shahzad, Xu and Javed, 2024; Shata and Hartley, 2025). These constructs have shown a strong capacity to predict technology adoption in educational contexts (Alshammari and Alshammari, 2024; Alshammari et al., 2025).

However, the TAM has conceptual limitations in the generative AI environment. The TAM views adoption as perception driven and implicitly assumes that cognitive evaluations occur primarily through individual assessment of system attributes. Interactive intensity, skill conditioning, and educational embedding also characterize AI-enabled learning environments (Zhao, An and Liu, 2025; Roe et al., 2025). Using generative AI involves constantly prompting and evaluating outputs to apply a certain judgment. Under such conditions, perceptions of usefulness and ease may not result simply from internal system characteristics but from instructional design and user effectiveness. The above suggests that the explanatory power of the TAM in AI contexts is dependent on the inclusion of context-sensitive antecedents that serve to actively construct cognitive perceptions rather than merely predicting intention from them.

This research paper argues that students' perceptions of AI tools are co-constructed through pedagogical framing and interactional competence. Specifically, we propose that authentic learning (AL) and prompt engineering competence (PEC) are proposed as predictors of PU and PEU.

Authentic learning is defined as learning that is designed to occur in contexts that approximate as closely as possible the real world (Larng, 2024). It is possible to use AI tools right away as they can be embedded in authentic tasks. As students engage AI to solve real problems and not just theoretical exercises, its usefulness is seen to be relevant (Salinas-Navarro et al., 2024). These alignments should greatly boost PU. Similarly, an authentic learning environment that scaffolds participation aids in reducing ambiguity and clarifying the task objective. A structure of this kind can reduce complexity perception at the time of interaction between humans and AI, enhancing PEU.

To interact with generative AI, PEC is the social interactional competence to utilize. Effective prompt design is crucial in obtaining relevant and high-quality output from AI (Woo et al., 2024). When prompt tactics are used effectively, there is less ambiguity and more predictability, leading to less frustration for readers (Bashardoust et al., 2024). Therefore, students with stronger PEC are likely to find AI tools easier to utilize. While competence may lead to better results, the main theoretical goal of competence is to reduce friction rather than making things seem genuinely more useful.

Previous studies have investigated AI adoption, pedagogical innovation, and skills related to AI. However, a structural framework that explicitly models pedagogical and competence-based determinants as antecedents of TAM constructs has rarely been the subject of integrated study. Existing research treats perceptions of usefulness and ease as unique cognitive evaluations that are independent from instructional design or skill. A gap therefore exists in understanding how pedagogical structure and AI-specific competence shape perceptions that will influence behavioral intention. As such, this study offers a proposal for, and empirical test of, an extended technology acceptance model that incorporates AL and PEC as predictors of PU and PEU that predict behavioral intention (BI) to use AI tools for learning. Employing structural equation modeling (SEM), the study aims to offer a theoretically enhanced explanation of AI tool adoption in higher education as it demonstrates how pedagogical and competence-based conditions shape students' cognitive evaluations and subsequent adoption intentions.

Accordingly, this study addresses the following research question:

RQ: How do authentic learning and prompt engineering competence influence students' behavioral intention to use AI tools through the mediating mechanisms of perceived usefulness and perceived ease of use within the Technology Acceptance Model framework?

2. Literature Review

2.1 AI Adoption in Higher Education

The quick speed at which generative AI technology is exploding is causing AI to enter more furiously into the higher education sector thereby impacting teaching and learning practices (Adamakis and Rachiotis, 2025; Kurtz et al., 2024). Recent empirical studies increasingly record experimentation with, and adoption of, AI tools among university students, who are embedding these systems in content generation, feedback, data analysis, and problem-solving tasks (Alshammari et al., 2025; Mukhamedkarimova and Umurkulova, 2025).

Two main themes are identified in the literature. First, AI allows for personalization and efficiency. Tools that rely on generative AI, such as adaptive tutoring systems, automated grading systems, and platforms that permit students to innovate and draft in addition to providing iterative feedback in real time, have been found to enhance productivity and perceived academic support (Vieriu and Petrea, 2025; Sallai et al., 2024). However, there is some ethical and pedagogical tension. Concerns about academic integrity, algorithmic bias, overreliance, and digital inequities continue to shape institutional debate (Hughes et al., 2025; Roe et al., 2025).

The use of these systems is determined not by students' exposure to them, but by students' own perceptions. While there are high rates of experimentation with these AI systems, their eventual incorporation into students' educational processes will likely depend on their perceptions of the systems' value (Wahdah, Kurniawan and Irawan, 2025; Flores-Velásquez et al., 2024). Theoretical models are therefore needed to explain students' perceptions of value and usability.

2.2 Technology Acceptance Model in AI Contexts

TAM continues to be utilized to examine technology use in educational environments (Shata and Hartley, 2025). According to TAM, BI is explained by PU, defined as the belief that the technology enhances students' performance, and PEU, defined as the belief that the technology requires minimal effort to use (Flores-Velásquez et al., 2024). The practical use of the TAM in AI contexts supports the theory that PU and PEU predict the acceptance intention of students (Zhao, An and Liu, 2025; Kanont et al., 2024). When students see real academic value and experience smooth and intuitive interactions, they are more likely to include AI into their workflow.

However, current generative AI environments significantly limit the applicability of traditional TAMs. According to this model, perceptions are largely determined by a person's valuation of system characteristics. AI-enabled learning environments depend on skills, more interaction, and enhanced pedagogy (Roe et al., 2025). Generative AI requires frequent prompting, critical assessment, and contextual coordination with learning tasks. In such a context, perceptions of usefulness and ease of use may be affected by instructional design and user competence, not only the features of the system. This indicates that the TAM needs to be enriched with context-sensitive antecedents that are able to account for how perceptions themselves are constructed.

2.3 Authentic Learning as a Pedagogical Predictor

Authentic learning is based on constructivist principles and involves a real, meaningful task that the learner must investigate in collaboration with others, and which requires the learner to apply their knowledge (Lang, 2024; Salinas-Navarro et al., 2024). Rather than focusing on abstract exercises, authentic learning environments ground academic work in realistic contexts, which leads to more engaged learning and higher-order thinking (Kitchen et al., 2025).

Within such scenarios, AI tools are functional. When AI is used for real projects, such as policy analysis or data-driven reports or essays, its usefulness becomes clear. According to Salinas-Navarro et al. (2024), it is likely that the extent to which AI can help students achieve things they value will strengthen PU. Authentic learning may also improve PEU. Specifying prompts, using structured steps, and guided application help reduce ambiguity. When the use of AI is embedded within clearly defined goals, students experience cognitive friction and clarity regarding the appropriate deployment of the technology (Fragkaki et al., 2025; Radović et al., 2021).

Based on the above arguments, the following hypotheses are proposed:

H1: Authentic learning positively influences perceived usefulness.

H2: Authentic learning positively influences perceived ease of use.

2.4 Prompt Engineering Competence as an AI-Specific Skill

Prompt engineering competence is the ability to make structured, accurate, and iterative prompts in a way that generates accurate and relevant outputs from generative AI systems (Woo et al., 2024; Federiakin et al., 2024). Prompt engineering competence refers to human-centric and operating-focused AI, in comparison to general AI literacy, which is conceptual and ethical (Lee and Palmer, 2025).

Efficiently creating a prompt will eliminate uncertainty during AI interaction. According to Woo et al. (2024), skilled users can harness this ability to tailor their commands strategically, increasing the predictability of responses and reducing trial and error; ultimately reducing perceived effort. Consequently, PEC is theoretically positioned as an antecedent of PEU, suggesting a more complex relationship with PU. While effective prompting may yield higher-quality outputs, proficiency alone does not alter an AI system's functional capabilities. Instead,

it simply provides access to these advanced capabilities (Anam, 2025). Therefore, PEC indirectly influence perceptions of usefulness, but its primary theoretical impact lies in enhancing interactional fluency.

Based on the above arguments, the following hypotheses are proposed:

H3: *Prompt engineering competence positively influences perceived ease of use.*

H4: *Prompt engineering competence positively influences perceived usefulness.*

3. Perceived Usefulness, Perceived Ease of Use, and Behavioral Intention

According to the TAM, global task–technology fit (TTF), and information and communication technology (ICT) in traditional schools, recent experiments conducted by Zhao, An and Liu (2025) have demonstrated that PU is the strongest predictor of BI in the field of education. Students’ intention to adopt AI tools increases when learning efficiency and academic performance are enhanced, while PEU also influences BI in situations requiring repeated contact (Bayaga, 2025). Interaction that is easier on cognition reduces load and encourages longer usage.

The TAM suggests that PEU could strengthen PU, as both constructs could serve as proximal determinants of intention. In the context of AI, the quality of interaction and the perceived value are closely linked, and these cognitive evaluations jointly influence the adoption of a system.

Based on the above arguments, the following hypotheses are proposed:

H5: *Perceived usefulness positively influences behavioral intention.*

H6: *Perceived ease of use positively influences behavioral intention.*

While the TAM can help establish a foundation for the adoption of new technologies, it is essential to extend the model to AI contexts. The abilities to engage in authentic learning and prompt engineering are expected to serve as key precursors influencing students’ PU and PEU. Although TAM suggests a relationship between ease of use and usefulness, these constructs are treated as separate mediators in the present model. The integration of these factors supports the development of a contextually grounded model for AI adoption in higher education. The validation of this model is examined empirically in this study. The proposed research model is presented in Figure 1.

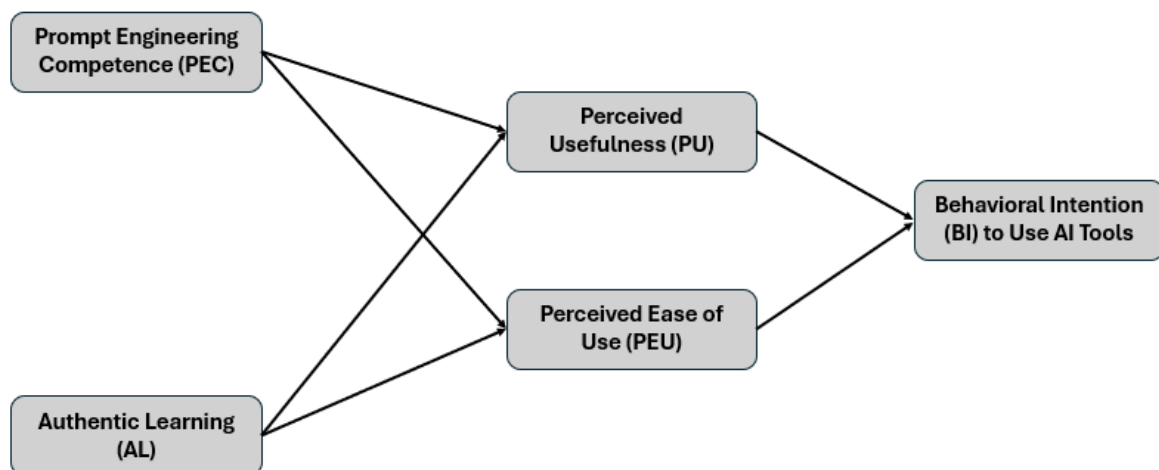


Figure 1: Proposed Research Model

4. Methodology

4.1 Research Design

The study utilized a quantitative cross-sectional research design to investigate the factors influencing students’ intention to use AI tools in higher education. To test the posited structural relationships among AL, PEC, PU, PEU, and BI, a survey-based approach was chosen. Many studies in the technology acceptance literature use cross-sectional survey designs, as they are useful to model latent constructs and predictive relationships (Hair et al., 2021; Kline, 2023).

The evaluation of the research model was performed using an AMOS covariance-based SEM test. According to Byrne (2013) and Hair et al. (2021), SEM is particularly useful for testing theoretically specified structural relationships among latent constructs and evaluating the extent to which empirical data support proposed explanatory models.

4.2 Participants and Data Collection

The collection of data took place in the first semester of 2025–2026. Respondents for this study were recruited from among students of various colleges of the University of Ha'il taking the course Introduction to AI. Google Forms was utilized to devise an online questionnaire. The link to the survey was shared with students via Blackboard and the university email, authorized by the institution. Potential participants were briefed about the study objectives, anonymity, and confidentiality. A total of 309 valid responses were collected and analyzed. The sample size is greater than the minimum recommendations for analysis using SEM, which normally requires a minimum of 200 cases or 10 participants per estimated parameter (Hair et al., 2006; Kline, 2023). This sampling frame was intentionally selected because students enrolled in the "Introduction to AI" course have structured and direct exposure to AI tools within an educational context. Such exposure ensures that participants possess sufficient familiarity and interaction experience to meaningfully evaluate perceived usefulness, perceived ease of use, and competence-related factors. Therefore, this context provides a suitable and theoretically relevant setting for examining AI adoption mechanisms in higher education.

However, it is acknowledged that the use of a single institution and course may limit the generalizability of the findings. This limitation is addressed in the limitations and future research section, and future research is encouraged to replicate the study across diverse institutional and disciplinary contexts.

4.3 Instrumentation

The survey tool consisted of five sections based on validated scales. All items were measured using a five-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The nine items used to evaluate PEC were adapted from the research of Gibreel and Arpacı (2025). The items examine students' ability to ask precise questions, improve AI instructions step by step, and communicate with generative AI.

Authentic learning was assessed using nine items adapted from Neo et al. (2012). The scale compares students' perceptions of learning in real life to an active and collaborative inquiry approach. The four items testing PU were adapted from Davis (1989) and measure the extent to which students agree that using AI tools could boost their academic performance and learning ability. The four items used to measure PEU were also adapted from Davis (1989). This scale measures the extent to which students believe that AI tools are easy to learn and use. Behavioral intention was measured using three items (Davis, 1989) that reflect students' intention to continue using AI tools for learning purposes. Prior to distribution, all items were reviewed for clarity and relevance.

4.4 Data Analysis

The data were analyzed in two stages. SPSS was used to help with the demographic and data analysis, examining missing values, normality, and outliers in our study. Statistics were calculated to describe participants' characteristics. All statistical analyses were conducted using IBM SPSS Statistics version 26 and AMOS version 24. Missing data were assessed prior to analysis. The dataset contained no missing values; therefore, no data imputation or deletion procedures were required. The two-step SEM method suggested by Anderson and Gerbing (1988) was used. Initially, confirmatory factor analysis (CFA) was performed to evaluate the measurement model in terms of construct reliability, convergent validity, and discriminant validity. Next, the relationships between constructs were examined using the structural model. Path coefficients, critical ratios, and significance levels were evaluated. Bootstrapping methods were employed for indirect effect and mediation path evaluation using 5,000 resamples with bias-corrected confidence intervals, which provide robust estimates without assuming normality of the sampling distribution (Preacher and Hayes, 2008). By separating measurement validation from structural testing, the two-step SEM procedure improves rigor and reduces bias, thereby enhancing the interpretability of SEM models (Anderson and Gerbing, 1988; Byrne, 2013).

Outliers were examined using boxplots for univariate outliers and Mahalanobis distance for detecting potential multivariate outliers. No significant outliers were identified that required removal.

5. Results

5.1 Demographic Information

Table 1 presents the demographic characteristics of the 309 respondents in this study. Of this number, the majority were male respondents, namely 66.7% (n = 206), with 33.3% (n = 103) being female. Despite the fact that gender distribution is skewed in favor of male respondents, both groups were adequately represented for meaningful structural modeling.

In terms of their academic programs, most respondents were enrolled in a bachelor's degree program (90.9%; n = 281), while the rest were diploma students (9.1%; n = 28). The outcomes mainly reflect those of individuals pursuing undergraduate degrees, which seems justifiable given the effect of increasing AI tool use at bachelor's level.

The study included a variety of participants from different colleges to increase the representation of various disciplines. The highest percentage of respondents came from the health college (25.2%; n = 78), business administration (19.4%; n = 60), science (13.9%; n = 43), arts (11.7%; n = 36), computer science and engineering (9.1%; n = 28), education (8.1%; n = 25), applied college (7.4%; n = 23), and engineering (5.2%; n = 16). The presence of various disciplines strengthens the applicability of the findings across different academic fields, especially as uses of AI increasingly straddle technical and non-technical disciplines.

On the question of previous exposure to AI tools, 94.8% of respondents (n = 293) reported previous use of AI, whereas 5.2% of respondents (n = 16) reported no previous use of AI. The high rate of exposure indicates that students are already using AI. Therefore, it is both timely and relevant to investigate the determinants of AI adoption.

Regarding the duration of utilization of AI tools, approximately half the respondents (47.2%; n = 146) had utilized AI tools for more than six months. Moreover, 20.1% (n = 62) had utilized AI tools for four to six months, while 17.8% (n = 55) and 14.9% (n = 46) had utilized AI for one to three months and for less than one month, respectively. The data shows that a large percentage of respondents had had ongoing experience with the tools, indicating that their answers were more likely based on experience and not on a first-time test.

The demographic indicators of the analyzed sample were heterogeneous in terms of academic discipline, level of AI experience, and gender. Due to the high levels of previous use of AI and extent of use, it would be appropriate to study the community's cognitive and pedagogical reasons for utilizing AI.

Table 1: Demographic and Background Characteristics of the Participants

		Frequency	Percent
Gender	Male	206	66.7
	Female	103	33.3
Academic Program	Bachelor	281	90.9
	Diploma	28	9.1
Colleges	Education	25	8.1
	Science	43	13.9
	Arts	36	11.7
	Business Administration	60	19.4
	Computer Science and Engineering	28	9.1
	Engineering	16	5.2
	Applied College	23	7.4
	Health	78	25.2
Previous_experience_with_AI	Yes	293	94.8
	No	16	5.2

		Frequency	Percent
Period_with_Using_AI_tools	Less than a month	46	14.9
	1-3 Months	55	17.8
	4-6 Months	62	20.1
	More than six months	146	47.2
Total		309	100.0

5.2 Confirmatory Factor Analysis

Following the approach advocated by Anderson and Gerbing (1988), a confirmatory factor analysis (CFA) was conducted to assess the measurement model. The initial measurement model did not achieve acceptable fit according to commonly recommended SEM criteria, yielding $\chi^2/df = 3.046$, CFI = .877, IFI = .878, and RMSEA = .082 (Awang, 2015). Figure 2 presents the initial measurement model before refinement.

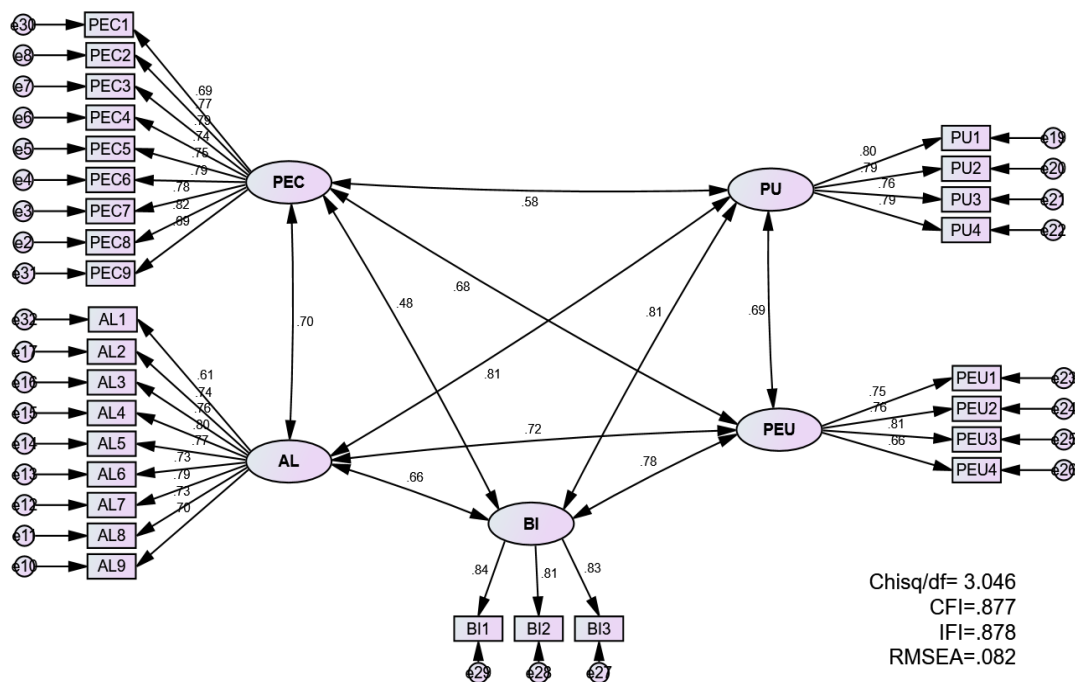


Figure 2: Initial Measurement Model Before Refinement

However, inspection of the factor loadings and modification indices revealed that three items (AL1, PEC1, and PEC9) exhibited relatively low factor loadings and contributed to model misfit. In accordance with SEM guidelines, indicators that do not adequately represent their underlying construct may be removed to improve measurement quality (Hair et al., 2010). Accordingly, these three items were removed from the model.

Furthermore, modification indices suggested correlating residual terms within the same constructs. Specifically, e3 and e2 corresponded to items within the Authentic Learning construct, while e12 and e15 corresponded to items within the Prompt Engineering Competence construct. These covariances were theoretically justified due to similarities in item wording and conceptual overlap. As recommended by Awang (2015), correlated residuals were introduced only when supported by substantive theoretical justification rather than solely to improve statistical fit.

Model fit was evaluated using multiple goodness-of-fit indices, including the chi-square to degrees of freedom ratio (χ^2/df), Comparative Fit Index (CFI), Incremental Fit Index (IFI), and Root Mean Square Error of Approximation (RMSEA). Following established SEM guidelines, χ^2/df values below 3.0, CFI and IFI values of .90 or higher, and RMSEA values of .08 or lower indicate acceptable model fit (Hair et al., 2010; Awang, 2015).

Following these theoretically justified modifications, the revised measurement model achieved acceptable fit ($\chi^2/df = 2.906$, CFI = .901, IFI = .901, RMSEA = .079). Compared with the initial model, the revised model demonstrated improved fit across all major indices, indicating a more adequate representation of the observed data. The limited and theoretically justified nature of the modifications ensured model parsimony and avoided excessive post-hoc adjustments. Accordingly, the revised measurement model was deemed suitable for subsequent structural analysis. Figure 3 presents the revised measurement model after refinement.

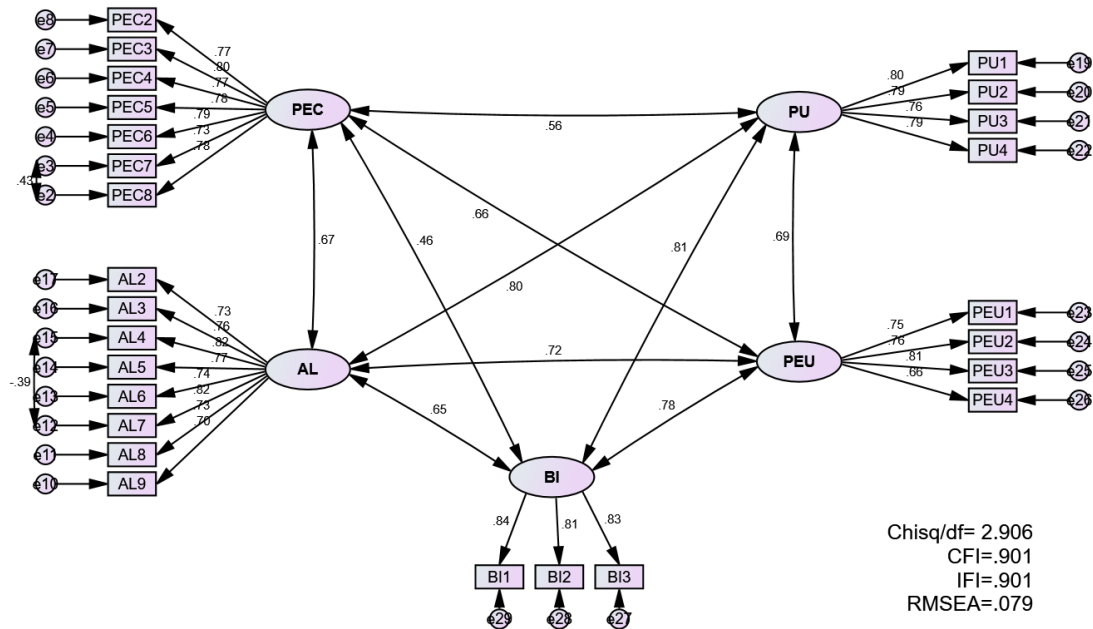


Figure 3: Revised Measurement Model After Refinement

To facilitate evaluation of the model refinement process and enhance methodological transparency, Table 2 compares the goodness-of-fit indices of the initial and revised measurement models.

Table 2: Comparison of Fit Indices for the Initial and Revised Measurement Models

Model	χ^2/df	CFI	IFI	RMSEA
Initial CFA Model	3.046	.877	.878	.082
Revised CFA Model	2.906	.901	.901	.079

As shown in Table 2, the revised model demonstrated improved fit across all major indices, supporting the adequacy of the measurement model for subsequent structural analysis.

Composite reliability (CR) and average variance extracted (AVE) were used to determine the convergent validity after the CFA verified the model fit. The SEM literature suggests that CR values should be higher than .60 to show adequate internal consistency. According to Hair et al. (2010), it is indicated by AVE values which should be more than .50 to show that the latent construct can explain half or more variance of that indicator. All constructs met all of these criteria (see Table 3). The observed CR values were between .834 and .915, all greater than .60. Likewise, the ranges of AVE values varied from .558 to .683, exceeding .50. The results show that the indicators sufficiently converge on their latent constructs. The overall CR and AVE statistics confirm that all the measurement model's constructs have satisfactory convergent validity (Awang, 2015; Hair et al., 2010).

Table 3: Composite Reliability and Convergent Validity of the Measurement Model

	CR	AVE
PEU	0.834	0.558
PEC	0.913	0.601
AL	0.915	0.575
PU	0.866	0.617
BI	0.866	0.683

The discriminant validity was assessed using the heterotrait–monotrait (HTMT) ratio of correlations, a more reliable and sensitive measure than the traditional method (Henseler, Ringle, and Sarstedt. 2015). According to Henseler, Ringle, and Sarstedt (2015), HTMT values less than .85 (conservative threshold) or .90 (liberal threshold) are considered ideal. Similarly, Hair et al. (2021) suggest that the HTMT value should not exceed .90 to ensure the constructs are conceptually distinct.

The HTMT values as shown in Table 4, range from .450 to .810. The highest value (PU and BI = .810) is below the conservative benchmark of .85 indicating that the constructs are empirically distinct. The inter-construct HTMT ratios did not approach, let alone exceed, the critical threshold of .90. Therefore, the model's latent constructs are sufficiently distinct from one another, as each one is expected to represent a particular concept (Henseler, Ringle, and Sarstedt 2015; Hair et al. 2021).

Table 4: Heterotrait–Monotrait Ratio (HTMT) for Discriminant Validity Assessment

Construct	PEC	AL	PU	PEU	BI
PEC	—	0.620	0.550	0.590	0.450
AL	0.620	—	0.780	0.720	0.670
PU	0.550	0.780	—	0.630	0.810
PEU	0.590	0.720	0.630	—	0.740
BI	0.450	0.670	0.810	0.740	—

5.3 Standardized Estimate

As suggested by Awang (2015), the strength of the factor loadings, structural paths, and coefficient of determination (R^2) were assessed by examining standardized estimates. Figure 4 illustrates that all the retained indicators, having standardized loadings over .60, add up to effective indicator reliability, indicating that the indicators adequately represent their constructs. The standardized structural coefficients show significant effect sizes, which allow for comparison of the weights of the predictors of the model. Significantly, this model was able to explain 71% of the variance in BI ($R^2 = .71$). According to Cohen (2013), R^2 estimates above .26 represent a large effect size for behavioral research. Thus, R^2 result demonstrates that the proposed model offers an adequate and practically significant representation of students' intention to adopt AI tools for higher education.

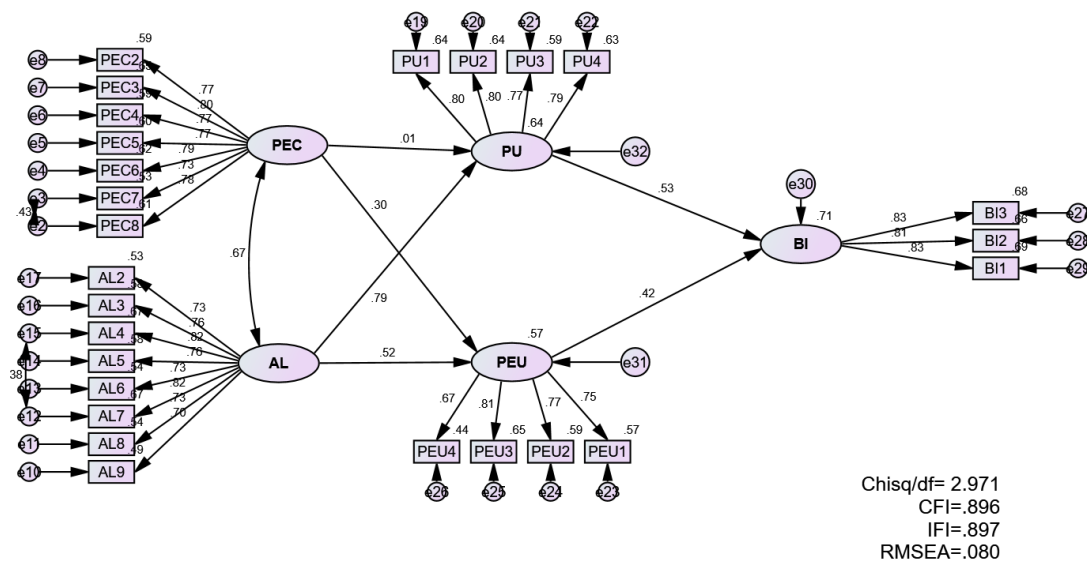


Figure 4: Final Structural Equation Model with Standardized Path Coefficients

5.4 Unstandardized Estimate

To evaluate the structural relationships and test the research hypotheses, the unstandardized estimates were examined. Unstandardized estimates, or actual regression weights (B), represent the relationship between variables in their true measurement units. To calculate the critical ratio (CR), the estimate is divided by its standard error. Awang (2015) notes that the unstandardized estimate is crucial for determining the statistical significance of relationships between variables. CR values are comparable to t-values; values greater than ± 1.96 are considered statistically significant. By analyzing the unstandardized estimates, their standard errors, CR values, and corresponding p-values, researchers can gain valuable insights into the statistical significance and validity of relationships in covariance-based SEM models. Standardized estimates reveal the strength of the relative effect, while unstandardized estimates help assess parameter significance for the acceptance or rejection of hypotheses. Figure 5 presents the unstandardized SEM estimates.

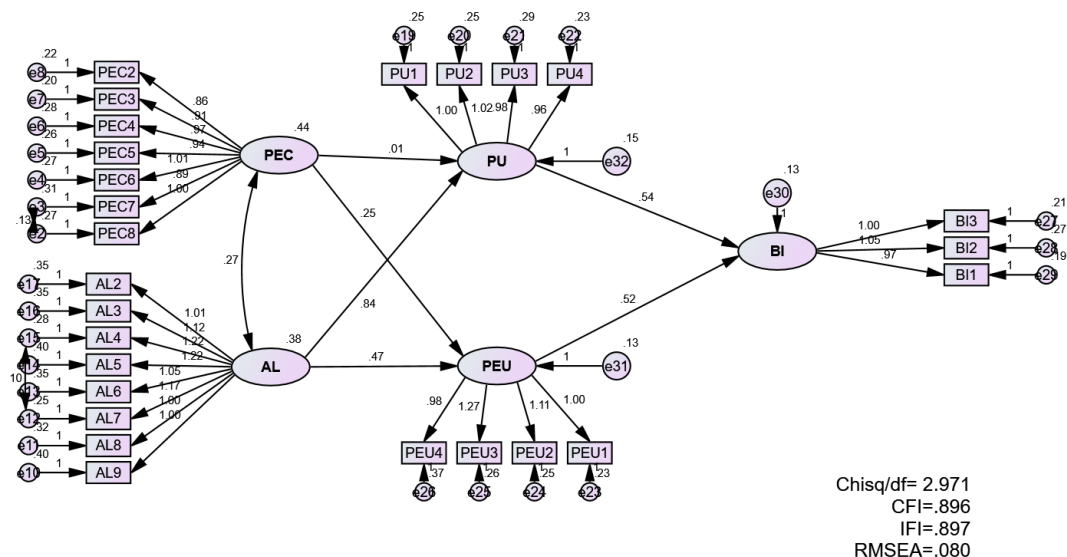


Figure 5: Unstandardized Structural Equation Model Estimates

6. Structural Model and Hypothesis Testing

The structural relationships were evaluated using unstandardized regression estimates with standard errors with CRs and p-values. According to Awang (2015), a value greater than ± 1.96 is significant at .05.

As indicated in Table 5, AL significantly predicted PU ($B = .844$; $CR = 9.360$; $p < .001$). Consequently, H1 is accepted. AL was positively associated with PEU ($B = .466$; $CR = 6.647$; $p < .001$), supporting H2. The current research establishes the importance of pedagogical design in affecting students' cognitive perception of AI tools.

The effect of PEC on PEU is positive and significant ($B = .247$; $CR = 4.180$; $p < .001$), supporting H3. However, PEC was not significantly associated with PU ($B = .013$; $CR = .206$; $p = .837$). Therefore, H4 is not supported. As the findings suggest, prompt engineering helps lower the level of effort to interact with AI tools, but prompt engineering skills do not enhance the PEU of the tool to achieve learning outcomes.

In accordance with the main propositions of the TAM, PU significantly predicted BI ($B = .544$; $CR = 8.144$, $p < .001$). Therefore, H5 is supported. In the same way, PEU was found to significantly predict BI ($B = .516$; $CR = 6.564$; $p < .001$), thus confirming H6.

Consequently, five of the six hypotheses were confirmed. According to the findings, AL plays a central role in determining the perceptions of usefulness and ease of use. However, PEC enhances only PEU. The cognitive perceptions (PU and PEU) appear as significant predictors of students' intention to use AI tools in higher education.

Table 5: Results of Structural Equation Modeling and Hypothesis Testing

Hypothesis	Path	B	S.E.	C.R.	P	Decision
H1	AL → PU	.844	.090	9.360	<.001	Supported
H2	AL → PEU	.466	.070	6.647	<.001	Supported
H3	PEC → PEU	.247	.059	4.180	<.001	Supported
H4	PEC → PU	.013	.063	.206	.837	Not Supported
H5	PU → BI	.544	.067	8.144	<.001	Supported
H6	PEU → BI	.516	.079	6.564	<.001	Supported

7. Indirect Effects Analysis

The indirect effects assessment of AL and PEC on BI was done using bootstrapping techniques. Both Awang (2015) and Hair et al. (2010) recommend using bootstrapping to analyze the indirect effects in SEM because bootstrapping does not require any assumption of normality of the sampling distribution. Additionally, bootstrapping provides bias-corrected (BC) confidence estimates for the indirect paths.

Table 6 shows that AL has a significant indirect influence on BI through PU and PEU. The corresponding values were found to be ($B = .700$; $p < .001$), confirming that the effect of AL on students' intention to use AI tools occurs through a cognitive evaluation mechanism. AI appears to be more useful and easier to use for students when embedded in real-world learning situations, resulting in a stronger intention to use the technologies.

Similarly, PEC has a significant indirect effect on BI through PEU ($B = .135$; $p = .044$). This result means that prompt engineering skills enhance students' intention to use AI tools. The path from PEC to PU was found to be non-significant, indicating that the mediation was only valid for PEU.

The bootstrapping results reflect that both the pedagogical determinant (AL) and the competence-based determinant (PEC) have an indirect effect on BI via cognitive perceptions. This shows that PU and PEU reinforce each other in the extended TAM framework.

Table 6: Bootstrapped Indirect Effects on Behavioral Intention (BI)

Predictor	Indirect Path(s)	Indirect Effect (B)	p-value	Decision
AL	AL → PU → BI; AL → PEU → BI	0.700	<.001	Supported
PEC	PEC → PEU → BI	0.135	.044	Supported

8. Discussion

The study extends the TAM by adding AL and PEC as upstream determinants of the adoption of AI tools in higher education. Using SEM in a staged approach on data from 309 undergraduate students, the results shed light on

how pedagogical design and AI-specific interactional skills affect cognitive assessments of AI tools, which, in turn, leads to use intention. The model explains 71% of the variance in BI. This indicates that the proposed extensions have a significant impact on the explanatory ability of TAM in AI-enabled learning.

The results show that AL is a significant predictor of PU and PEU. As a result, the design of instructions influence the beliefs students have about AI technologies. Students who engage with AI tools that are embedded in authentic, problem-oriented and contextually relevant tasks are more likely to use them as tools for achieving real academic and professional results (Kitchen et al., 2025; Lang, 2024; Salinas-Navarro et al., 2024). When learners enter a workplace or classroom-like context, AI is no longer experienced as a technological artefact but as a resource that contributes towards better performance. When technology delivers visible and consequential benefits, it is perceived as being useful.

At the same time, structured and scaffolded AL appears to lessen interactional levels of perceived difficulty. Fraggaki et al. (2025) and Radović et al. (2021) argue that interacting with AI is made easier through guided engagement, having a clear task objective, and working as a team. Students do not see AI as a standalone tool but as part of a goal-directed learning ecology. As a result, it requires less effort and is easy to use. Seemingly, usability is not only dependent on interface features but also on pedagogical embedding.

Another important finding is that PEC was significantly associated with PEU but not PU. The underlying rationale for this distinction lies in the nature of skill-based interaction with AI systems, where students' perceptions are shaped by their ability to effectively engage with prompt creation rather than by the perceived value of the outcomes. Students who can create effective prompts often experience smoother interactions with AI, leading to less friction and frustration with the application (Anam, 2025; Bashardoust et al., 2024; Woo et al., 2024). Thus, students are more likely to perceive AI tools as user-friendly and easier to use.

However, skilled prompting alone does not increase perceptions of usefulness. Perceived usefulness, rather than the user's ability to use AI, seems to be based on the belief that AI systems can lead to significant improvements in learning outcomes. In other words, leveraging AI operationally allows organizations to make use of it but does not necessarily make AI valuable. Even if a student is skilled at crafting prompts, the fact that AI outputs do not significantly advance learning tasks will not change usefulness perceptions (Lee and Palmer, 2025; Federiakin et al., 2024). The non-significant effect of prompt engineering competence on perceived usefulness can be interpreted through expectancy-value theory. While competence enhances students' ability to interact effectively with AI systems, perceived usefulness depends on the extent to which students believe that the technology contributes to meaningful learning outcomes. Thus, operational proficiency alone may not translate into higher perceived value unless it is aligned with task relevance and learning goals. This distinction highlights that competence reduces interactional effort but does not necessarily shape outcome-based evaluations. Our understanding of competence-based determinants is refined by this finding, which highlights that interactional skills primarily affect effort-related beliefs and not outcome-related ones.

As indicated by TAM, PU and PEU significantly predict BI (Chukwuere, 2024; Flores-Velázquez et al., 2024; Ali et al., 2025; Zhao, An and Liu, 2025). Overall, students are more likely to adopt AI tools in a learning context when they perceive them as useful and easy to use. Furthermore, the R^2 value of .71 indicates that the model explains a substantial proportion of the variance in AI adoption. Finally, the findings indicate that PU and PEU demonstrate strong explanatory power within the model, alongside the additional antecedent variables introduced in this study.

9. Theoretical Implications

Unlike prior TAM extensions that primarily introduce additional perceptual or attitudinal variables, this study advances the model by integrating pedagogical and competence-based antecedents as upstream antecedents of cognitive evaluations. This reconceptualization positions perceived usefulness and perceived ease of use as outcomes of learning design and interactional capability, rather than solely as evaluations of system characteristics. In doing so, the study contributes to a more context-sensitive understanding of AI adoption in educational environments. This study improves the TAM in three ways. First of all, cognitive beliefs that are core to the TAM are pedagogically constructed and not merely individually, influencer, or technology based. Conventional use of the TAM assumes that PU and PEU emanate from user-technology interaction in isolation. According to the present findings, learning contexts shape authenticity perceptions. By linking pedagogical design to the TAM's core elements and locating technology acceptance to the educational context's broader learning ecology, this study improves the TAM's contextual validity, especially in educational contexts.

Prompt engineering competence exerts a distinct influence that enhances the understanding of competence-based determinants. The reduction of user effort appears to drive the value of AI-specific operational skills. The distinction between operational and functional skills provides a more nuanced understanding of how competence in specific skills influences the decision to adopt AI applications. However, this finding suggests that not all forms of AI literacy equally enhance perceptions of usefulness and ease of use.

Furthermore, by integrating both pedagogical and competence-based influences on AI adoption, the TAM framework can better account for the role of educational context in shaping adoption decisions, particularly for technologies that require advanced skills, such as generative AI in higher education. Cognitive evaluations related to AI adoption are not formed in isolation; rather, they are shaped by the learner's educational environment and their competence in relevant skills.

10. Practical Implications

The findings offer valuable insights for higher education stakeholders, especially for curriculum designers in terms of integrating AI with real-world authentic tasks. According to the study results, the use of AI was relevant and useful for respondents when using project-based and solution-based activities. This strengthens the PU of users. Classroom use of AI is therefore a more meaningful application than incidental exposure to AI (Kitchen et al., 2025; Salinas-Navarro et al., 2024). Instructors should deliberately scaffold how learners interact with AI. Mzwri and Turcsányi-Szabó (2025) highlight that through structured opportunities to practice creating prompts, revising outputs and evaluating AI responses can reduce perceived difficulty, ultimately leading to increased confidence in utilizing AI. However, tasks should be devised in a way that shows how AI contributes to learning outcomes. This is important, because developing competence is not enough for PU. The policy-makers of educational institutions should note the high explanatory power, that is, the predictability, of the model for integrated AI strategy design. Concurrent investment in innovation in teaching methodologies and AI-specific skills is the way forward with regard to adopting AI. According to Adamakis and Rachiotis (2025), Kurtz et al. (2024), and others, policies should not be limited to frameworks restricting integrities. Competences should be responsibly and effectively trained to engage in AI.

11. Limitations of the Study and Future Research

The study has several limitations. Because it is cross-sectional in nature, it is not possible to establish causality or to assess how these variables may evolve over time. Therefore, longitudinal research is needed to examine the evolution of these variables. Another limitation of the study is the generalizability of the findings. The sample was drawn from a single institution, and institutional culture may influence the variables examined in this study. Consequently, the study should be replicated across multiple institutions and diverse cultural contexts. All variables were measured through self-reporting, introducing the potential for common method bias. Future studies could incorporate objective measures of these variables. As the extended model suggests, AI adoption is a multilevel phenomenon. Therefore, future research should examine the influence of additional factors, such as trust in AI technologies, ethical considerations, social influence, and AI anxiety.

12. Conclusion

This study provides a theoretically grounded extension of the Technology Acceptance Model by demonstrating that students' adoption of AI tools in higher education is not solely driven by traditional cognitive evaluations of usefulness and ease of use but is significantly shaped by pedagogical context and interactional competence. Specifically, authentic learning emerges as a foundational antecedent that enhances both perceived usefulness and perceived ease of use, while prompt engineering competence primarily reduces interactional effort rather than influencing outcome-based evaluations. These findings highlight that AI adoption is a contextually constructed process, where meaningful integration into real-world learning tasks plays a critical role in shaping students' perceptions and intentions. The study thus contributes to the refinement of TAM by positioning pedagogical design and competence as upstream factors influencing cognitive evaluations. Future research should extend this model across different institutional and cultural contexts to enhance generalizability. Longitudinal designs are also needed to examine how perceptions and competencies evolve over time. In addition, further studies may incorporate additional variables such as AI literacy, trust, ethical awareness, and social influence to provide a more comprehensive understanding of AI adoption in higher education. Overall, the findings suggest that maximizing the educational value of AI requires not only accessible technologies but also thoughtfully designed learning environments and targeted development of students' interactional competencies.

AI Statement: AI tools were used solely for language refinement during the preparation of this manuscript. The authors wrote, reviewed, and approved all manuscript content and take full responsibility for the accuracy, integrity, and originality of the work.

Ethics statement: All participants provided informed consent prior to participation in the study. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Ethical approval was obtained from the Research Ethics Committee (REC) at the University of Hail, Saudi Arabia (Approval No. H-2026-090). Participation was voluntary, and all responses were collected anonymously and analyzed in aggregated form

References

- Adamakis, M., and Rachiotis, T. (2025). Artificial intelligence in higher education: A state-of-the-art overview of pedagogical integrity, artificial intelligence literacy, and policy integration. *Encyclopedia*, 5(4), 180.
- Alenezi, F. Y. (2023). Artificial Intelligence Versus Arab Universities: An Enquiry into the Saudi Context. *Scientific Journal of King Faisal University, Humanities and Management Sciences*, 24(1).
- Ali, F., Ahmed, A., Alipour, M. A., and Terashima-Marin, H. (2025). Adoption of AI-coding assistants in programming education: exploring trust and learning motivation through an extended technology acceptance model. *Journal of Computers in Education*, 1-39.
- Alshammari, S. H., Almankory, A. Z., and Alrashidi, M. E. (2025). The effects of awareness and trust on students' willingness to use ChatGPT: an integrated TAM-ECM model. *RIED-Revista Iberoamericana de Educación a Distancia*, 28(2), 155-180.
- Alshammari, S. H., Alshammari, A. E. A., Alghamdi, B. A., Alrashidi, M. E., Rosli, M. S., and Risdianto, E. (2026). Artificial Intelligence in Educational Technology: A Systematic Review of Pedagogical Opportunities, Implementation Challenges, and Ethical Considerations. *International Journal on Studies in Education*, 8(1), 17-37.
- Alshammari, S. H., and Alshammari, M. H. (2024). Modelling the effects of perceived system quality and personal innovativeness on the intention to use metaverse: a structural equation modelling approach. *PeerJ Computer Science*, 10, e2331.
- Alshammari, S. H., Alrashidi, M. E., Alshammari, M. H., Alshammari, A. E. A., and Alkhwalidi, A. F. (2025). Determinants of student adoption of artificial intelligence applications in higher education. *Scientific Reports*, 15(1), 35921.
- Alshammari, S. H., and Alshammari, M. H. (2024). Factors affecting the adoption and use of ChatGPT in higher education. *International Journal of Information and Communication Technology Education (IJICTE)*, 20(1), 1-16.
- Anam, R. K. (2025). Prompt engineering and the effectiveness of large language models in enhancing human productivity. *arXiv preprint arXiv:2507.18638*.
- Anderson, J. C., and Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin*, 103(3), 411.
- Awang, P. (2015). *SEM made simple: A gentle approach to learning Structural Equation Modeling*. MPWS Rich Publication, Bangi.
- Bashardoust, A., Feng, Y., Geissler, D., Feuerriegel, S., and Shrestha, Y. R. (2024). The Effect of Education in Prompt Engineering: Evidence from Journalists. *arXiv preprint arXiv:2409.12320*.
- Bayaga, A. (2025). Leveraging AI-enhanced and emerging technologies for pedagogical innovations in higher education. *Education and Information Technologies*, 30(1), 1045-1072.
- Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. routledge.
- Chukwuere, J. E. (2024). Developing generative AI chatbots conceptual framework for higher education. *arXiv preprint arXiv:2403.19303*.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. routledge.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 13(3), 319-340.
- Federiakina, D., Molerov, D., Zlatkin-Troitschanskaia, O., and Maur, A. (2024). Prompt engineering as a new 21st century skill. In *Frontiers in Education* (Vol. 9, p. 1366434). Frontiers.
- Flores-Velázquez, C. H., Olivares-Zegarra, S., Dávila-Ignacio, C., Arévalo-Tuesta, J. A., Morales-Romero, G. U. I. L. E. R. M. O., Trinidad-Loli, N. I. C. É. F. O. R. O., ... and Aldana-Trejo, F. L. O. R. C. I. T. A. (2024). A bibliometric review of studies about the acceptance of artificial intelligence technologies in teaching and learning in higher education. *International Journal of Learning, Teaching and Educational Research*, 23(3), 275-292.
- Fragkaki, M., Kouzeli, A., Stavropoulou, E., Karagiannopoulou, S., Bekiari, I., and Karava, A. M. M. (2025). Integrating ICT and AI Tools: From Children's Rights and Human Rights Curriculum Theories to Authentic Actions in Preliminary Education. In *Human Rights in the Digital Age*. IntechOpen.
- Gibreel, O., and Arpaci, I. (2025). Development and validation of the prompt engineering competence scale (PECS). *Information Development*, 02666669251336455.
- Hair Jr, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2010). Multivariate data analysis. In *Multivariate data analysis* (pp. 785-785).
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., and Ray, S. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook*. Springer international publishing.

- Henseler, J., Ringle, C. M., and Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the academy of marketing science*, 43(1), 115-135.
- Hughes, L., Malik, T., Dettmer, S., Al-Busaidi, A. S., and Dwivedi, Y. K. (2025). Reimagining higher education: Navigating the challenges of generative AI adoption. *Information Systems Frontiers*, 1-23.
- Kanont, K., Pingmuang, P., Simasathien, T., Wisnuwong, S., Wiwatsiripong, B., Poonpirome, K., ... and Khlaisang, J. (2024). Generative-AI, a learning assistant? Factors influencing higher-ed students' technology acceptance. *Electronic journal of e-Learning*, 22(6), 18-33.
- Kitchen, E., Goh, E., Steriopoulos, E., Harkison, T., Drake, C., Robertson, M., ... and Waterston, L. (2025). Preparing tourism, hospitality and events graduates to be industry ready: extending the three-factor model of authentic learning. *Studies in Higher Education*, 50(11), 2448-2466.
- Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications.
- Kurtz, G., Amzalag, M., Shaked, N., Zaguri, Y., Kohen-Vacs, D., Gal, E., ... and Barak-Medina, E. (2024). Strategies for integrating generative AI into higher education: Navigating challenges and leveraging opportunities. *Education Sciences*, 14(5), 503.
- Lang, J. C. (2024). Embracing generative AI for authentic learning. *Creative Education*, 15(1), 1-20.
- Lee, D., and Palmer, E. (2025). Prompt engineering in higher education: a systematic review to help inform curricula. *International Journal of Educational Technology in Higher Education*, 22(1), 7.
- Mukhamedkarimova, D. F., and Umurkulova, M. M. (2025). Factors Contributing to Higher Education Students' Acceptance of Artificial Intelligence: A Systematic Review. *European Journal of Educational Research*, 14(4), 1373-1388.
- Mzwri, K., and Turcsányi-Szabo, M. (2025). The impact of prompt engineering and a generative AI-driven tool on autonomous learning: A case study. *Education Sciences*, 15(2), 199.
- Neo, M., Neo, K. T. K., and Tan, H. Y. J. (2012). Applying Authentic Learning Strategies in a Multimedia and Web Learning Environment (MWLE): Malaysian Students' Perspective. *Turkish Online Journal of Educational Technology-TOJET*, 11(3), 50-60.
- Özbay, M., Özbay, F., and Tutkunca, S. S. (2025). Artificial Intelligence In Universities: A Study On Academics' Views. *Romaya Journal*, 5(Special Issue (SI) IXASC 2025), 73-85.
- Preacher, K. J., and Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, 40(3), 879-891.
- Radović, S., Firsova, O., Hummel, H. G., and Vermeulen, M. (2021). Strengthening the ties between theory and practice in higher education: An investigation into different levels of authenticity and processes of re-and de-contextualisation. *Studies in Higher Education*, 46(12), 2710-2725.
- Roe, J., Perkins, M., Somoray, K., Miller, D., and Furze, L. (2025). To Deepfake or Not to Deepfake: Higher Education Stakeholders' Perceptions and Intentions towards Synthetic Media. *arXiv preprint arXiv:2502.18066*.
- Rughiniş, C., Vulpe, S. N., Turcanu, D., and Rughiniş, R. (2025). AI at the knowledge gates: institutional policies and hybrid configurations in universities and publishers. *Frontiers in Computer Science*, 7, 1608276.
- Salinas-Navarro, D. E., Vilalta-Perdomo, E., Michel-Villarreal, R., and Montesinos, L. (2024). Using generative artificial intelligence tools to explain and enhance experiential learning for authentic assessment. *Education Sciences*, 14(1), 83.
- Sallai, D., Cardoso-Silva, J., Barreto, M., Panero, F., Berrada, G., and Luxmoore, S. (2024). Approach generative AI tools proactively or risk bypassing the learning process in higher education. *LSE Public Policy Review*, 3(3).
- Shahzad, M. F., Xu, S., and Javed, I. (2024). ChatGPT awareness, acceptance, and adoption in higher education: the role of trust as a cornerstone. *International Journal of Educational Technology in Higher Education*, 21(1), 46.
- Shata, A., and Hartley, K. (2025). Artificial intelligence and communication technologies in academia: faculty perceptions and the adoption of generative AI. *International Journal of Educational Technology in Higher Education*, 22(1), 14.
- Vafadar, M., and Amani, A. M. (2024). Academic education in the era of generative artificial intelligence. *Journal of Electronics and Electrical Engineering*, 117-133.
- Vieriu, A. M., and Petrea, G. (2025). The impact of artificial intelligence (AI) on students' academic development. *Education Sciences*, 15(3), 343.
- Wahdah, S. I., Kurniawan, R. Y., and Irawan, N. (2025). Adoption of ChatGPT in higher education: Insights from the Unified Theory of Acceptance and Use of Technology model. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 312-327.
- Woo, D. J., Wang, D., Yung, T., and Guo, K. (2024). Effects of a prompt engineering intervention on undergraduate students' AI self-efficacy, AI knowledge and prompt engineering ability: A mixed methods study. *British Educational Research Journal*.
- Zhao, Z., An, Q., and Liu, J. (2025). Exploring AI tool adoption in higher education: Evidence from a PLS-SEM model integrating multimodal literacy, self-efficacy, and university support. *Frontiers in Psychology*, 16, 1619391.

Appendix A: Measurement Items

Prompt Engineering Competence for Learning with ChatGPT

PEC1: I can write academic prompts that help ChatGPT provide accurate and relevant information for my learning.

PEC2: I can break down academic tasks into clear steps and write prompts that guide ChatGPT through each step.

PEC3: I can adjust my prompts using different formats (such as questions, lists, or scenarios) to improve the quality of ChatGPT's academic responses.

PEC4: I can improve ChatGPT's answers by specifying requirements such as role, tone, length, or academic format.

PEC5: I can enhance ChatGPT's responses by adding background information and recognizing possible limitations or biases in its output.

PEC6: I can write prompts that encourage ChatGPT to perform higher-order thinking, such as analyzing, comparing, or evaluating academic concepts.

PEC7: I can use real examples or academic scenarios in my prompts to make ChatGPT's responses more relevant for my learning.

PEC8: I can improve ChatGPT's answers by refining my prompts based on previous responses.

PEC9: I can evaluate ChatGPT's responses for accuracy and credibility when completing academic tasks.

Perceived Authentic Learning with ChatGPT

AL1: I feel that the information generated by ChatGPT is relevant to my learning.

AL2: The activities I do using ChatGPT feel authentic and relevant to my academic tasks.

AL3: I enjoy using ChatGPT as a tool to learn academic topics.

AL4: The examples and explanations provided by ChatGPT help me understand course content better.

AL5: ChatGPT helps me reflect more deeply on the topics I am studying.

AL6: The information suggested by ChatGPT (such as references or resources) is useful for my learning.

AL7: The explanations and content generated by ChatGPT are relevant to my academic needs.

AL8: I like that ChatGPT helps me explore additional sources and external information while learning.

AL9: I can apply what I learn from using ChatGPT in real academic tasks or real-life situations.

Perceived Usefulness of ChatGPT for Learning

PU1: Using ChatGPT improves my learning performance.

PU2: Using ChatGPT enhances my effectiveness in studying academic topics.

PU3: Using ChatGPT increases my learning productivity.

PU4: I find ChatGPT useful for my academic learning.

Perceived Ease of Use of ChatGPT

PEU1: Learning to use ChatGPT is easy for me.

PEU2: I find it easy to get ChatGPT to do what I want for my learning.

PEU3: My interaction with ChatGPT is clear and understandable.

PEU4: It is easy for me to become skillful in using ChatGPT.

Behavioral Intention to Use ChatGPT

BI1: Assuming I have access to ChatGPT, I intend to use it for my learning.

BI2: Given that I have access to ChatGPT, I predict that I will use it for academic purposes.

BI3: I plan to continue using ChatGPT in the future for my learning.

Validating an Instrument of Adult Learning Needs in Online and Blended Learning

Anh Nguyet Diep^{1,6}, Khuyen Dinh², Anne-Françoise Donneau¹, Thanh Ngoc Phuong Phan³, Minh Hien Vo⁴, Celine Cocquyt⁵, Maurice de Greef⁵, Tom Vanwing⁵ and Chang Zhu⁵

¹Department of Public Health, Faculty of Medicine, University of Liege, Belgium

²Faculty of Business and Law, University of Roehampton, London, UK

³Department of General English and English for Specific Purposes, School of Foreign Languages, Can Tho University, Vietnam

⁴Graduate School, Can Tho University, Vietnam

⁵Department of Educational Sciences, Faculty of Psychology and Educational Sciences, Vrije Universiteit Brussel, Belgium

⁶Department of Information Technology, College of Information and Communication Technology, Can Tho University, Vietnam

anhnguyet.diep@uliege.be (Corresponding author)

khuyen.dinh@roehampton.ac.uk

afdonneau@uliege.be

ptnphuong@ctu.edu.vn

vmhien@ctu.edu.vn

celine.cocquyt@vub.be

maurice.de.greef@vub.be

tvwing@vub.be

chang.zhu@vub.be

<https://doi.org/10.34190/ejel.24.4.4684>

An open access article under [CC Attribution 4.0](#)

Abstract: It is a challenge to promptly respond to the learning needs of students, particularly adult learners characterized by rich lived-experiences, (sometimes) negative prior learning experience, multiple and contrasting motivation types alongside considerable socio-demographic heterogeneity. Furthermore, developing a typology of adults' learning needs remains a challenging endeavour given the different perspectives in learning needs conceptualization and the different theories concerning adult learning in both traditional and online and blended learning (OBL). In particular, no validated instrument currently exists to identify and measure adult learners' needs in OBL contexts, representing a gap that needed to be addressed. This paper developed and validated an instrument to identify adults' learning needs based on adult learning, self-regulated and motivational theories, and technological acceptance models. We deconstructed adults' learning needs into three categories based on the existence, relatedness, and growth (ERG) theory of human needs. Data (N=209) were collected from adult learners following blended learning programs in eight centres of adult education in Flanders, Belgium. Confirmatory factor analysis (CFA) revealed that the instrument comprising of nine dimensions of learning needs displayed adequate model fit and factor loadings. Multivariate analysis of covariance (MANCOVA) highlighted that learners who differed in educational attainment, age, and ICT skills reported differences in their needs across the sub-dimensions. Motivational orientations, however, did not significantly differentiate learning needs at the multivariate level. The validated instrument, hence, offered practitioners a diagnostic tool for needs-based OBL design. Theoretically, the study challenged the assumption that motivational orientation alone predicted differentiated need profiles in OBL, opening new avenues for debate on the interplay between motivation, prior experience, and adult learning needs.

Keywords: Adult learners, Learning needs, Socio-cognitive theories, Technology acceptance model, Motivational orientations

1. Introduction

Adult learners bring to the classroom a broad range of lived and prior learning experiences (Ross-Gordon, 2003), alongside diverse motivation profiles and population-specific constraints such as time pressures and self-esteem concerns (Cercone, 2008). Adult learning needs, conceptualized not as generic skill deficiencies, but as an integrated set of psychological, situational, and instructional requirements that are foundational to sustaining adult motivation and performance in technology-mediated spaces (Diep et al., 2019), do substantially

differ. Understanding how to capitalize on adult learners' strengths, i.e., rich lived experiences, and reduce the constraints is of crucial importance to help adults better achieve in their registered program.

From a socio-constructivist perspective, online and blended learning (OBL) combining online and face-to-face instruction (Graham, 2006) postulated that through online interaction, adults exchange ideas and reflect on existing knowledge, resulting in reinforced or transformed frames of reference (Mezirow, 2000). In what ways can instructors and practitioners take into account both adult learners' characteristics and the requirements of learning in a blended learning environment to support them during the learning process? The research field has presented both theoretical and empirical works on adult learners' profiles, adult learners' motivation and self-regulated learning (Mertens, Maeyer, and Donche, 2024), the challenges of learning in a blended learning environment for adults, and also how learning activities can be best designed to optimize learning. To address those characteristics specific to adult learners, prior research has leaned on separate tools, such as scales evaluating digital skill literacy (van Deursen, Helsper, and Eynon, 2014), academic self-directedness (Botha and Masenge, 2022), or generalized e-learning readiness instruments (Loock et al., 2022). However, these existing measurement tools often evaluate psychological, technical, or environmental traits in isolation, and many remain grounded in traditional face-to-face andragogical models rather than modern digital dynamics. Few instruments are explicitly tailored to capture the overlapping systemic and psychological constraints inherent to adults navigating OBL spaces. Consequently, an instrument to identify adults' learning needs remains scarce, particularly in an OBL environment.

A closer examination of existing instruments revealed specific methodological shortcomings that further justified the present study. The Distance Education Learning Environments Survey (DELES; Walker and Fraser, 2005), while widely used, was developed for higher education distance learners, and did not account for the face-to-face component inherent in blended learning, nor did it address the motivational heterogeneity characteristic of adult populations. Similarly, the e-learning readiness scales reviewed by Loock et al. (2022) primarily assessed learners' preparedness to engage with digital platforms rather than the full spectrum of psychological and instructional needs that sustained adult motivation during the learning process. The digital skills framework proposed by van Deursen, Helsper, and Eynon (2014) captured one important dimension of adult learners' needs but ignored relational and structural needs that were equally critical in OBL. Notably, the accelerated transition to OBL learning during and after the COVID-19 pandemic has further exposed these gaps. For instance, studies conducted in post-pandemic adult education contexts highlighted that digital exclusion, social isolation, and the lack of face-to-face support structures have intensified adults' needs in ways that pre-pandemic, single-domain instruments were not adequate to capture (Chambers and Whitfield, 2025; Mertens, Maeyer, and Donche, 2024). Together, these limitations underscored the need for an integrated, OBL-specific instrument that simultaneously addressed the existence, relational, and growth dimensions of adult learning needs.

To address this gap, the present study develops an ERG-based framework (Alderfer, 1969) and operationalizes it into a validated instrument measuring adult learning needs across existence, relatedness, and growth dimensions. We further examine need differences by motivational profiles and socio-demographics.

2. Theoretical Framework of Adult Learners' Needs in OBL

2.1 Conceptualization of Adult Learners' Needs

Although the conceptualization of adults' learning needs has been widely discussed in the literature, a large body of research has shown that this notion could be described from multiple approaches. The functional approach, which was based on scientific empiricism, highlighted that educational needs can be identified objectively (Sava, 2012; Pearce, 1995). Along these same lines, the needs of learners were focused on what people need to know in order to do something better than they were currently doing (Pearce, 1995). The empowerment approach, which has relied on the subjectivist perspective, considered classifying learners' needs as a process in which instructors play a central role (Sava, 2012; Pearce, 1995). Within this paradigm, Pearce (1995) postulated that course designers tended to follow the functionalist approach when it came to formal needs assessments and suggested that adult learners were key actors in identifying needs rather than instructors. Sava (2012) primarily applied the empowerment perspective in informal needs analyses.

Combining the two paradigms mentioned above, Laurillard (2012) proposed a new conceptual viewpoint in which instructors played a vital role in identifying educational needs. To that end, educators needed to clarify learners' prior experiences, their strengths and weaknesses in order to help them transcend their preferred learning styles. In the current study, we adopted the conceptualization of learning needs from Diep et al. (2019),

defining adults' learning needs as "those that motivate learners and substantially enhance their learning, the lack of which will lead to demotivation and failure to achieve major learning and personal goals" (p.227-228).

2.2 The Construction of a Conceptual Framework and the Classification of Adult Learners' Needs in OBL

2.2.1 Literature review

A body of research demonstrated that there was no universal framework for adapting learners' needs in OBL. Instead, there were several theoretical models which equally supported designing OBL addressed to learners' needs. In the current study, several relevant theories have been selected and critically examined in terms of their relevance and limitations for understanding adult learners' needs in OBL.

The Community of Inquiry theory (Col) (Garrison, Anderson, and Archer, 2000) was the first framework devised for guiding construction and facilitation in an online learning environment. Rooted in the socio-constructivist theory, the theoretical model emphasizes the construction of knowledge as a collaborative process. Accordingly, it required the active participation of learners, well-structured design and effective facilitation strategies developed by the instructors. However, the Col model has been noted to inadequately explore how individual motivational factors influence learning, an aspect of particular relevance when working with adult populations (Yang and Lay, 2024). Furthermore, the rapid expansion of blended and/or fully online delivery in adult education following the COVID-19 pandemic has revealed new challenges that the original Col framework did not anticipate, namely heightened digital inequalities among adult learners, the erosion of informal peer support that typically occurs in face-to-face settings, and increased demands for instructor presence in asynchronous environments (Chambers and Whitfield, 2025). These developments call for updated theoretical lenses that account for the structural and motivational vulnerabilities specific to adults in contemporary blended contexts.

While the Col focuses on guiding the design and facilitation of online learning activities, the self-determination theory (SDT) (Ryan and Deci, 2000) closely examined human motivation and personality. This theory highlighted three essential components of human needs: competence, autonomy, and relatedness. The first element, competence, was closely related to self-efficacy and indicated a person's belief and confidence in whether he/she could successfully accomplish a set task independently. Autonomy, the second component, refers to self-directed action that will result in real changes. Niemiec and Ryan (2009) posited that autonomy must go hand in hand with competence as both elements were equally essential for maintaining intrinsic motivation. The third feature, relatedness, was rooted in the viewpoint that human needs were closely related to innate tendencies toward connectedness, effectiveness, and coherence (Ryan and Deci, 2000). Based on the SDT, various empirical studies have been conducted to gain insight into learners' needs in adult learning environments. For example, Reeve (2002) found that autonomy plays an essential role in learners' motivation. Furthermore, the study of Knowles, Holton, and Swanson (2005) provided empirical evidence that adult learners are internally motivated and self-directed.

While SDT addresses motivational needs, it does not account for the technological conditions that mediate learning in OBL. Technology affordance theory thus offers a complementary lens. The theory of affordance was initially proposed by Gibson (1979), who indicated the characteristics of an environment, the agents and the connection between them. When it came to virtual learning settings, the social values of mobile learning devices and online learning processes could be implied by the concept of technology affordance. Not surprisingly, a substantial body of research aiming to connect technology to the affordance theory (Oliver, 2013). For example, Gaver (1991) posited that the study of affordance should "focus not on technologies or users alone, but the fundamental interactions between the two" (p. 5).

While the theories discussed above illuminate distinct aspects of OBL design, none offers a sufficiently integrated account of adult learners' needs. The ERG theory, discussed in the following section, addresses this limitation. Furthermore, this theoretical fragmentation is mirrored in the limitations of existing measurement instruments. Col-based scales (e.g., Arbaugh et al., 2008) focused primarily on cognitive and teaching presence, overlooking the existence-level needs (technical access, transparency) that were particularly salient for adult learners with limited digital experience. SDT-based instruments (e.g., Walker and Fraser, 2005) captured motivational and relational dimensions well but did not integrate the technological acceptance conditions that shaped whether adults could activate or exercise their autonomy in OBL environments. Technology acceptance models such as TAM (Davis, 1989) addressed system usability but neglected the relational and growth needs that deemed critical to adults' cognitive engagement beyond platform adoption. Taken together, these frameworks addressed adult learning needs in silos rather than as an integrated system — a limitation that becomes particularly pronounced in blended learning contexts where existence, relational, and growth demands co-occur

and interact simultaneously. A theoretical model that considers the different demands of learners was put forward by Milheim (2012) to create an extensible framework for OBL learning design based on current ideas. This "hierarchy of learner's needs" paradigm, which had its roots in Maslow's theory, is composed of five levels: self-actualization, relationships, safety, self-esteem, and physiology. Nonetheless, the model focused more on material preparation than on explaining why content matters in terms of fulfilling learners' needs. Additionally, communication technologies were brought up frequently in a variety of categories, making it challenging to explain their role in addressing learners' needs within the instructional framework. Furthermore, according to Abela (2009), Maslow's needs typology might be overly strict, causing categories to overlap. For instance, Milheim (2012) notes that pre-course preparation was a safety requirement and that delivering a checklist of necessary items should be received as a physiological need. These two fell within both categories.

2.2.2 Adult learners' needs viewed from the perspective of the ERG theory

Alderfer (1972) created a new model based on the findings of empirical research and an expansion of Maslow's hierarchy of needs to address the shortcomings previously mentioned. Accordingly, the three primary components of human needs have been categorized as existence, relatedness, and growth. Existence needs concerned various safety-related, physiological, and material needs; relatedness needs involved a sense of security, belonging and respect; and growth needs included self-esteem and self-actualization.

Compared to Maslow's rigid hierarchy, the ERG theory offers greater flexibility by allowing simultaneous activation of multiple need levels, which better reflects the non-linear nature of adult learning trajectories in OBL (Alderfer, 1972). Furthermore, the model might offer a helpful technologically assisted learning environment to enhance the process of knowledge development. In terms of student motivation, the ERG theory can effectively illustrate different forms of motivation. One way to understand the growth needs in the ERG model is to consider the motivation of learners to advance their knowledge and skills for personal or professional purposes. On this basis, Diep et al. (2019) proposed a conceptual framework for designing OBL to address adult learning needs, which is summarized below.

Existence needs

The first dimension, existence, emphasized the role of the technical and structural conditions that supported adult learners in achieving their learning goals as well as organizing their learning process. Thus, three sub-dimensions emerged under existence needs: system functionalities, technical support, and transparency. As for system functionalities, usefulness and ease of use of the LMS were highlighted as consistent predictors of learning retention (Davis, 1989; Venkatesh, Thong and Xu., 2012). Concerning technical support, several studies posited that learners achieved the most and felt more engaged in virtual learning courses when they received adequate technical support (Addae, Abakah, and Amuzu, 2025). Put differently, providing technical support was vital to enhancing learners' confidence in using the online system so that they could be more active and interactive in their learning (Prior et al., 2016). On the contrary, a lack of support on technical issues may lead to a negative attitude toward the use of the LMS in their learning. Thus, it was essential that learners did not encounter technical issues in the first place. The last component under existence, transparency, highlighted the importance of the transparent clarification of learning goals and outcome expectations when it came to the nature and functioning of the course (Poll, Widen and Weller, 2014; Dixon, 2014). In addition, Locke and Latham (2002) proposed that clear expectations of assessment were useful in order to provide learners with a global picture of what they must do, how they could self-evaluate their learning progress, as well as how to make more effort when necessary to address the discrepancies in their knowledge acquisition.

Relatedness needs

The motivation for fostering new social relationships and relief from daily routines (Boshier, 1991) was aligned with the dimension of relatedness. The ERG theory, viewed from a social inclusion perspective, displayed the potential to demonstrate how engaging in educational programs could nurture the two vital features of social inclusion, namely connections and self-activation (De Greef, Segers, and Verté, 2012). By blending the literature on knowledge construction and socio-constructivism with this dimension of learner motivational orientations, three main components, namely online interaction, instructor facilitation, and sense of belonging, were presumed to vary as a function of relatedness—the second category of adult learners' needs. Concerning online interaction, Laurillard (2012) highlighted that peer interaction played a vital role in the construction and cultivation of knowledge. Accordingly, collaborative and interactive activities with peers enabled learners to utilize and activate their prior knowledge. In addition, participating in a collaborative community provided them with useful feedback to further improve their learning performance through a reflective process. Furthermore,

to ensure the effectiveness of peer collaboration in virtual learning communities, the instructors' role has been emphasized in a large body of literature (Hsieh and Tsai, 2012; Chambers and Whitfield, 2025). The responsibilities of online facilitators generally involved moderating the discussion and keeping it on track, responding to learners' questions and concerns in a timely and effective manner, and motivating them to express their views and engage in the course (Peltier, John and William, 2007). Drago and Peltier (2004) posited that instructor-to-student interactions were critical for enhancing learner connectedness to the online learning community. Thus, the instructor's online facilitation served as a second sub-category under relatedness. Furthermore, beyond the need to interact with peers, sense of belonging was another essential feature of the blended learning environment as it contributed to encouraging learners to interact with their peers (Koul, Fisher, and Shaw, 2011). In this respect, a number of studies have found that sense of belonging played a crucial role in motivating learners to persevere and complete the course when under pressure (Mebane et al., 2008).

Growth needs

Drawing from constructivism, autonomy, relevance, and competence, as well as andragogical assumptions (Knowles, 1984), self-determination theory (Ryan and Deci, 2000), and self-directed learning, we suggested three components that fall under this classification, namely autonomy, relevance, and competence. Autonomy, which was rooted in the self-determination theory, refers to affordable/accessible opportunities for learners to seek knowledge, to achieve, to be independent and to feel competent in learning (Bangert, 2004; Ross-Gordon, 2003; Walker and Fraser, 2005). Along these same lines, Bangert (2004) and Knowles (1984) highlighted the importance of considering a range of diverse talents, learning strategies, and prior learning in order to nurture learning autonomy. Relevance, which was based on constructivism and andragogy, emphasized authentic activities that were highly relevant to real-world experience (Reeves, Herrington, and Oliver, 2002). In addition, Woo and Reeves (2007) recommended meaningful online interaction in which learners have opportunities to engage, negotiate, and collaborate, etc. with their peers to resolve authentic tasks. Competence needs refer to "the experience of behaviour as effectively enacted" (Niemic and Ryan, 2009, p. 135). In order to enhance competence, feedback and learning activities of a formative nature were necessary. While access to feedback enables learners to enhance their learning, experience, self-efficacy (Bandura, 1986) and/or achieve the learning goals at hand (Niemic and Ryan, 2009), learning activities accompanied by cognitive challenges assist them in testing and transcending their academic capacities.

2.3 Adult Learners' Motivational Orientations and Learning Needs

Adult learners generally brought different types of motivation when registering for a course or a program. These motivation types could be summarized as extrinsic or intrinsic motivation (Deci and Ryan, 2000). The extrinsic dimension referred to the reward that could be obtained as a result of the invested effort whereas intrinsic motivation pertained to the internal pleasure and fulfilment achieved through the learning per se. Obtaining a diploma and the pursuit of new knowledge and skills were specific examples of extrinsic and intrinsic motivation, respectively. Intrinsically motivated learners were more likely to employ a deep learning approach to learning, which included seeking peer and instructors' support and proactively managing time (Pintrich and Zusho, 2007). On the other hand, extrinsic learners tended to apply a surface learning approach such as memorization and exerted minimal effort on required learning tasks (Pintrich and Zusho, 2007). Therefore, it was argued that learners with different types of motivation will have different learning needs - a distinction that existing instruments have rarely operationalized, representing a gap that the present study addresses. As a consequence, it is hypothesized that adult learners of different motivational orientations, namely diploma-oriented and knowledge-and-skills oriented, will differ in their learning needs.

Next to motivation, adult learners' socio-demographics including gender, age, employment status, ICT skills, and educational levels proved to be influential factors in learning in a blended environment. Older adults might have lower computer self-efficacy and thus presented more needs for technical support. Additionally, Ke and Kwak (2013) discovered that older learners favoured online interactions more than younger learners. Adult learners who were more advanced in computer skills may not experience technical problems, and hence probably required less support from the instructors. Furthermore, adult learners who were heavily occupied with other obligations might opt for an individual approach to learning. Consequently, online interaction or collaborative learning was less favoured. For example, female learners were found to be more engaged in social interactions and working learners were more active in facilitating discussion (Diep et al., 2017). As for educational levels, learners who had a higher degree or were more capable than their peers seemed to attribute less value to collaborative work, mentioning that they needed more time for peer correction (Kollock, 1999). In other words,

we recommended that socio-demographic variables including gender, age, employment status, ICT skills, and educational levels should be investigated to yield more understanding of adults' learning needs.

2.4 Research Questions

In order to examine adults' learning needs in OBL, this research focused on the following research questions.

RQ1. To what extent could the nine dimensions of adults' learning needs in an OBL environment be distinguished?

RQ2. Did adult learners with different motivational orientations differ in their learning needs, controlling for age groups, gender, ICT skills, and educational levels?

The first main objective was to establish the structural validity and distinctiveness of the nine learning needs dimensions (**RQ1**). This serves as a psychometric prerequisite for the subsequent analysis. Based on the validation of the dimensions, specific adults' learning needs were analyzed related to adult learners' background variables and their motivational orientations (**RQ2**).

3. Methodology

3.1 Research Design and Participants

The study adopted a cross-sectional design, in which adult learners following a blended learning program across different adult education centers (Centrum voor Volwassenenonderwijs – CVO) in Flanders, Belgium were invited to participate in the questionnaire. The learners completed the questionnaire online in most cases. When possible, the researcher distributed the questionnaire in the classroom at the beginning of the class. At the time of completing the questionnaire, the adult learners had completed at least one semester of a one-year program or were in their final semester. In total, 209 participants completed the questionnaire, drawn from 14 campuses of 8 CVOs. As for the programs they were following, participants were enrolled in one of the 13 programs as shown in Table 1. As the number of learners following each program was not equal and there were programs with few learners, all 13 programs were collapsed into 4 categories based on their course objectives. Accordingly, four categories of programs were created, i.e., category 1 comprised pedagogical programs, category 2 administrative, category 3 caring, and category 4 technical.

Table 1: Participants' current training programs

Program	Number	Percentage	Category
Specific Teacher Training	108	51.7	(1) Pedagogical
Freelance coaching	19	9.1	
Subtotal	127	60.8	
Librarian	5	2.4	(2) Administrative
Administrative commercial assistant	4	1.9	
Medical administrative assistant	3	1.4	
Real Estate	9	4.3	
Dispatcher	7	3.3	
Subtotal	28	13.4	
Beauty	18	8.6	(3) Caring
Childcare	24	11.5	
Subtotal	42	20.1	
Computer and networks techniques	2	1	(4) Technical
Graphical designer	1	0.5	
Computer-Aided-Design draftsman in construction	7	3.3	
Webmaster and Social media manager	2	1	
Subtotal	12	5.7	
Total	209	100	

Table 2 provided information regarding the learners' socio-demographics. As can be seen, the number of female participants was twice as that of male learners. Young learners (≤ 25 years of age) constituted the largest group, followed by those aged between 26-34 (28.7%), and 35-44 (23.4%). Most of the learners were employed, with full-time working learners being the majority. Roughly half of them were married or cohabiting with a partner and the other half was single. As for education levels, more than half of the learners had obtained a tertiary education, i.e., holding a bachelor's or post-graduate degree. Two-thirds of the participants had no prior experience with blended learning.

Table 2: Adult learners' socio-demographics (N=209)

Categories	Number	Percentage
Gender		
Male	65	31.10
Female	144	68.90
Age group (M=33.09; SD=10.83)		
≤ 25	65	31.1
26-34	60	28.7
35-44	49	23.4
≥ 45	35	16.7
Marital status		
Married or living together	106	50.7
Single	96	45.9
Employment status		
Full-time employed	93	44.5
Part-time employed	27	12.9
Fully-registered students	89	42.6
Education levels		
Lower secondary	20	9.6
Higher secondary	72	34.4
Bachelor	66	31.6
Post-graduate	51	24.4
ICT skills		
Beginner	9	4.3
Immediate	109	52.2
Advanced	91	43.5
Blended learning experience		
Yes	80	38.3
No	126	60.3

3.2 Instrument

The instrument adopted existing scales that have been validated in previous research. In this study, we brought them together in one instrument to reflect three layers of adults' learning needs. The questionnaire consisted of two main sections as follows.

Section 1

The general information included the socio-demographics, the information regarding learners' current blended learning program and their (expected) grades, their experience with ICT skills and blended learning, and finally, their motivational orientations when following the program.

As for the motivational orientations, the participants were asked to indicate a Yes or No response regarding their motivation when enrolling in the program, including the following:

1. I'd like to get the diploma.
2. I'd like to improve my knowledge and skills for my current and/or future job
3. I'd like to make new friends and meet new people
4. I'd like to get more fun and relief from my daily routine and boredom.
5. I'd like to learn new knowledge and skills.
6. I participate because this course will prepare me for service to the community.
7. I participate because I need to comply with my employer's policy or to secure social benefits.

Descriptive statistics revealed that there were only two learners indicated their motivation as type (3), namely looking for new friends and relationships, as shown in Figure 1. Additionally, the motivational orientations related to daily routine relief (4), service to community (6) and securing social benefits (7) received little recognition from the learners. As a consequence, it was decided that only motivation 1, namely diploma-oriented, and motivation 2, knowledge-and-skills oriented, were included subsequently in discerning the differences in learners' needs as indicated by the adult learners. Once a learner indicated both motivation 1 and motivation 2, a third category was created. In other words, after recoding, the motivation orientation variable now consists of three categories: diploma-oriented (n=87), knowledge-and-skills oriented (n=60), and both diploma and knowledge oriented (n=24).

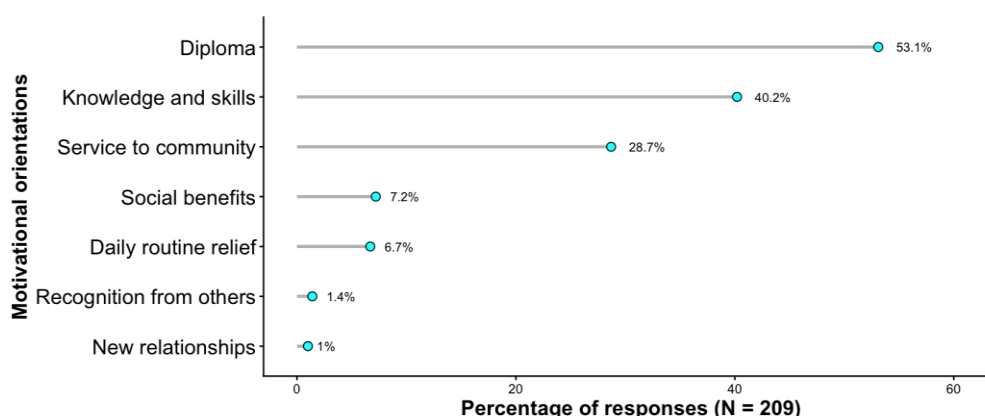


Figure 1: Proportion of responses for each of the motivational orientations of course participation

Section 2

The second section of the questionnaire comprised the main variables examined in the study. Accordingly, corresponding to each dimension of learning needs, four items were developed to capture the scale. In total, 36 items captured the 9 dimensions of learning needs. The items were phrased such that the learners indicated the extent to which they agreed with the statements describing the learning needs. The scale ranged from (1) strongly disagree to (5) strongly agree. All the scales were translated from English into Dutch using back-translation.

The adaptation process followed a systematic multi-step procedure. First, the original English items were translated into Dutch by a bilingual researcher with expertise in adult education, after which an independent bilingual translator back-translated the Dutch version into English. Discrepancies between the original and back-translated versions were discussed and resolved by consensus among the research team. Beyond linguistic equivalence, the wording of selected items was reviewed for contextual fit within the Flemish adult education (CVO) setting. For instance, references to "the learning platform" were standardized to reflect the specific LMS infrastructure used across the participating CVOs, and items relating to instructor roles were aligned with the facilitation practices typical of blended CVO programmes. These adaptations were minor and did not alter the theoretical meaning of the original constructs. The adapted Dutch questionnaire was subsequently reviewed by two experienced CVO instructors for face validity prior to distribution.

An overview of the variables is presented in Table 3. The Dutch version of the questionnaire is presented in Appendix A and English-version (not validated, for reference purpose) Appendix B.

Table 3: Overview of the adopted scales

Learners' needs	Authors	No. of items	Sample items
Existence			
Transparency	Ginns and Ellis (2009)	4	It is necessary that ... The instructor makes it clear right from the start what they expected from students.
System functionalities	Diep et al. (2016)	4	I can find the information on the learning platform easily.
Technical support	Lee et al. (2011)	4	I feel that I can get technical support when I need it.
Relatedness			
Facilitation (instructor)	Arbaugh et al. (2008)	4	The instructor encourages course participants to explore new concepts in this course.
Peer interaction	Walker and Fraser (2005); Laurillard (2012)	4	I can work collaboratively with peers in a project/task.
Sense of belonging	Rovai (2002)	4	I feel that I can rely on others in this course.
Growth			
Personal relevance	Walker and Fraser (2005)	4	I can apply my everyday experiences in class.
Competence support	Ausburn, (2004); Laurillard (2012)	4	The instructor offers learners timely guidance
Autonomy support	Walker and Fraser (2005)	4	I am in control of my learning.

3.3 Data Analysis Methods

First, confirmatory factor analysis (CFA) was conducted to examine the nine components of adult learners' needs. CFA was preferred over exploratory factor analysis (EFA) given that the factor structure was theoretically specified a priori based on the ERG framework, thus allowing for confirmatory rather than exploratory testing of the proposed dimensions. Regarding sample size adequacy for CFA, the present study's sample of $N=209$ satisfied commonly recommended thresholds. Following the widely cited 5:1 to 10:1 participant-to-parameter ratio (Hair et al., 2010), the model's 36 observed variables and nine latent factors require a minimum of approximately 180 observations, which the current sample exceeded. Additionally, simulation studies suggested that CFA models with strong factor loadings (above 0.50) and adequate communalities can achieve acceptable parameter recovery with samples as small as 150–200 (Wolf et al., 2013). The adequacy of the present sample was further supported empirically by the stable factor loadings (all above 0.50) and acceptable fit indices obtained in both the continuous (ML) and ordinal (WLSMV) estimation approaches, indicating that the sample was sufficient to reliably estimate the proposed factor structure. As the scale responses ranged from (1) strongly disagree to (5) strongly agree, there has not been a consensus whether to treat them as continuous or ordinal. The treatment of the scale might have consequences on estimation procedures. In this case, when treating the scale as ordinal, Byrne (1998) suggested Weighted Least Squares (WLS) estimation as the procedure, whereas with continuous scaling, Maximum Likelihood was suggested. Model fit was evaluated by goodness-of-fit indices including the Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Squared Residual (SRMR). Factor loadings above the threshold of 0.500 were deemed satisfactory. McDonald's Omega (ω) was employed to examine composite reliability of the resulting scales, with values above 0.7 being considered reliable.

Subsequently, a factor score was calculated for each dimension before further analyses were conducted. Descriptive statistics and measures of skewness and kurtosis were calculated to check the variability of the dependent variables, namely the nine learning need dimensions. Afterwards, multivariate analysis of covariance (MANCOVA) was conducted to answer RQ2. MANCOVA was chosen over separate univariate tests as it simultaneously examined group differences across all nine learning need dimensions while controlling for socio-demographic covariates, thereby reducing the risk of Type I errors. This was preceded by assumption verification by the Box's M and Levene's test of homogeneity of covariance matrices and variances, respectively. For both

tests, a non-significant p-value suggested the acceptance of the null hypothesis. For the multivariate tests, Pillai's Trace and Wilks' Lambda were chosen as tests of significance because they were considered as more robust. Once a dependent variable or a dimension of learning need was significant at the multivariate level, univariate tests were performed to examine the nuanced differences in the perceptions of learning needs among learners of different profiles.

All analyses were performed using R software (R Core Team, 2023).

4. Findings

RQ1. The dimensions of the learning needs

Validation of learning needs instrument

As the scales were developed based on a predefined theoretical framework, we specified a priori model with nine separate scales making up the three layers of learning needs. Accordingly, the Existence layer consisted of three factors: transparency, the quality of the learning platform (LMS quality), and technical support. The Growth layer was made up of three factors, namely the relevance (of the learning content), competence support, and the autonomy. Finally, the Relatedness layer comprised three factors, i.e., instructor's facilitation, peer collaboration, and sense of belonging.

When treating the variables as continuous and using Maximum Likelihood as the estimation procedure, the model revealed adequate fit with $\chi^2(558)=981.276$, $p<.001$. The CFI had a value close to 0.9 ($=0.885$); RMSEA was slightly above 0.05 ($=0.062$); and SRMR was smaller than 0.08 ($=0.067$). All factor loadings were well above the cut-off value of 0.500. Before deciding whether the model should be further improved based on the modification indices, the model was verified by treating the observed variables as ordinal. Accordingly, Robust Weighted Least Squares (WLSMV) was used as the estimation method. The result was similar to the first approach, with fit indices displaying better model fit: $\chi^2(558)=686.759$, $p<.000/.001$; CFI = 0.985, RMSEA=0.034; SRMR=0.074. All factor loadings once again revealed adequate loadings ($>.500$). Therefore, no further adaptation with respect to freeing error covariances was adopted in order to keep the parsimony of the final model. Regarding composite reliability, ω revealed good reliability for all nine dimensions, ranging from 0.76 to 0.93. The results are presented in Table A1 in Appendix C.

Further, we tested a second-order factor structure, i.e., whether the three latent variables confirmed in the previous step could be explained by a higher-order factor. In other words, we examined if transparency, LMS quality, and technical support could be subsumed under the higher order factor of existence; relevance of learning content, competence support, and autonomy under growth; and instructor's facilitation, peer collaboration, and sense of belonging under relatedness. The model fit indices were $\chi^2(24)=124.916$, $p<.001$; CFI = 0.857 ($<.09$), RMSEA=0.146 ($>.05$); SRMR=0.090 ($>.08$). This showed a lack of fit as all the indices did not satisfy the cut-off value. The factor loadings indicated that all first-order variables loaded adequately to their respective higher order factor, except autonomy (0.322). Table A2 the Appendix C presents the factor loadings resulting from the analysis.

Thus, in response to RQ1, the nine dimensions of adults' learning needs could be distinguished from one another. Accordingly, most first order variables could be explained by their respective higher order factor, except autonomy. A close examination of model fit indices suggested that a structural model with higher-order factor displayed poorer fit than the model with only first-order factors. Thus, the latter, which was more parsimonious was selected, i.e., treating the nine dimensions of learning needs as separate from one another.

RQ2. Adult learning needs as a function of motivation orientations, controlling for the effect of program type, age groups, gender, education levels, and ICT skills?

Descriptive statistics of the dependent variables

The descriptive statistics of the dependent variables are shown in Table 4 based on the factor scores obtained for each participant after the CFA. At the univariate level, the distributions of all the nine variables displayed no extremes, with none was skewed or lacking variability.

Table 4: Descriptive statistics of nine dimensions of learning needs

Dependent variables: The needs of ...	Mean	SD	Skewness	Kurtosis
Transparency	-0.014	0.533	0.108	-0.733
LMS quality	-0.048	0.688	-0.363	-0.596
Technical support	-0.025	0.649	-0.115	-0.699
Relevance of learning content	-0.024	0.703	0.059	-0.775
Competence support	-0.005	0.601	0.315	-0.539
Autonomy	-0.036	0.615	-0.336	-0.611
Instructor's facilitation	-0.002	0.753	0.129	0.216
Peer collaboration	-0.013	0.792	-0.166	0.190
Sense of belonging	-0.012	0.867	-0.091	0.003

Examining the possible effects of CVO and program categories

As learners were attending different courses at different CVOs, it was necessary to examine the possible effects of CVOs and program categories. If the effect proved to be significant, then the CVOs and/or program categories should be included in the model as covariates next to socio-demographics. The results from multivariate analyses in Table A3 and Table A4 in Appendix C indicated that neither CVOs nor program categories had an effect on learners' perceptions of different dimensions of learning needs. Thus, the two categorical variables were not included in subsequent analyses.

Multivariate analysis of covariance examining effect of motivational orientations on learning needs, controlling for age groups, gender, educational levels, and ICT skills

As there were nine dependent variables, we included age groups, gender, education levels, and ICT skills as the covariates, and motivational orientation, including three categories, namely diploma-oriented, knowledge-oriented, and both diploma and knowledge oriented as the independent variables.

Checking assumptions

Box's M test, $F(90,15495)=0.932$, $p=0.661$, was not significant, suggesting that the assumption of equal covariance matrices had been met. Levene's univariate tests of homogeneity of variances presented in Table A5 in Appendix C revealed that all assumptions were upheld.

The results of the multivariate tests of group mean differences are presented in Table 5. Accordingly, as for motivational orientations, Wilks' Lambda=0.848, $F(18,306)=1.456$, $p=0.105$. Thus, the means of the nine dimensions of learning needs did not differ significantly among the three motivational orientations. Regarding the control variables, the results from the multivariate analyses showed that age groups, educational levels, and ICT skills had a significant effect on learners' perceptions of their learning needs.

Table 5: Results from multivariate tests upon the effect of motivational orientations

Effect		Value	F	Hypothesis df	Error df	p-value
Age group	Pillai's Trace	0.193	4.062	9	153	0.000
	Wilks' Lambda	0.807	4.062	9	153	0.000
Gender	Pillai's Trace	0.059	1.073	9	153	0.386
	Wilks' Lambda	0.941	1.073	9	153	0.386
Educational levels	Pillai's Trace	0.199	4.219	9	153	0.000
	Wilks' Lambda	0.801	4.219	9	153	0.000
ICT skills	Pillai's Trace	0.103	1.954	9	153	0.048
	Wilks' Lambda	0.897	1.954	9	153	0.048
Motivational orientations	Pillai's Trace	0.156	1.447	18	308	0.108
	Wilks' Lambda	0.848	1.456	18	306	0.105

Univariate tests

When the result at the multivariate level was significant, it was more informative to look at the univariate tests in order to examine the nuanced differences in the perceptions of learning needs among learners of different profiles. Accordingly, significant differences were found among different age groups regarding the needs of transparency ($F(1,161)=4.541$, $p=0.035$), LMS quality ($F(1,161)=15.73$, $p<0.001$), technical support ($F(1,161)=11.892$, $p=0.001$), and sense of belonging ($F(1,161)=4.555$, $p=0.034$). Results indicated that the older the learners, the greater the degree of need expressed for all four categories.

Regarding educational levels, significant differences were found for the needs of competence support ($F(1,161)=12.344$, $p=0.001$), instructor's facilitation ($F(1,161)=21.344$, $p<0.001$), peer collaboration ($F(1,161)=13.583$, $p<0.001$), and sense of belonging ($F(1,161)=17.854$, $p<0.001$). Further analyses revealed that the higher the degree the learners had obtained, the lower the degree of need was expressed concerning competence support, instructors' facilitation, peer collaboration, and sense of belonging.

Learners with different ICT skill levels also differed in their perceptions of learning needs regarding competence support ($F(1,161)=4.881$, $p=0.029$) and sense of belonging ($F(1,161)=8.105$, $p=0.005$). Interestingly, learners with higher ICT skills revealed lower mean scores for competence support and sense of belonging. The results are presented in Table 6.

Table 6: Parameter estimates for the independent variables

Dependent Variable	Parameter	B	SE	p-value
Transparency	Intercept	-0.197	0.303	0.516
	Age groups	0.085	0.040	0.035
	Gender	0.197	0.091	0.032
	Education levels	0.049	0.043	0.253
	ICT skills	-0.091	0.071	0.202
	Motivation orientation 1 ^a	-0.251	0.120	0.037
	Motivation orientation 2 ^a	-0.427	0.126	0.001
LMS quality	Intercept	-0.661	0.384	0.087
	Age groups	0.200	0.050	<0.001
	Gender	0.267	0.115	0.022
	Education levels	0.086	0.055	0.116
	ICT skills	-0.082	0.090	0.363
	Motivation orientation 1 ^a	-0.303	0.152	0.047
	Motivation orientation 2 ^a	-0.563	0.160	0.001
Technical support	Intercept	-0.282	0.370	0.447
	Age groups	0.167	0.049	0.001
	Gender	0.218	0.111	0.051
	Education levels	-0.056	0.053	0.287
	ICT skills	-0.058	0.087	0.506
	Motivation orientation 1 ^a	-0.175	0.146	0.233
	Motivation orientation 2 ^a	-0.339	0.154	0.029
Relevance of learning content	Intercept	0.198	0.420	0.637
	Age groups	0.039	0.055	0.478
	Gender	0.192	0.126	0.129
	Education levels	-0.074	0.060	0.216
	ICT skills	-0.026	0.099	0.796

Dependent Variable	Parameter	B	SE	p-value
	Motivation orientation 1 ^a	-0.391	0.166	0.019
	Motivation orientation 2 ^a	-0.431	0.175	0.015
Competence support	Intercept	0.664	0.334	0.049
	Age groups	0.042	0.044	0.343
	Gender	0.190	0.100	0.060
	<i>Education levels</i>	-0.167	0.048	0.001
	<i>ICT skills</i>	-0.173	0.078	0.029
	Motivation orientation 1 ^a	-0.255	0.132	0.055
	Motivation orientation 2 ^a	-0.300	0.139	0.033
Autonomy	Intercept	-0.456	0.365	0.212
	Age groups	0.073	0.048	0.131
	Gender	0.061	0.109	0.577
	Education levels	0.043	0.052	0.407
	ICT skills	0.126	0.086	0.143
	Motivation orientation 1 ^a	-0.254	0.144	0.080
	Motivation orientation 2 ^a	-0.366	0.152	0.017
Instructor's facilitation	Intercept	0.749	0.399	0.063
	Age groups	0.095	0.052	0.072
	Gender	0.169	0.120	0.160
	<i>Education levels</i>	-0.262	0.057	0.000
	ICT skills	-0.118	0.094	0.209
	Motivation orientation 1 ^a	-0.290	0.158	0.068
	Motivation orientation 2 ^a	-0.353	0.167	0.035
Peer collaboration	Intercept	1.196	0.432	0.006
	Age groups	-0.037	0.057	0.516
	Gender	-0.004	0.130	0.975
	<i>Education levels</i>	-0.227	0.061	0.000
	ICT skills	-0.192	0.102	0.060
	Motivation orientation 1 ^a	0.029	0.171	0.864
	Motivation orientation 2 ^a	-0.045	0.180	0.801
Sense of belonging	Intercept	1.880	0.466	0.000
	<i>Age groups</i>	-0.131	0.061	0.034
	Gender	-0.017	0.140	0.902
	<i>Education levels</i>	-0.280	0.066	0.000
	<i>ICT skills</i>	-0.312	0.109	0.005
	Motivation orientation 1 ^a	-0.063	0.184	0.734
	Motivation orientation 2 ^a	-0.022	0.194	0.909

^areference category: motivation orientation 3

5. Discussions

The paper aimed to validate the instrument measuring adult learners' needs in a blended learning environment. Subsequently, multivariate analyses of covariance were conducted to provide insights into the learning needs of adult learners of different socio-demographic and motivational profiles. The findings are discussed as follows.

First, based on a review of the literature on adult learning, the technical acceptance model, and blended learning, a theoretical framework was synthesized to guide the learning need instrument development. This resulted in the construction of nine dimensions of learning needs captured by 36 items. The CFA results showed that the instrument displayed adequate fit based on satisfactory global fit indices and factor loadings obtained. Therefore, the instrument could be employed by adult education practitioners and researchers to evaluate the learning needs of adults in order to better cater for their needs and to improve the blended course design and practices accordingly. The second-order factor analysis suggested that the higher order factors, namely existence, relatedness, and growth needed further confirmation from empirical studies and theoretical elaboration before such clustering could be established. The lack of fit could be attributed to the low loading of autonomy in the Growth dimension. It could be hypothesized that being and becoming an autonomous learner can be a preference, thus differing from learner to learner. On the other hand, the need for relevance and instructor's competence support was shared by most learners, which was confirmed in a recent study (Chambers and Whitfield, 2025). Thus, in terms of content, the need for autonomy could be slightly deviated from the other two. It was recommended that future research further validate autonomy and decide how this dimension should be interpreted under the overarching needs theory. More broadly, the weak loading of autonomy on the Growth factor raises a theoretical question about whether autonomy, as conceptualized in SDT and andragogy, constitutes a learner *need* or rather an instructional *outcome* — a distinction with significant implications for OBL design that warrants further theoretical debate.

Second, learners of three different motivational profiles, namely diploma-oriented, skills-and knowledge-oriented, and both, did not display significant differences in their learning needs. It was expected that knowledge and skills-oriented learners would express a higher degree of needs in terms of growth goals (Pintrich and Zusho, 2002), including competence support from the instructors, the relevance of the learning content, and autonomy. This, on the one hand, was somewhat contradictory to motivational theories. At the methodological level, this non-significant result may partly reflect the unequal distribution of motivational profiles in the sample, with diploma-oriented learners constituting the largest group ($n=87$), which may have limited statistical power to detect more nuanced differences. At the theoretical level, this finding invites a reconsideration of whether motivational orientation, as a single categorical variable, is sufficiently sensitive to capture the complexity of adult learning needs in OBL contexts. However, recent literature suggests that motivation is a relative construct rather than a fixed trait and is subject to changes due to environmental and social contexts. According to Reeve (2006), there were a number of learning conditions that may facilitate or impede an individual's motivation. Among these were motivation styles of the instructors, the affordances made to the learners, the use of external events and the expectations placed on the learners (Reeve, 2006) and perceived task value (Fryer, Bovee, and Nakao, 2014). Therefore, a possible explanation is that there has been a fluctuation in the evolution of adults' motivational orientations that could not be directly observed or noticed by the learners themselves or by the questioning method. Therefore, it was advised that future research may employ a longitudinal approach to depict learners' motivation and capture the relationships between those changes in motivation and the learning needs during the course. For example, it is hypothesized that a learner starting with an extrinsic motivation will transfer to intrinsic motivation if he or she receive substantial support in terms of learning strategies and find more relevance in the learning content.

Thirdly, in line with existing literature on adult learning, the results confirmed that adults' socio-demographics still played an important role regarding learning needs identification. More specifically, we found that age groups, ICT skills, educational levels were variables that could help explain differences in adults' learning needs. For example, learners with higher educational levels tended to be more independent learners, i.e., expressing a lesser degree of needs regarding competence support, facilitation from the instructors, peer collaboration, sense of belonging, and more autonomy in learning. This finding suggested that highly educated adult learners tended to prefer a more independent approach to learning when they were in a heterogeneous class. This pattern may be explained by the fact that highly educated adults enter OBL with well-established self-regulatory strategies and a stronger sense of academic identity, rendering external scaffolding less necessary (Knowles, 1984). From an ERG theory perspective, their growth needs are more likely to be pursued independently, whereas relatedness and existence needs recede in salience as prior academic experience reduces uncertainty about the learning environment. Despite supporting learners to be more independent and autonomous is an

important role of education, encouraging learners to involve in collaborative learning is of equal importance as research showed that collaboration in blended learning results in enriched and highly metacognitive and cognitive gains (Mebane et al., 2008). To motivate adults to invest effort in the collaborative learning, Vanslambrouck et al. (2018) showed that the instructors should make explicit the task values or the benefits of engaging in such interactions as a cost-benefits mechanism might come into play given that adults need to combine their learning with life and work responsibilities.

On the contrary, older learners expressed higher existential needs including the transparency, LMS quality, technical support as well relatedness needs, i.e., sense of belonging (Ke and Kwak, 2013). Thus, in catering for the learning needs of adults of different profiles participating in the same program, the instructors should pay more attention to these differences. Differentiated instructions with various choices of learning activities and learning paths aimed at reaching the same learning objectives could be a possible approach. In so doing, the learners would be more motivated to stay the course through.

6. Limitations

First, due to the nature of the data collection method, i.e., convenience sampling, it should be acknowledged that the findings cannot not be generalized. Second, the sample was skewed in terms of program type, with pedagogical programs (Specific Teacher Training and Freelance Coaching) accounting for over 60% of participants, which may have introduced program-specific patterns in learning need profiles that are not representative of the broader adult learner population. Third, the gender distribution was unequal, with female participants constituting approximately 69% of the sample, which may reflect the nature of certain CVO programs (e.g., childcare, teaching) rather than the adult learner population at large. These characteristics should be considered when interpreting the findings and when applying the validated instrument to other adult education contexts. Moreover, the researchers were not successful in recruiting learners of different motivational profiles as expected, hence a clear picture of how adult learners differed in their learning needs as a function of motivation could not be presented. Nevertheless, initial findings from the study revealed that motivational orientations were not significantly related to learning needs identification. Therefore, it is encouraged that future research might adopt a more valid approach to track the motivational fluctuation of adult learners during the course so that helpful implications can be offered to instructors and instructional designers. Finally, as the data were collected at the end of the program, the identification of learning needs might have been somewhat influenced. Although we believe that at this point the learners could more accurately specify their needs after a process of engaging with and experiencing the challenges and difficulties in their blended courses, it could be worthwhile to track adults' learning needs from the beginning so that prompt support could be provided. That being said, monitoring adults' learning needs "on-the-go" is a strongly recommended study to examine what specific needs are urgent in which specific stage of the learning trajectory.

7. Conclusions

In this paper, an instrument measuring adults' learning needs in a blended learning environment was validated. Employing the ERG theory, we elaborated on nine learning needs dimensions underpinned by both learning theories and empirical findings. The instrument revealed satisfactory model fit and factor loadings; therefore, it could be used as a valid tool to support instructors and adult education institutions in determining their learners' needs in order to better adapt the curriculum and instructional design when necessary. Significant differences in learning needs identified among learners of different profiles also call for more attention when a needs-based curriculum is implemented. More specifically, learners of lower educational levels, older learners, and those with lower ICT skills should receive more support in the existence and relatedness need categories. The findings that higher-educated learners expressed a lesser extent of competence support in the dimension of Growth needs and Relatedness help explained why it is difficult to engage this group in a heterogeneous classroom and call for more deliberate strategies aimed at engaging them in collaborative learning within a blended learning environment.

Beyond these immediate findings, this study offered critical insights for the broader landscape of adult education. For adult education policy, these results argued against "one-size-fits-all" funding and institutional frameworks, highlighting the necessity of macro-level policies that allocate dedicated resources for digital equity, socio-emotional onboarding, and technological support for vulnerable learner demographics. In terms of OBL design, our framework provided instructional designers with an empirical blueprint to deliberately structure OBL environments—ensuring that structural coherence and technical support (Existence), collaborative learning and sense of belonging (Relatedness), and intellectual challenges (Growth) were

intentionally mapped into the curriculum. For future learner-support systems, this validated instrument can serve as an early-stage diagnostic tool within the learning management systems (LMS) to proactively deploy automated, adaptive interventions tailored to learners' unique socio-demographic and motivational profiles.

Taken together, these findings not only confirmed the utility of the ERG framework as a theoretical lens for adult learning needs but also challenge the assumption that motivational orientation alone was sufficient to predict differentiated need profiles in OBL — an assumption prevalent in SDT-based research. In this sense, the present study opened new avenues for theoretical debate regarding the interplay between motivation, prior experience, and learning needs in OBL adult education contexts. Nevertheless, it should be acknowledged that the reliance on convenience sampling and the unequal distribution of motivational profiles in the present study limited the generalizability of these findings. Future replication with more diverse and larger samples, combined with longitudinal designs, would strengthen the validity of the instrument, and yield more nuanced insights into the dynamic relationship between adult learners' motivational orientations and their learning needs across different stages of a blended program.

Acknowledgements

The authors would like to thank all the students from the eight Centrum voor Volwassenenonderwijs (CVO) in Flanders (Belgium), who have participated in the study, despite the challenges in time.

The use of AI: The authors reviewed and edited the content as required, and they take full responsibility for the publication. The authors declared no AI-tools were used to write, generate, or revise the text.

Ethics approval and consent to participate: The study did not entail interventions on adults or the implemented instructions, hence was not subject to an ethical approval. Before proceeding to the online questionnaire, participants signed the consent form which indicated their voluntary participation in the study and that the data collected to be published anonymously. The study was conducted in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Consent for publication: Before proceeding to the online questionnaire, participants signed the consent form which indicated their voluntary participation in the study and that the data collected to be published anonymously.

Availability of data and material: The datasets are available from the corresponding author on reasonable request, which is subject to an approval from the ethical committee.

Competing interests: The authors declare that they have no competing interests.

Funding: The research was funded by the SRP 28 (Vrije Universiteit Brussel).

Authors' contributions: All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by AD, AFD, CC, TW, MDG and CZ. The first draft of the manuscript was written by AD and KD, TNPP, MV and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

- Abela, J. (2009) 'Adult learning theories and medical education: a review', *Malta Medical Journal*, 21(1), pp. 11–18.
- Addae, D., Abakah, E. and Amuzu, D. (2025) 'Adults who learn online: Exploring the online learning readiness of nontraditional undergraduate students in a Ghanaian university', *International Journal of Educational Research Open*, 8, p. 100407.
- Alderfer, C. P. (1969) 'An empirical test of a new theory of human needs', *Organizational behavior and human performance*, 4(2), pp. 142–175.
- Alderfer, C.P. (1972). *Existence, relatedness, and growth: Human needs in organizational settings*. New York: The free press.
- Arbaugh, J. Ben, et al. (2008). "Developing a community of inquiry instrument: Testing a measure of the community of inquiry framework using a multi-institutional sample." *The internet and higher education* 11 (3-4), pp.133-136.
- Bandura, A. (1986) 'The explanatory and predictive scope of self-efficacy theory', *Journal of social and clinical psychology*, 4(3), pp. 359–373.
- Bangert, A. W. (2004) 'The seven principles of good practice: A framework for evaluating on-line teaching', *The Internet and Higher Education*, 7(3), pp. 217–232.
- Boshier, R. (1991) 'Psychometric properties of the alternative form of the education participation scale', *Adult Education Quarterly*, 41(3), pp. 150–167.

- Botha, J. A. and Masenge, A. (2022) 'The adult learner self-directedness scale: Validity and reliability assessment', *SA Journal of Industrial Psychology*, 48, p. a1926.
- Byrne, B. M. (1989) *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York: Springer-Verlag.
- Cercone, K. (2008) 'Characteristics of adult learners with implications for online learning design', *AACE journal*, 16(2), pp. 137–159.
- Chambers, M. A. and Whitfield, C. (2025) 'Factors influencing postgraduate nursing students' engagement with online learning in higher education: A mixed methods literature review', *Nurse Education in Practice*, p. 104570.
- Davis, F. D. (1989) 'Perceived usefulness, perceived ease of use, and user acceptance of information technology', *MIS quarterly*, pp. 319–340.
- De Greef, M., Segers, M. and Verté, D. (2012) 'Understanding the effects of training programs for vulnerable adults on social inclusion as part of continuing education', *Studies in Continuing Education*, 34(3), pp. 357–380.
- Deci, E. L., and Ryan R. M.. 'The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior.' *Psychological inquiry*. 11 (4), pp. 227-268.
- Diep, A. N., Zhu, C., Cocquyt, C., De Greef, M., Vo, M. H. and Vanwing, T. (2019) 'Adult learners' needs in online and blended learning', *Australian Journal of Adult Learning*, 59(2), pp. 223–253.
- Diep, A. N., Zhu, C., Struyven, K. and Blicq, Y. (2017) 'Who or what contributes to student satisfaction in different blended learning modalities?', *British Journal of Educational Technology*, 48(2), pp. 473–489.
- Diep, N. A., Cocquyt, C., Zhu, C. and Vanwing, T. (2016) 'Predicting adult learners' online participation: Effects of altruism, performance expectancy, and social capital', *Computers and Education*, 101, pp. 84–101.
- Diep, N. A., Cocquyt, C., Zhu, C. and Vanwing, T. (2017) 'Online interaction quality among adult learners: The role of sense of belonging and perceived learning benefits', *Turkish Online Journal of Educational Technology*, 16(2), pp. 71–78.
- Dixon, C. S. (2014). "The three E's of online discussion." *Quarterly Review of Distance Education* 15 (1), pp. 1-8.
- Drago, W. and Peltier, J. (2004) 'The effects of class size on effectiveness of online courses', *Management Research News*, 27(10), pp. 27–36.
- Fryer, L. K., Bovee, H. N. and Nakao, K. (2014) 'E-learning: Reasons students in language learning courses don't want to', *Computers and Education*, 74, pp. 26–36.
- Garrison, D. R., Anderson, T. and Archer, W. (2000) 'Critical inquiry in a text-based environment: Computer conferencing in higher education', *The internet and higher education*, 2(2-3), pp. 1–19.
- Gaver, W. (1991) 'Technology affordances', in S. Robertson and G. Olson (eds.) *SIGCHI conference on human factors in computing systems*. ACM, pp. 79–84.
- Gibson, J. (1979) *The ecological approach to visual perception*. Houghton Mifflin.
- Graham, C. R. (2006) 'Blended learning systems: Definition, current trends, and future directions', in C. J. Bonk and C. R. Graham (eds.) *Handbook of blended learning: Global perspectives, local designs*. San Francisco: Pfeiffer Publishing, pp. 3–21.
- Hair, J. F., Black, W. C., Babin, B. J. and Anderson, R. E. (2010) *Multivariate Data Analysis*. 7th edn. Pearson Prentice Hall.
- Hsieh, Y. H. and Tsai, C. C. (2012) 'The effect of moderator's facilitative strategies on online synchronous discussions', *Computers in Human Behavior*, 28(5), pp. 1708–1716.
- Ke, F. and Kwak, D. (2013) 'Online learning across ethnicity and age: A study on learning interaction participation, perception, and learning satisfaction', *Computers and education*, 61, pp. 43–51.
- Knowles, Malcolm S., Elwood Holton, and Richard Swanson. "The adult learner: the definitive classic in adult education and human resource development (6th)." *Burlington, MA: Elsevier*(2005).
- Knowles, M. S. (1984) *Andragogy in action: Applying principles of adult learning*. San Francisco: Jossey-Bass.
- Kollock, P. (1999) 'The economies of online cooperation', *Communities in cyberspace*, p. 220.
- Koul, R. B., Fisher, D. L., and Shaw, T. (2011). An application of the TROFLEI in secondary-school science classes in New Zealand'. *Research in Science & Technological Education*, 29(2), pp. 147–167
- Laurillard, D. (2012) *Teaching as a design science. building pedagogical patterns for learning and technology*. Routledge.
- Lee, R. M., Draper, M. and Lee, S. (2001) 'Social connectedness, dysfunctional interpersonal behaviors, and psychological distress: Testing a mediator model', *Journal of Counseling Psychology*, 48(3), p. 310.
- Locke, E. A. and Latham, G. P. (2002) 'Building a practically useful theory of goal setting and task motivation: A 35-year odyssey', *American psychologist*, 57(9), p. 705.
- Loock, V. S., Fleischer, J., Scheunemann, A., Froese, L., Teich, K. and Wirth, J. (2022) 'Narrowing down dimensions of e-learning readiness in continuing vocational education—perspectives from the adult learner', *Frontiers in Psychology*, 13, p. 1033524.
- Mebane, M., et al. (2008). 'Evaluation of the efficacy of affective education online training in promoting academic and professional learning and social capital', *Intl. Journal of Human-Computer Interaction*, 24(1), pp. 68-86.
- Mertens, B., De Maeyer, S. and Donche, V. (2024) 'Exploring learner profiles among low-educated adults in second-chance education: individual differences in quantity and quality of learning motivation and learning strategies', *European Journal of Psychology of Education*, 39(4), pp. 3963–3987.
- Mezirow, J. (2000) *Learning as Transformation: Critical Perspectives on a Theory in Progress*. Jossey-Bass Publishers.
- Milheim, K. L. (2012) 'Towards a better experience: Examining student needs in the online classroom through Maslow's hierarchy of needs model', *Journal of Online Learning and Teaching*, 8(2), p. 159.

- Niemiec, C. P. and Ryan, R. M. (2009) 'Autonomy, competence, and relatedness in the classroom Applying self-determination theory to educational practice', *Theory and research in Education*, 7(2), pp. 133–144.
- Oliver, M. (2013) 'Learning technology: theorising the tools we study', *British Journal of Educational Technology*, 44, pp. 31–43.
- Pearce, S. D. (1995) 'Needs assessment: constructing tacit knowledge from practice', *International Journal of Lifelong Education* 14.5, pp. 405–419.
- Peltier, J. W., John A. S., and William D. (2007). "The interdependence of the factors influencing the perceived quality of the online learning experience: A causal model." *Journal of Marketing Education*. 29 (2), pp. 140-153.
- Pintrich, P. R. and Zusho, A. (2007) 'Student motivation and self-regulated learning in the college classroom', in *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Springer, Dordrecht, pp. 731–810.
- Poll, K., Widen, J. and Weller, S. (2014) 'Six instructional best practices for online engagement and retention', *Journal of Online Doctoral Education*, 1(1).
- Prior, D. D., Mazanov, J., Meacheam, D., Heaslip, G. and Hanson, J. (2016) 'Attitude, digital literacy and self efficacy: Flow-on effects for online learning behavior', *The Internet and Higher Education*, 29, pp. 91–97.
- Reeve, J. (2002) 'Self-Determination Theory applied to educational settings', in E. Deci and R. Ryan (eds.) *Handbook of Self-Determination Research*. University of Rochester Press, pp. 183–203.
- Reeve, J. 'Teachers as facilitators: What autonomy-supportive teachers do and why their students benefit.' *The elementary school journal*. 106 (3). pp 225-236.
- Reeves, T. C., Herrington, J. and Oliver, R. (2002) 'Authentic activities and online learning', *Annual Conference Proceedings of Higher Education Research and Development Society of Australasia*. Perth, Australia.
- R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ross-Gordon, J. M. (2003) 'Adult learners in the classroom', *New Directions for Student Services*, 2003(102), pp. 43–52.
- Rovai, A. P. (2002) 'Building sense of community at a distance', *The International Review of Research in Open and Distributed Learning*, 3(1).
- Ryan, R. M. and Deci, E. L. (2000) 'Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being', *American psychologist*, 55(1), p. 68.
- Sava, S. (2012) *Needs Analysis and Programme Planning in Adult Education*. 1st edn. Verlag Barbara Budrich. doi: 10.2307/j.ctvbkjvs2.
- van Deursen, A. J. A. M., Helsper, E. J. and Eynon, R. (2014) *Measuring digital skills: From digital skills to tangible outcomes* project report. London School of Economics and Political Science.
- Venkatesh, V., Thong, J. Y. and Xu, X. (2012) 'Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology', *MIS quarterly*, 36(1), pp. 157–178.
- Vanslambrouck, S., et al. 'Students' motivation and subjective task value of participating in online and blended learning environments.' *The Internet and Higher Education*. 36, pp 33-40.
- Walker, S. L. and Fraser, B. J. (2005) 'Development and validation of an instrument for assessing distance education learning environments in higher education: The Distance Education Learning Environments Survey (DELES)', *Learning Environments Research*, 8(3), pp. 289–308.
- Wolf, E. J., Harrington, K. M., Clark, S. L. and Miller, M. W. (2013) 'Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety', *Educational and Psychological Measurement*, 73(6), pp. 913–934.
- Woo, Y. and Reeves, T. C. (2007) 'Meaningful interaction in web-based learning: A social constructivist interpretation', *The Internet and higher education*, 10(1), pp. 15–25.
- Yang, Y. and Lay, Y. F. (2024) 'What are the roles of positive psychological construct in blended learning contexts? Integrating academic buoyancy into the Community of Inquiry framework', *Frontiers in psychology*, 15, p. 1354156.

Appendix A

Onderzoek 'Noden en tevredenheid van cursisten in online en gecombineerde leeromgeving'

Beste heer, mevrouw,

In het kader van een onderzoeksproject aan de Universiteit X, nodigen we u uit om deel te nemen aan ons onderzoek. Het onderzoek bestudeert de noden die cursisten hebben in een gecombineerde leeromgeving.

Doel van de vragenlijst

We nodigen u uit om deel te nemen aan een vragenlijst, om informatie te verzamelen over welke noden u ervaart tijdens het studeren in een gecombineerde leeromgeving. Deze informatie is zeer waardevol om de cursistenbegeleiding te kunnen verbeteren.

Deelname aan de vragenlijst

Er wordt een vragenlijst afgenomen. Dit zal ongeveer 20 minuten van uw tijd in beslag nemen.

Verwerking en rapportering van gegevens

De verwerking en de rapportering van de gegevens gebeurt anoniem. De verzamelde gegevens worden in alle discretie en met het nodige respect behandeld. Dit houdt in dat de gegevens niet gebruikt worden voor commerciële doeleinden en anoniem verwerkt worden in wetenschappelijke publicaties zoals rapporten, artikels en boeken.

Geïnformeerde toestemming

Door akkoord te gaan, verklaart u dat u als deelnemer aan dit onderzoek:

- de uitleg over de aard van het onderzoek hebt gelezen en dat u de mogelijkheid werd geboden om bijkomende informatie te verkrijgen;
- uit vrije wil deelneemt aan dit onderzoek;
- de toestemming geeft om het onderzoeksmateriaal op vertrouwelijke wijze te bewaren, te verwerken en erover te rapporteren;
- op de hoogte bent van de mogelijkheid om uw deelname aan het onderzoek op ieder moment stop te zetten;

Men heeft mij voldoende ingelicht over dit onderzoek. Ik weet dat ik bij verdere vragen steeds contact kan opnemen via << information omitted for peer review >>

☐ Ik ga akkoord met deelname aan dit vragenlijstenonderzoek

Deel 1. Achtergrondinformatie

1. Geslacht

- ☐ Man
☐ Vrouw

2. Geboortjaar: _____

3. Nationaliteit: _____

4. Burgerlijke staat:

- ☐ Ik ben getrouwd
☐ Ik ben gescheiden
☐ Ik ben weduwe of weduwnaar
☐ Ik ben ongehuwd
☐ Ik ben alleenstaand
☐ Ik ben samenwonend

- ☐ Andere: _____
5. Welke van de volgende omschrijvingen geeft je huidige werksituatie het beste weer?
- ☐ Voltijds tewerkgesteld
 - ☐ Deeltijds tewerkgesteld
 - ☐ Werkzoekende
 - ☐ Gepensioneerd of bruggepensioneerd
 - ☐ Voltijds student
 - ☐ Huisvrouw/huisman
 - ☐ Andere: _____
6. Duid jouw hoogst behaalde diploma aan
- ☐ Geen
 - ☐ Basisonderwijs
 - ☐ Lager secundair onderwijs
 - ☐ Hoger secundair onderwijs
 - ☐ Professionele bachelor of graduaat
 - ☐ Academische bachelor of kandidaat
 - ☐ Master of licentiaat
 - ☐ Doctor
 - ☐ Andere: _____
7. In welk centrum volg je momenteel les?
Meer specifiek, welk centrum heeft jou in contact gebracht met deze vragenlijst?
- _____
8. Hoe heet de opleiding die je momenteel volgt?
- _____
9. Wanneer ben je aan deze opleiding begonnen? (jaar + semester indien mogelijk)
- _____
10. Heb je al ervaring met gecombineerd of afstandsonderwijs?
- ☐ Ja ☐ Neen
11. Hoeveel modules heb je al afgewerkt in deze opleiding? _____
12. Over het algemeen, hoe schat u uw ICT-vaardigheden in (vb. opzoeken op internet, werken met Microsoft Office, sociale media zoals Facebook en Twitter)
- ☐ Beginner ☐ Gemiddeld ☐ Gevorderd
13. Waarom heeft u zich ingeschreven voor deze opleiding of cursus? Duid de meest belangrijke reden aan.
- ☐ Ik wou graag het diploma behalen.
 - ☐ Ik wou graag mijn kennis en vaardigheden verbeteren voor mijn huidige of toekomstige job.
 - ☐ Ik wou graag nieuwe mensen leren kennen en nieuwe vrienden maken.
 - ☐ Ik wou breken met de sleur en verveling in mijn (dagelijks) leven.
 - ☐ Ik wou graag nieuwe kennis en vaardigheden opdoen.
 - ☐ Ik volg deze cursus omdat dit me voorbereid om me in te zetten voor de gemeenschap.
 - ☐ Ik volg deze cursus omdat ik moet van mijn werkgever of om te kunnen genieten van sociale voordelen.
14. Welke gemiddelde score behaalde je reeds gedurende deze opleiding?
- Opgelet, deze score is uitgedrukt in een score op 20. Indien je een andere score kreeg, gelieve “andere” aan te duiden en in te vullen. Indien je nog geen punten kreeg, gelieve **vraag 15** te beantwoorden.
- ☐ ≤ 10
 - ☐ 11-12
 - ☐ 13-14
 - ☐ 15-16
 - ☐ 17-18
 - ☐ 19-20
 - ☐ Andere: _____

Niet van toepassing → ga naar vraag 15

15. Vul deze vraag enkel in indien je vraag 14 niet kon beantwoorden wegens ‘niet van toepassing’.

Hoe verwacht je gemiddeld te scoren doorheen deze opleiding? Opgelet, deze score is uitgedrukt in een score op 20. Indien je centrum met een ander scoresysteem werkt, gelieve “andere” aan te duiden en in te vullen.

- ☐ ≤ 10
- ☐ 11-12
- ☐ 13-14
- ☐ 15-16
- ☐ 17-18
- ☐ 19-20
- ☐ Andere: _____

Deel 2. 1. Duidelijke verwachtingen

Welke noden ervaar je in verband met het creëren van duidelijke verwachtingen tijdens de cursus? We willen graag weten in welke mate je nood hebt aan duidelijkheid over wat van jou verwacht wordt tijdens de cursus. Duid aan in welke mate je akkoord gaat met de volgende stellingen.

(1) helemaal niet akkoord, (2) niet akkoord, (3) neutraal, (4) akkoord, tot (5) helemaal akkoord

	1	2	3	4	5
Ik heb nood aan een duidelijk idee over wat er van mij verwacht wordt in deze cursus.					
Ik vind het nodig dat ik gemakkelijk te weten kan komen welke kwaliteit de docent van mijn werk verwacht.					
Ik vind het nodig dat de docent duidelijke instructies geeft over wat hij/zij van de studenten verwacht vanaf het begin van de cursus.					
Ik vind het nodig dat ik gemakkelijk kan ontdekken wat er van mij verwacht wordt.					

Deel 2.2. Het leerplatform

Welke behoeftes heb je tijdens het werken met het online leerplatform (vb. Moodle of Smartschool)?

Duid aan in welke mate je tevreden bent over de volgende zaken.

(1) helemaal niet akkoord, (2) niet akkoord, (3) neutraal, (4) akkoord, tot (5) helemaal akkoord

	1	2	3	4	5
Ik vind het nodig dat ik de informatie op het leerplatform gemakkelijk kan terugvinden).					
Ik vind het nodig dat de aanwijzingen en navigatie op het leerplatform duidelijk zijn).					
Ik heb behoefte aan overzichtelijke informatie en inhoud op het leerplatform.					
Ik heb behoefte aan een goede gebruikersinterface op het leerplatform.					

Deel 2. 3. Technische hulp

Wat vind je nodig op vlak van technische ondersteuning? Duid aan in welke mate je nood hebt aan de volgende zaken.

(1) helemaal niet akkoord, (2) niet akkoord, (3) neutraal, (4) akkoord, tot (5) helemaal akkoord

	1	2	3	4	5
Ik vind het nodig dat ik zo weinig mogelijk technische problemen heb tijdens de cursus.					
Ik vind het nodig dat ik gemakkelijk technische hulp kan vragen.					
Ik vind het nodig dat ik weet waar ik hulp kan vragen in geval van technische problemen.					
Ik vind het nodig dat ik tijdig hulp krijg bij technische problemen.					

Deel 2. 4. Relevantie

In welke mate heb je nood aan relevante onderwerpen tijdens de cursus? Relevante onderwerpen zijn thema's die jou interesseren.

Duid aan in welke mate je akkoord gaat met de volgende stellingen.

(1) helemaal niet akkoord, (2) niet akkoord, (3) neutraal, (4) akkoord, tot (5) helemaal akkoord

	1	2	3	4	5
Ik vind het nodig dat ik mijn kennis en vaardigheden in het dagelijks leven (vb. thuis, op het werk) kan inzetten.					
Ik vind het nodig dat ik onderwerpen kan studeren die mij interesseren.					
Ik vind het nodig dat ik mijn studies kan linken aan mijn dagelijks leven.					
Ik vind het nodig dat ik mijn persoonlijke ervaringen kan toepassen in de cursus.					

Deel 2. 5. Leerondersteuning

In welke mate heb je behoefte aan ondersteuning van de docent?

Duid aan in welke mate je akkoord gaat met de volgende stellingen.

(1) helemaal niet akkoord, (2) niet akkoord, (3) neutraal, (4) akkoord, tot (5) helemaal akkoord

	1	2	3	4	5
Ik vind het nodig dat de docent mij voldoende feedback geeft.					
Ik heb nood aan tijdige begeleiding van de docent.					
Ik vind het nodig dat de docent mij opdrachten geeft die op mijn capaciteiten zijn afgestemd.					
Ik vind het nodig dat de docent voldoende mogelijkheden voorziet om theorie om te zetten in de praktijk.					

Deel 2.6. Zelfstandigheid

In welke mate vind je het nodig dat je zelfstandig en autonoom kan studeren? Duid jouw mening hierover aan.

(2) helemaal niet akkoord, (2) niet akkoord, (3) neutraal, (4) akkoord, tot (5) helemaal akkoord

	1	2	3	4	5
Ik vind het nodig om te studeren wanneer het mij past.					
Ik vind het nodig dat ik controle heb over mijn eigen leerproces.					
Ik vind het nodig dat ik een belangrijke rol in mijn eigen leerproces speel.					
Ik heb behoefte om mijn leerproces op mijn eigen manier aan te pakken.					

Deel 2.7. Betrokkenheid van de docent

In welke mate vind je de betrokkenheid van de docent tijdens de online leermomenten nodig ? Duid jouw mening hierover aan.

(1) helemaal niet akkoord, (2) niet akkoord, (3) neutraal, (4) akkoord, tot (5) helemaal akkoord

	1	2	3	4	5
Ik vind het nodig dat de docent alle cursisten aanmoedigt om betrokken te zijn bij en om productief deel te nemen aan online gesprekken.					
Ik vind het nodig dat de docent alle cursisten aanstuurt om gefocust te blijven op de online opdracht.					
Ik vind het nodig dat de docent alle cursisten aanmoedigt om online nieuwe concepten in deze cursus te ontdekken.					

	1	2	3	4	5
Ik heb nood aan ondersteuning die het online groepsgevoel tussen de cursisten versterkt.					

Deel 2.8. Samenwerking

Welke noden ervaar je tijdens het samenwerken met medecursisten? Samenwerking kan verwijzen naar groepswerk of groepspresentatie.

(1) helemaal niet akkoord, (2) niet akkoord, (3) neutraal, (4) akkoord, tot (5) helemaal akkoord

	1	2	3	4	5
Ik vind het nodig dat ik samen kan werken met medecursisten aan taken of projecten.					
Ik heb behoefte om mijn ideeën met mijn medecursisten te bespreken.					
Ik heb nood aan feedback van mijn medecursisten.					
Ik vind het nodig dat ik toegang heb tot het werk van medecursisten (vb. essays, blogs, wiki's, ingediende opdrachten, mondelinge presentaties).					

Deel 2.9. Gemeenschapsgevoel

In welke mate vind je het nodig om je verbonden te voelen met jouw medecursisten? Duid jouw mening aan.

(1) helemaal niet akkoord, (2) niet akkoord, (3) neutraal, (4) akkoord, tot (5) helemaal akkoord

	1	2	3	4	5
Ik vind het nodig dat alle cursisten voor elkaar zorgen tijdens deze cursus.					
Ik vind het nodig dat ik me verbonden voel met mijn medecursisten.					
Ik vind het nodig dat ik op mijn medecursisten kan rekenen.					
Ik heb nood aan hulp van mijn medecursisten.					

SLOT

Hartelijk bedankt voor jouw medewerking aan dit onderzoek.

Tenslotte stellen we jou nog een laatste vraag. Werk je graag mee aan vervolgonderzoek? Of word je graag op de hoogte gehouden van de resultaten van dit onderzoek? Of beter, maak je graag kans op het winnen van een prijs verbonden aan de deelname aan dit onderzoek?

Duid dan aan wat voor jou van toepassing is.

- ☐ Ik ben bereid om mee te werken aan vervolgonderzoek

- ☐ Ik word graag op de hoogte gehouden van de resultaten van dit onderzoek
- ☐ Ik maak graag kans om een prijs te winnen verbonden aan de deelname aan dit onderzoek.

Vul hieronder jouw naam en e-mailadres in.

Naam:

E-mailadres:

Appendix B

Part 1. Background information

16.

- ☐ Male
- ☐ Female

17. Year of birth: _____

18. Nationality: _____

19. Civil status:

- ☐ I am married
- ☐ I am divorced
- ☐ I am a widow/widower
- ☐ I am unmarried
- ☐ I am single
- ☐ I am living with a partner
- ☐ Other: _____

20. Which of the following descriptions best reflects your current employment situation?

- ☐ Employed full-time
- ☐ Employed part-time
- ☐ Job seeker
- ☐ Retired or early retired
- ☐ Full-time student
- ☐ Homemaker
- ☐ Other: _____

21. Please indicate your highest obtained degree/diploma

- ☐ None
- ☐ Primary education
- ☐ Lower secondary education
- ☐ Higher secondary education
- ☐ Professional bachelor's degree or associate degree
- ☐ Academic bachelor's degree
- ☐ Master's degree
- ☐ Doctor degree
- ☐ Other: _____

22. At which center are you currently following the classes?

More specifically, which center has contacted you for this questionnaire?

23. Which program are you following?

24. When did you begin the program? (year + semester if applicable)

25. Do you have previous online/blended learning experience?

- ☐ Yes
- ☐ No

26. How many modules have you completed in this program? _____

27. In general, how would you rate your ICT (Information and Communication Technology) skills (e.g., searching the Internet, using Microsoft Office, social media such as Facebook and Twitter)?

☐ Beginner

☐ Intermediate

☐ Advanced

28. Why did you enroll for the program? Please indicate the most important reason (s).

- ☐ I'd like to get the diploma.
- ☐ I'd like to improve my knowledge and skills for my current and/or future job.
- ☐ I'd like to make new friends and meet new people.
- ☐ I'd like to get more fun and relief from my daily routine and boredom.
- ☐ I'd like to learn new knowledge and skills.
- ☐ I participate because this course will prepare me for service to the community.
- ☐ I participate because I need to comply with my employer's policy or to secure social benefits.

29. What average score have you obtained so far during this course/program?

Please note that this score should be expressed on a **20-point scale**. If you received a score using a different grading system, please select **"Other"** and specify it. If you have **not yet received any grades**, please proceed directly to **Question 15**.

- ☐ ≤ 10
- ☐ 11-12
- ☐ 13-14
- ☐ 15-16
- ☐ 17-18
- ☐ 19-20
- ☐ Other: _____

Not applicable → Go to question 15

30. Please answer this question only if you could not answer Question 14 because it was "not applicable."

What average score do you expect to achieve throughout this course/program?

Please note that this score should be expressed on a **20-point scale**. If your center uses a different grading system, please select **"Other"** and specify the grading scale used.

- ☐ ≤ 10
- ☐ 11-12
- ☐ 13-14
- ☐ 15-16
- ☐ 17-18
- ☐ 19-20
- ☐ Other: _____

Part 2. 1. Transparency

What needs do you have regarding setting clear expectations during the course? We would like to know to what extent you need clarity about what is expected from you during the course.

Please indicate the extent to which you agree with the following statements, using the scale below:

(1) totally disagree, (2) disagree, (3) neutral, (4) agree, or (5) totally agree

	1	2	3	4	5
I have a clear idea of where I am going and what is expected of me in this course.					
It is always easy to know the standard of work expected.					
The instructor makes it clear right from the start what they expected from students.					

	1	2	3	4	5
It is easy to find out what is expected of me in this course.					

Part 2.2. The learning platform

What needs do you have while working with the online learning platform (e.g., Modell or Smartschool)?

Please indicate the extent to which you agree with the following statements, using the scale below:

(1) totally disagree, (2) disagree, (3) neutral, (4) agree, or (5) totally agree

	1	2	3	4	5
I can find the information on the learning platform easily.					
The program directions and navigations are clear.					
The material and content uploaded on the learning platform are well-organized.					
The learning platform's graphical user interface is good.					

Part 2. 3. Technical support

What do you consider as necessary regarding technical support. Please indicate the extent to which you agree with the following statements, using the scale below:

(1) totally disagree, (2) disagree, (3) neutral, (4) agree, or (5) totally agree

	1	2	3	4	5
I have as few as technical issues in this course (program) as possible.					
I can get technical support when I need.					
I know where to ask for help when I had any technical issues.					
Technical support (team) responds to my issues in a timely manner.					

Part 2. 4. Relevance

To what extent do you need relevant topics during the course?

Relevant topics are subjects or themes that interest you personally.

Please indicate the extent to which you agree with the following statements, using the scale below:

(1) totally disagree, (2) disagree, (3) neutral, (4) agree, or (5) totally agree

	1	2	3	4	5
I can relate what I learn to my life outside the campus.					
I am able to pursue topics that interest me.					
I link class work to my life outside the campus.					
I can apply my everyday experiences in class.					

Part 2. 5. Learning support

To what extent do you need the instructors' support?

Please indicate the extent to which you agree with the following statements, using the scale below:

(1) totally disagree, (2) disagree, (3) neutral, (4) agree, or (5) totally agree

	1	2	3	4	5
The instructor gives learners adequate feedback and enable them to calculate and track their performance.					
The instructor offers learners timely guidance.					
The instructor provides learners tasks and assignments appropriate to learners' capacity and prior experience.					
The instructor provides learners opportunities to apply knowledge from the course into practice.					

Part 2.6. Autonomy

To what extent do you think it is important to be able to study independently and autonomously? Please indicate your opinion

(1) totally disagree, (2) disagree, (3) neutral, (4) agree, or (5) totally agree

	1	2	3	4	5
I can work during times I find convenient.					
I am in control of my learning.					
I play an important role in my learning.					

	1	2	3	4	5
I approach learning in my own way.					

Part 2.7. Facilitation

To what extent do you think the instructor's involvement during online learning sessions necessary?

Please indicate the extent to which you agree with the following statements, using the scale below:

(1) totally disagree, (2) disagree, (3) neutral, (4) agree, or (5) totally agree

	1	2	3	4	5
The instructor helps to keep course participants engaged and participating in productive dialogue.					
The instructor helps keep the course participants on task in a way that helped me to learn.					
The instructor encourages course participants to explore new concepts in this course.					
Instructor actions reinforce the development of a sense of community among course participants.					

Part 2.8. Peer interaction

What needs do you experience when collaborating with fellow learners? Collaboration may refer to group work or group presentations.

Please indicate the extent to which you agree with the following statements, using the scale below:

(1) totally disagree, (2) disagree, (3) neutral, (4) agree, or (5) totally agree

	1	2	3	4	5
I can work collaboratively with peers in a project/task.					
I can discuss my ideas with other students.					
I can receive feedback from other students to improve my work.					
I can have access to other students' output, e.g. essays in blogs or wikis, submitted assignments, or oral presentation.					

Part 2.9. Sense of belonging

Please indicate the extent to which you agree with the following statements, using the scale below:

(2) totally disagree, (2) disagree, (3) neutral, (4) agree, or (5) totally agree

	1	2	3	4	5
I have a feeling that students in this course care about each other.					
I feel connected to other in this course.					
I feel that I can rely on others in this course.					
I feel confident that others will support me.					

Appendix C

Table A1: Factor loadings for the learners' needs instrument with scales treated as continuous and ordinal

Scales	Items	Factor loadings		Composite reliability (ω)
		Continuous scale	Ordinal scale	
Transparency	n1	0.556	0.604	0.76
	n2	0.777	0.826	
	n3	0.638	0.710	
	n4	0.682	0.813	
LMS quality	n5	0.709	0.819	0.84
	n6	0.816	0.906	
	n7	0.807	0.902	
	n8	0.659	0.792	
Technical support	n9	0.596	0.728	0.84
	n10	0.836	0.858	
	n11	0.831	0.876	
	n12	0.711	0.837	
Relevance of learning content	n13	0.726	0.808	0.79
	n14	0.627	0.727	
	n15	0.775	0.769	
	n16	0.648	0.885	
Competence support	n17	0.565	0.670	0.76
	n18	0.753	0.769	
	n19	0.708	0.740	
	n20	0.623	0.743	
Autonomy	n21	0.596	0.722	0.86
	n22	0.868	0.890	
	n23	0.840	0.932	
	n24	0.772	0.810	
Instructor's facilitation	n25	0.779	0.799	0.84
	n26	0.931	0.920	
	n27	0.845	0.820	
	n28	0.401	0.813	

Scales	Items	Factor loadings		Composite reliability (ω)
		Continuous scale	Ordinal scale	
Peer collaboration	n29	0.788	0.845	0.84
	n30	0.931	0.916	
	n31	0.845	0.905	
	n32	0.401	0.733	
Sense of belonging	n33	0.886	0.918	0.93
	n34	0.903	0.937	
	n35	0.874	0.896	
	n36	0.840	0.903	

Table A2: The second order factor and the respective components

Higher order factor	First order factor	Factor loadings
Existence	Transparency	0.691
	LMS quality	0.728
	Technical support	0.763
Relatedness	Peer collaboration	0.803
	Sense of belonging	0.886
	Instructor's facilitation	0.713
Growth	Relevance of learning content	0.599
	Competence support	0.825
	Autonomy	0.322

Table A3: Multivariate analysis: Effect of CVO on adults' perceptions of learning needs

Effect		Value	F	Hypothesis df	Error df	p-value
CVO B ^a	Pillai's Trace	0.027	.572	9	189	0.819
	Wilks' Lambda	0.973	.572	9	189	0.819
CVO C ^a	Pillai's Trace	0.024	.514	9	189	0.864
	Wilks' Lambda	0.976	.514	9	189	0.864
CVO D ^a	Pillai's Trace	0.043	.946	9	189	0.487
	Wilks' Lambda	0.957	.946	9	189	0.487
CVO E ^a	Pillai's Trace	0.054	1.203	9	189	0.295
	Wilks' Lambda	0.946	1.203	9	189	0.295
CVO F ^a	Pillai's Trace	0.077	1.744	9	189	0.082
	Wilks' Lambda	0.923	1.744	9	189	0.082
CVO G ^a	Pillai's Trace	0.034	.740	9	189	0.672
	Wilks' Lambda	0.966	.740	9	189	0.672
CVO H ^a	Pillai's Trace	0.014	.305	9	189	0.973
	Wilks' Lambda	0.986	.305	9	189	0.973

^areference category: CVO A

Table A4: Multivariate analysis: Effect of program categories on adults' perceptions of learning needs

Effect		Value	F	Hypothesis df	Error df	p-value
Program category 1^a	Pillai's Trace	0.070	1.615	9	194	0.113
	Wilks' Lambda	0.930	1.615	9	194	0.113
Program category 2^a	Pillai's Trace	0.060	1.386	9	194	0.196
	Wilks' Lambda	0.940	1.386	9	194	0.196
Program category 3^a	Pillai's Trace	0.021	.463	9	194	0.898
	Wilks' Lambda	0.979	.463	9	194	0.898

^areference category: program category 4

Table A5: Levene's tests of equality of covariance matrices

Dependent variables	F	df1	df2	p-value
Transparency	1.002	2	165	0.369
LMS quality	0.518	2	165	0.597
Technical support	0.940	2	165	0.393
Relevance of learning content	0.016	2	165	0.984
Competence support	0.281	2	165	0.755
Autonomy	1.103	2	165	0.334
Instructor's facilitation	0.674	2	165	0.511
Peer collaboration	1.930	2	165	0.148
Sense of belonging	0.578	2	165	0.562

Multi-level Emergency Remote Teaching Framework for K–12 Distance Education: A Stakeholder-based Qualitative Study

Bünyami Kayali¹ and Selçuk Karaman²

¹Bayburt University, Türkiye

²Ankara Hacı Bayram Veli University, Türkiye

bunyamikayali@bayburt.edu.tr

selcuk.karaman@hbv.edu.tr

<https://doi.org/10.34190/ejel.24.4.4847>

An open access article under [CC Attribution 4.0](#)

Abstract: This study develops a multi-level Emergency Remote Teaching (ERT) framework for K–12 distance education, grounded in the experiences of students, teachers, parents, and administrators during the COVID-19 pandemic. Prior ERT research has predominantly focused on single stakeholder groups or produced descriptive findings without synthesizing perspectives into a structured response model. Using a qualitative case study design, semi-structured interviews were conducted with 100 participants (25 per stakeholder group) selected through maximum variation sampling. Thematic content analysis was employed to identify convergent and divergent priorities across groups, which were subsequently mapped onto the five layers of the Holistic Analytical Layer Framework (HALF): policy, system, program, course, and activity, based on their functional relevance to each layer. Findings indicate that digital infrastructure functions as a foundational system-level prerequisite shared across all stakeholder groups, while pedagogical, psychosocial, and governance priorities diverge by role. Teachers emphasized assessment reliability and instructional continuity; students highlighted engagement and psychosocial support; parents stressed communication and guidance mechanisms; while administrators prioritized governance flexibility and decision-making autonomy. The resulting framework positions ERT not as an isolated emergency improvisation but as a coordinated, multi-layered response architecture that aligns stakeholder expectations across interconnected educational layers and supports the coordination of policy, infrastructure, programmatic planning, and instructional practices during crisis periods. The study contributes to the distance education literature by transforming multi-stakeholder perspectives into an integrated framework and extending the application of HALF to emergency education contexts. For e-learning practice, the framework offers school administrators and policymakers a concrete diagnostic tool for auditing crisis-preparedness across governance, infrastructure, and instructional layers, while offering teacher educators a basis for designing professional development that addresses both technical and psychosocial readiness for remote instruction. Given its single-region, qualitative design, the framework should be understood as a context-sensitive foundation for future empirical validation rather than a universally transferable model.

Keywords: Emergency remote teaching, Distance education, K–12, Stakeholder perspectives, Crisis preparedness, Digital equity

1. Introduction

The coronavirus, first identified in China in December 2019, rapidly evolved into a global pandemic, profoundly disrupting social, economic, and educational systems worldwide (Bozkurt, 2020). The crisis challenged established perceptions of work, education, and social life, compelling individuals and institutions to navigate unprecedented circumstances (Balci, 2020). Although the immediate health influences of the pandemic have diminished, its long-term consequences continue to reshape educational practices, technological engagement, and institutional preparedness (Zhao, 2020).

The pandemic accelerated transformations already underway in Information and Communication Technologies (ICT), digital learning environments, and evolving conceptions of knowledge and 21st-century skills (Kyllonen, 2012; McLoughlin and Lee, 2007; Van Weert, 2006). These developments sped up ongoing debates about digital transformation, educational flexibility, and institutional resilience (Aktan, 2021).

During the pandemic, ERT was widely implemented as a temporary solution to sustain educational continuity. In the post-pandemic discourse, it has frequently been stated that certain emergency distance education practices will continue to shape future educational models (Telli Yamamoto and Altun, 2020). Similar institutional statements were made by the Ministry of National Education (MoNE, 2020). These reflections highlight a broader recognition that education systems were insufficiently prepared for large-scale disruptions and must develop structured preparedness strategies for future crises (Brooks et al., 2020; Han, Demirbilek and Demirtaş, 2021).

Within the scope of K–12 ERT research, studies have largely focused on stakeholder perceptions of distance education (e.g., Ball, 2022; Kruczek et al., 2022), evaluations of emergency practices (e.g., Fiş Erümit, 2021; Gören et al., 2020; O'Keefe et al., 2020), identification of challenges encountered during implementation (e.g., Trust and Whalen, 2021), the broader effects of the pandemic on educational processes (e.g., Apaydın and Manolova Yalçın, 2024; Parveen et al., 2022), and recommendations for the post-pandemic period (e.g., Javorsky et al., 2022; Nisiforou, Kosmas and Vrasidas, 2021; Sofianidis et al., 2021). Although some studies have incorporated multiple stakeholder groups, research simultaneously involving students, teachers, parents, and administrators remains limited. Furthermore, even studies that include multiple stakeholders tend to present descriptive findings or fragmented recommendations rather than synthesizing these insights into a comprehensive and operational framework. This pattern is empirically documented in the literature. Bond (2021), in a living rapid systematic review of 89 K–12 ERT studies across five international databases, identified the underrepresentation of diverse stakeholder perspectives as a critical gap, noting that research remained predominantly teacher-focused. Saroglou and Kostas (2023) similarly found, in a systematic review of 46 K–12 ERT studies, that teachers constituted the primary sample in 67.3% of studies, while only one study incorporated all stakeholder groups, including parents. Al Mazrooei et al. (2022), drawing on 51 studies indexed in Web of Science and Scopus, likewise confirmed that K–12 ERT research has largely been confined to teacher and student experiences, producing taxonomies of challenges rather than governance-oriented response architectures. Even in studies that incorporate multiple stakeholder groups (e.g., An et al., 2022), findings remain largely descriptive and role-specific, offering limited guidance for developing integrated crisis-response frameworks. Accordingly, the present study adopts the HALF as an integrative analytical lens that systematically aligns governance, infrastructural, pedagogical, and psychosocial dimensions within a coherent multi-level architecture. This approach enables the synthesis of stakeholder experiences and expectations into a structured ERT framework for K–12 education, thereby extending existing research beyond descriptive accounts toward a governance-oriented model for crisis preparedness and educational continuity.

Despite the proliferation of pandemic-related research, a significant gap persists in translating multi-stakeholder experiences into a structured, multi-layered response model for K–12 distance education. Existing research often examines isolated perspectives, thereby overlooking the interconnected and systemic nature of educational ecosystems. More importantly, few studies attempt to transform qualitative stakeholder insights into a coherent governance-oriented framework capable of guiding future crisis preparedness. Unlike prior studies that list stakeholder challenges without integrating them into a systemic model, the present study moves beyond descriptive aggregation by mapping stakeholder expectations onto the hierarchical layers of the HALF framework, thereby generating a governance-oriented framework that organizes stakeholder priorities within a coherent multi-level structure rather than presenting them as a thematic list of recommendations.

Rather than offering a fragmented list of recommendations, this study integrates the perspectives of students, teachers, parents, and administrators into a structured Emergency Remote Teaching framework for K–12 distance education. By employing the HALF as a synthesis mechanism, the study proposes a multi-level model encompassing policy, system, program, course, and activity layers. In doing so, it moves beyond descriptive reporting to develop a stakeholder-driven, operational roadmap aimed at strengthening institutional resilience and sustaining educational continuity in future crisis-driven disruptions. Accordingly, this study seeks to answer the following research question:

- How do the experiences and expectations of students, teachers, parents, and administrators shape the requirements for sustaining educational continuity during crisis-driven disruptions in K–12 education?

2. Theoretical Framework

2.1 Distance Education

Distance education can be defined as a planned, institutional, and administrative process in which students and teachers are not in the same environment, different technologies are used, and course designs and teaching methods are determined by the nature of distance education (Moore and Kearsley, 2011).

Considering both the above definition and other definitions in the literature (Hawkins, 1999; Holmberg, 1995), it can be stated that the primary purpose of distance education activities is to maintain education and training activities with the help of technologies that can contribute to education by providing people with time-space flexibility (Özbay, 2015).

2.2 Emergency Remote Teaching

ERT allows learners to safely access educational activities and materials during crisis periods, unlike planned, voluntary distance education (Hodges et al., 2020). Bozkurt and Sharma (2020) defined ERT as online education activities carried out in an unplanned and improvised manner due to force majeure, such as epidemics, wars, or disasters. Although the concept of ERT, which entered our lives during the COVID-19 pandemic, appears similar to that of distance education, there are many differences between the two. When these differences are examined, it is seen that distance education is a preference, while ERT is a necessity; distance education offers permanent solutions, while ERT provides temporary solutions to crisis periods; distance education activities are carried out in a planned and systematic manner, while ERT activities are improvised (Hodges et al., 2020).

2.3 Holistic Analytical Layer Framework

HALF was proposed by Karaman and Kurşun (2020) as a conceptual tool to guide both research and practice in distance education. The framework emerged in response to persistent challenges in the field, including terminological inconsistencies, the dominance of face-to-face education assumptions in distance education research, methodological ambiguities, and the gap between theoretical discourse and practical implementation.

HALF conceptualizes distance education systems as interconnected and multi-layered structures rather than isolated instructional environments. Within this perspective, the components of the distance education process and their interrelationships are organized across five hierarchical yet interdependent levels: policy, system, program, course, and activity. This layered structure enables the alignment of governance decisions, infrastructural capacities, curricular organization, instructional design, and classroom-level practices within a coherent analytical framework.

The policy level represents the broadest and most strategic dimension, where national or institutional distance education policies are formulated, and long-term goals are established. The system level encompasses technological infrastructure, including software, hardware, and organizational support mechanisms that sustain the implementation of distance education. The program level focuses on the structural organization of diploma or certification programs delivered through distance education. The course level addresses the design, implementation, and evaluation of individual courses in alignment with program-level principles. Finally, the activity level, considered the core operational layer, includes the design and execution of specific instructional activities. At this level, variables such as learning environment, tools and materials, participant roles, timing, interaction patterns, and instructional configurations are determined.

From a theoretical standpoint, HALF aligns with systems-oriented approaches in education by recognizing the interdependence among institutional layers. In crisis contexts such as emergency remote teaching, disruptions at one level inevitably influence those at other levels. Therefore, analyzing ERT exclusively at the classroom or course level may overlook broader structural determinants. By organizing stakeholder experiences across these five interconnected layers, HALF serves as an integrative synthesis mechanism that transforms qualitative insights into a structured, operational response model. In this study, the framework serves as an analytical lens to systematically map stakeholder recommendations onto governance, infrastructure, programmatic, and pedagogical dimensions, thereby enabling the development of a coherent multi-level ERT framework for K–12 distance education.

3. Method

This study used a qualitative case study design (Merriam, 1988). The case study approach was selected because it enables an in-depth exploration of complex educational phenomena within their real-life contexts. In the context of ERT, where multiple stakeholder perspectives intersect within dynamic institutional environments, a case study design allows for a comprehensive examination of experiences, practices, and contextual conditions. The case in the current study concerns the emergency remote teaching experiences of students, teachers, parents, and administrators within the K–12 education system of a single northeastern province of Türkiye during the 2020–2021 academic year.

As widely acknowledged in educational research (Gall, Gall and Borg, 2007), case studies are particularly suitable for investigating contemporary events over which researchers have limited control. By focusing on stakeholder experiences during the pandemic, this approach provides insight not only into how educational processes unfolded under crisis conditions but also into the structural and systemic factors that shaped those experiences

(Yin, 2018). Consequently, the case study design facilitated the identification of context-sensitive recommendations and the development of a structured framework grounded in lived realities.

3.1 Research Group

Participants were recruited from public schools in a province in Türkiye's northeast during the 2020–2021 academic year, a period characterized by intensive ERT implementation. The province was selected as it reflects diverse socio-economic and technological conditions, thereby providing a suitable context for examining variations in distance education experiences.

The study sample consisted of 100 participants: 25 teachers, 25 students, 25 school administrators, and 25 parents. To ensure representation across educational stages, three schools were selected from each of the primary, middle, and high school levels (Grades 1–12), for a total of 9 schools. Participants were then recruited voluntarily from these schools.

Maximum variation sampling, a purposive sampling strategy, was employed to capture a broad spectrum of experiences and contextual differences (Suri, 2011). This approach enables the exploration of rich and diverse perspectives rather than aiming for statistical generalization. To enhance variation, schools were selected to represent different institutional characteristics, including project high schools, academically selective (qualified) high schools, village schools, and transportation-centered schools. This diversity allowed the study to reflect differences in technological infrastructure, student profiles, and administrative capacity, thereby strengthening the transferability of findings. In addition, variation was ensured across educational level (primary, middle, and high school), geographic location (urban, peri-urban, and rural), and stakeholder role (teacher, student, parent, and administrator).

3.2 Data Collection Tools

Data were collected through semi-structured interviews designed for each stakeholder group. Semi-structured interviews were preferred as they allow for in-depth exploration of participants' lived experiences while maintaining flexibility to probe emerging themes (McMillan and Schumacher, 2010). This approach was particularly well-suited to capturing nuanced perspectives on emergency remote teaching practices during the pandemic.

The development of the interview protocols followed a systematic process. First, a comprehensive literature review of emergency remote teaching and distance education was conducted, resulting in an initial question pool. The draft interview forms were then reviewed by two field experts—one Professor and one Assistant Professor—to ensure content relevance and clarity. Based on their feedback, revisions were made prior to pilot testing.

The pilot study involved two participants from each stakeholder group. Following the pilot interviews, necessary revisions were implemented, including refining question wording, reordering items for clarity, and removing redundant prompts. To enhance linguistic accuracy and comprehensibility, the finalized forms were reviewed by a language expert with eight years of experience in Turkish Language Teaching.

Due to pandemic-related restrictions, all interviews were conducted online via Zoom. With participants' informed consent, interviews were audio-recorded to ensure accuracy and prevent data loss. Each interview lasted approximately 20–25 minutes.

3.3 Data Analysis

The qualitative data were analyzed using thematic content analysis. Content analysis was selected as it enables the systematic examination of textual data and facilitates the identification of patterns, categories, and themes aligned with the research purpose (Krippendorff, 1980; Patton, 2002). The analysis aimed to transform stakeholder narratives into structured thematic categories that could be systematically mapped onto the layers of the HALF.

The analysis process followed a structured coding procedure (Zhang and Wildemuth, 2009). First, all interview recordings were transcribed verbatim and reviewed multiple times to ensure transcription accuracy. During this stage, transcripts were cleaned and anonymized. Second, an initial coding scheme was developed through open coding. To enhance consistency, eight interviews selected randomly were jointly coded by the researchers to establish shared definitions, coding boundaries, and interpretative alignment. Following this calibration phase, the remaining transcripts were divided between the two researchers for independent coding. To protect

participant confidentiality, each participant was assigned a coded identifier: T (Teacher), S (Student), P (Parent), and A (Administrator), followed by a number (e.g., T-3, S-15).

To ensure inter-coder reliability, 20 randomly selected interviews were re-coded independently by both researchers after the initial coding phase. Inter-coder agreement was calculated at the category level using Cohen's kappa coefficient based on the consistency of coding across identified meaning units. The resulting kappa value was $\kappa = .86$, indicating almost perfect agreement between coders. Discrepancies were discussed and resolved through consensus to finalize the coding structure.

To further enhance analytical rigor, the preliminary coding structure and thematic categorizations were reviewed by two external field experts. Based on their feedback, three types of revisions were implemented: removing codes deemed conceptually inconsistent; refining or renaming codes for clarity and precision; and merging overlapping codes to improve conceptual coherence. This iterative refinement process ensured both internal consistency and external validation of the analytical framework. Themes were mapped onto the HALF layers only after the completion of thematic analysis. Codes were first derived inductively from the data and subsequently organized within the HALF structure according to their functional relevance to each layer. Consistent with theory-guided thematic analysis, the framework served as an analytical lens for organizing findings rather than as a deductive coding template (Braun and Clarke, 2006).

3.4 Reliability and Validity

In qualitative research, rigor is commonly evaluated through the criteria of credibility, transferability, dependability, and confirmability rather than traditional notions of validity and reliability (Guba, 1981; Guba and Lincoln, 1994; Yıldırım and Şimşek, 2013). Accordingly, several strategies were implemented to enhance the study's trustworthiness.

To ensure credibility, expert and peer feedback were obtained at multiple stages of the research process, including the development of the interview protocols and the data analysis phase. In addition, voluntary participation and prolonged engagement with the data through repeated transcript review strengthened interpretive accuracy.

Transferability was supported by providing detailed descriptions of the research context, sampling strategy, participant characteristics, and analytical procedures, thereby enabling readers to assess the applicability of findings to similar contexts.

To enhance dependability, interviews were audio-recorded and systematically transcribed. Inter-coder consistency was evaluated using Cohen's kappa coefficient, and coding decisions were reviewed collaboratively to ensure procedural stability (see Data Analysis section).

Finally, confirmability was ensured through a comprehensive audit trail. All raw data, coding files, and analytical notes were digitally archived to enable external review and ensure that findings were grounded in the data rather than in researcher bias.

4. Findings

This section presents the findings derived from the qualitative analysis of stakeholder perspectives regarding educational practices during crisis periods. The data, organized around six main themes, were interpreted through the HALF to provide a multi-layered understanding of ERT requirements. Frequency counts reported in the table are intended to indicate the relative salience of themes across stakeholder groups rather than to support statistical inference; they are presented as a transparency measure consistent with the qualitative nature of the study, following established practice in qualitative content analysis (Krippendorff, 1980; Patton, 2002). The distribution of stakeholder recommendations and their frequencies are summarized in Table 1.

Table 1: Thematic Analysis of Teacher, Student, Parent, and Administrator Views on Educational Strategies in Times of Crisis

Theme	Subtheme / Code	Teacher (f)	Student (f)	Parent (f)	Administrator (f)
1. Educational Planning and Implementation	Evaluation of absenteeism	20	6	9	—
	Hybrid education models	10	6	7	9
	Development of instructional content and materials	14	5	7	—
	Regulation of lesson duration	13	4	7	12
	Small-group instruction	—	7	—	—
	Ensuring active participation in lessons	—	8	—	—
2. Supportive and Preventive Practices	Remedial education planning	5	4	9	10
	Classroom management strategies	—	4	—	10
	Guidance and counseling services	—	—	—	3
	Education on screen and technology addiction	—	2	—	—
3. Capacity Building and Professional Development	Professional development for teachers	18	—	—	—
	Training on technology use	13	13	10	17
	Parental education programs	—	—	10	—
	Student mentoring and peer-support practices	—	—	9	—
4. Governance and Structural Arrangements	Definition of legal responsibilities	—	—	—	9
	Participation in decision-making processes	—	—	—	9
	Timely dissemination of official decisions and information	—	—	—	15
	Crisis and emergency preparedness planning	—	—	—	10
	Development of implementation guidelines	—	—	—	3

Theme	Subtheme / Code	Teacher (f)	Student (f)	Parent (f)	Administrator (f)
5. Equity and Access in Education	Access to the internet and digital devices	22	16	11	22
	Provision of printed learning materials	—	—	10	—
	Ensuring educational continuity in rural areas	—	—	—	6
	Use of mobile and digital learning tools	—	2	—	—
6. Assessment, Evaluation, and Monitoring	Assessment and evaluation practices	15	—	—	13
	Reliability of online assessments	—	5	3	10
	Alternative assessment methods	—	—	—	4
	Monitoring of student performance	—	3	—	—

The thematic distribution in Table 1 indicates that, while certain challenges, particularly access to digital infrastructure, are universally recognized, stakeholders prioritize interventions differently based on their functional roles within the educational ecosystem. The following analysis interprets these findings across the systemic layers defined in HALF.

Access to internet connectivity and digital devices emerged as the most frequently cited barrier across all stakeholder groups (total $n = 71$), indicating a foundational system-level prerequisite for educational continuity. While teachers and administrators emphasized infrastructure provision, parents offered a pragmatic alternative: printed learning materials ($n = 10$), highlighting a policy-level solution to digital inequities. The lived experience behind this finding was particularly acute in rural settings. One student explained:

“Our house is in a village, and we could not get internet connected. They should provide internet connectivity to our village, or if that is not possible, they should distribute printed materials to students living in villages” (S-15).

A parent similarly described the financial burden of adaptation:

“When the process began, we bought a tablet and had internet installed at home so that my son could attend lessons” (P-14).

These accounts indicate that digital inequality was experienced not only as a technological challenge but also as an economic burden for families.

Governance-related themes were almost exclusively articulated by administrators, who emphasized timely dissemination of official decisions ($n = 15$), clarification of legal responsibilities ($n = 9$), and crisis preparedness planning ($n = 10$). One administrator stated:

“Administrators should be given more authority to make decisions based on the conditions of their own region. During periods when the pandemic was less severe, they should be able to decide on hybrid education or alternative approaches” (A-23).

Another emphasized the procedural dimension:

“Attendance requirements should be enforced for students, and assessment and evaluation must be conducted. Necessary measures should be taken to ensure that assessment and evaluation are carried out reliably” (A-9).

These findings indicate that administrators attached particular importance to decision-making processes, institutional responsibilities, and crisis preparedness.

Capacity building emerged as a multidimensional program-level requirement. Teachers (n = 18) and administrators (n = 17) strongly advocated professional development and technology training, framing ERT as a competence-based challenge. Parents broadened this perspective by emphasizing parental education (n = 10) and student mentoring practices (n = 9), suggesting that educational continuity extends beyond institutional boundaries into the home-learning environment. An administrator captured this multi-stakeholder dimension of capacity development:

“Administrators and teachers must definitely be included in the decision-making process. Technology use training should be provided for parents, teachers, and students” (A-8).

This finding suggests that stakeholders viewed technology-related competencies as a shared responsibility involving multiple actors.

Additionally, administrators underscored classroom management strategies (n = 10), whereas students uniquely highlighted the need for guidance on technology addiction (n = 2). These findings indicate that program-level preparedness must incorporate academic, technological, and psychosocial dimensions. Students directly articulated the psychosocial burden of ERT. One student described the cumulative anxiety of sustained distance learning:

“I was very anxious this year because of the high school entrance exam. Schools were closed, and I could not follow the lessons well through distance education. Over time, this situation began to cause me serious stress” (S-5).

Another student highlighted the protective role of in-school guidance:

“During the guidance sessions with our classroom teacher, I learned to take deep breaths and think about positive things when I felt overwhelmed. I was also glad that my teacher paid attention to me” (S-17).

Some students additionally expressed a preference for hybrid instructional arrangements:

“I strongly believe that in such situations, some lessons should be held face-to-face with the necessary precautions. Spending so much time in front of a computer made it difficult for me to concentrate, and I could not understand the lessons well” (S-7).

These accounts demonstrate that students experienced emergency remote teaching not only as an academic challenge but also as a psychosocial experience affecting well-being and engagement. Teachers also reported several assessment-related difficulties during emergency remote teaching.

“Internet disconnections during examinations and the lack of security in online assessments are among the problems that must be acknowledged. Students were also forced to answer questions on small screens, which posed an additional challenge” (T-3).

In response to these challenges, some administrators described alternative monitoring practices:

“For students preparing for high-stakes exams, we organized online practice tests as a substitute for the face-to-face mock exams we could not conduct, and in this way we tried to monitor student achievement” (A-8).

Overall, the findings reveal both shared and stakeholder-specific priorities regarding educational continuity during crisis periods. While access to digital infrastructure emerged as a common concern across groups, other issues—including governance, capacity building, psychosocial support, and assessment practices—were emphasized differently according to stakeholders’ roles and responsibilities within the educational system.

4.1 Synthesis of Findings: An ERT Framework for K–12 Education

The synthesis of findings demonstrates that stakeholder recommendations do not operate as isolated interventions but as interdependent elements embedded within a layered educational ecosystem. Rather than aggregating themes descriptively, the HALF was employed as an integrative synthesis mechanism to reorganize stakeholder expectations across policy, system, program, course, and activity levels.

Importantly, the analytical positioning of themes within the framework was determined not by frequency or magnitude but by their functional relevance to sustaining educational continuity during crisis conditions. This functional mapping enabled the transformation of dispersed qualitative insights into a structured and operational ERT architecture for K–12 education.

The comparative analysis indicates that stakeholder priorities extend across multiple interconnected layers of the educational system. By aligning stakeholder priorities within the hierarchical yet interdependent layers of the HALF, the study advances a coherent crisis-response model grounded in lived experience rather than abstract prescription (see Figure 1).

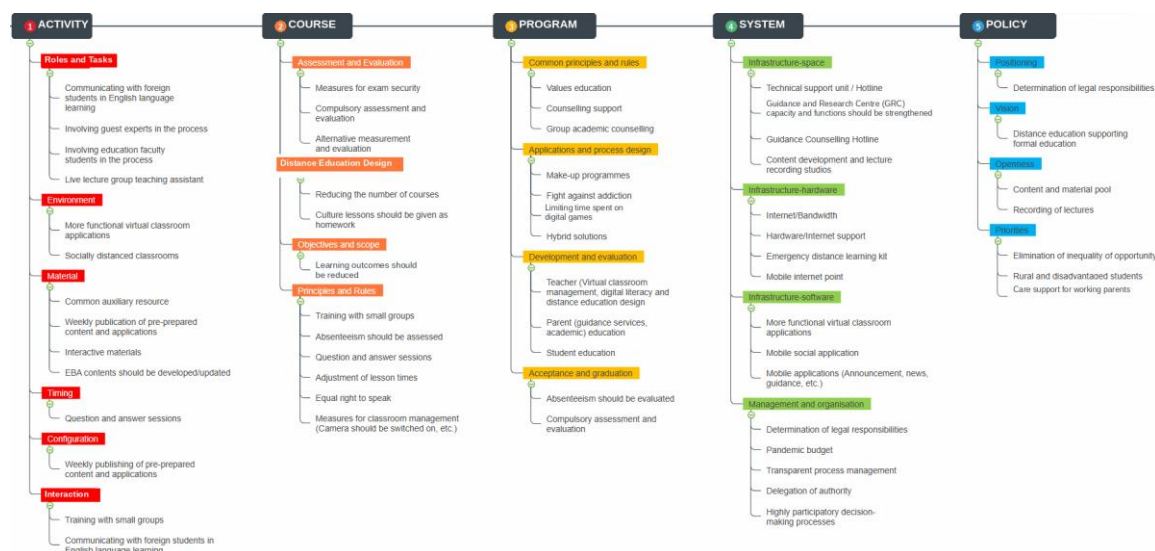


Figure 1: ERT Framework for K–12 Education Based on Stakeholder Findings and the HALF (Developed by the authors based on Karaman and Kurşun's (2020) HALF)

As illustrated in Figure 1, the proposed ERT framework operationalizes stakeholder-derived expectations across five interconnected layers. At the policy level, equity and hybridity are positioned as normative standards, reframing digital access as a foundational policy priority and hybrid flexibility as a structural safeguard rather than a temporary accommodation. At the system level, infrastructure provision is integrated with participatory governance, suggesting that technological readiness must be complemented by administrative agility and decentralized decision-making to enable context-responsive implementation. At the program level, the framework extends beyond curricular continuity to incorporate psychosocial resilience, including digital well-being, addiction prevention, and value-based education, thereby signalling a shift from academic recovery toward holistic learner protection. At the course level, procedural formalization addresses concerns about reliability and accountability while encouraging a transition toward process-oriented, diversified assessment practices. Finally, at the activity level, the inclusion of external instructional actors, such as education faculty students and mentors, introduces a collaborative instructional model aimed at alleviating teacher overload and fostering distributed responsibility.

Collectively, the framework aligns governance, infrastructure, programmatic, and pedagogical expectations across the layers defined in the HALF, providing a coherent structure for translating stakeholder-derived priorities into a crisis-response model.

5. Discussion

The ERT framework proposed in this study, synthesized from multi-stakeholder experiences, demonstrates that crisis-period education extends beyond a technical modality shift and instead constitutes a multi-layered systemic transformation. The findings indicate that sustainable emergency remote teaching requires coordinated alignment across infrastructural, administrative, pedagogical, and psychosocial dimensions.

At the system level, the pronounced demand for hardware, software, and stable internet connectivity confirms that infrastructure is a foundational prerequisite for ERT. While prior studies have identified digital inequality as a structural barrier (Bozkurt and Sharma, 2020; Jackson, 2021), the present study advances this discussion by revealing a rare cross-stakeholder consensus regarding access. Unlike role-specific concerns observed in other

thematic areas, digital equity emerged as a universally articulated necessity. This convergence suggests that unresolved infrastructural disparities render pedagogical innovations at higher layers structurally fragile.

In addition to infrastructure, administrators emphasized management autonomy and dedicated crisis budgets, echoing findings by Thornton (2021) and Yoskovitz and Schechter (2023). However, this study reframes these demands not merely as managerial preferences but as requirements for systemic flexibility. By institutionalizing decentralized authority, the proposed framework challenges centralized, uniform crisis responses and instead promotes context-sensitive adaptation, consistent with World Health Organization (WHO, 2020) recommendations regarding localized decision-making during public health disruptions.

At the activity level, the recommendation to involve education faculty students and group advisors signals a structural reconfiguration of instructional responsibility. While Avcı and Akdeniz (2021) documented increased teacher burnout during pandemic teaching, the present findings propose a distributed instructional model as a resilience strategy. This transition from the “isolated teacher” paradigm toward a “supported instructional team” reflects broader shifts in professional identity aligned with contemporary lifelong learning perspectives (Van Weert, 2006).

Similarly, the call for centralized material distribution underscores the need for content standardization to alleviate the time-intensive burden of producing individual materials, a concern long emphasized in the distance education literature (Rowntree, 1990). At the program level, stakeholder emphasis on screen dependency and digital addiction (Iivari, Sharma and Ventä-Olkkonen, 2020; King et al., 2020) broadens the scope of ERT beyond academic continuity. Within the layered logic of the HALF, this indicates that crisis-responsive education must incorporate a preventive dimension that safeguards psychosocial well-being alongside instructional delivery.

At the course level, tensions surrounding assessment and evaluation reveal an unresolved balance between institutional accountability and online reliability. While systematic assessment has been emphasized in distance education research (Cabi, 2016; Özdoğan and Berkant, 2020), participants in this study articulated a need for alternative, process-oriented evaluation strategies to mitigate reliability concerns associated with remote examinations (Koç, 2021). The advocacy for small-group instruction and active participation further reinforces learner-centered design principles (Cavanaugh, 2001), suggesting that effective ERT requires a pedagogical shift from broadcast-oriented instruction toward interaction-intensive engagement structures. Beyond confirming earlier findings, the present study suggests that the tension between institutional accountability and learner-centered evaluation represents not only a technical challenge but also a governance issue requiring coordinated responses across course and program layers. Despite this apparent convergence around assessment reform, stakeholder priorities differed substantially by role. Teachers framed assessment primarily as an issue of institutional validity and procedural reliability. In contrast, parents showed limited direct engagement with formal assessment structures, instead prioritizing their children’s psychosocial well-being and academic support at home. This tension between systemic accountability and relational care mirrors the findings of Huber and Helm (2020), who observed that institutional actors and family stakeholders frequently operate with differing priorities during crisis-period education.

Taken together, these dimensions—infrastructure, administrative autonomy, instructional collaboration, psychosocial protection, and assessment reform—operate not as isolated interventions but as interdependent components of a layered governance challenge. It should be acknowledged, however, that identifying convergent stakeholder concerns is a necessary but not sufficient basis for developing a comprehensive governance model. The proposed framework represents a first-order synthesis of stakeholder priorities organized through the analytical lens of HALF. Its translation into institutional policy would require iterative validation across diverse regional and socio-economic contexts.

6. Conclusion and Recommendations

This study offers a comprehensive roadmap for K–12 education during crisis periods by synthesizing the multi-perspective experiences of students, teachers, parents, and administrators into a structured, multi-layered framework. Organized through the HALF, the proposed model moves beyond temporary, reactive solutions and advances a systemic resilience approach grounded in governance, infrastructure, pedagogy, and psychosocial support.

The findings demonstrate that successful ERT cannot be reduced to technological delivery alone; rather, it requires coordinated consistency across policy, system, program, course, and activity levels. Although stakeholders expressed growing acceptance of digital integration, the sustainability of this transformation

remains contingent upon addressing structural inequalities, particularly those related to digital access and connectivity.

A central conclusion of this study is the necessity of stakeholder-specific yet systemically aligned support mechanisms. Teachers require pedagogical autonomy, professional development, and workload redistribution, whereas students and parents emphasize psychosocial protection, reliable communication, and clear governance structures. Consequently, future crisis preparedness should adopt an ecosystem-centered rather than teacher-centered orientation, integrating hybrid flexibility with inclusive and context-sensitive implementation strategies.

From a policy and practice perspective, several implications emerge. The following implications are presented in order of structural dependency, from foundational prerequisites to pedagogical adaptations, rather than as an exhaustive or equally weighted list. First, national curricula should incorporate emergency-ready modular structures that enable continuous transitions between face-to-face, online, and hybrid modalities without compromising learning outcomes. Second, digital equity should be regarded as a foundational policy priority, with hardware provision and reliable internet access serving as essential prerequisites for disadvantaged students. Third, assessment regulations should shift from high-stakes, exam-centered models toward continuous, process-oriented, and diversified evaluation strategies better suited to remote contexts. Fourth, school counseling systems should integrate digital well-being frameworks to address screen dependency, gaming risks, and other psychosocial vulnerabilities identified by stakeholders. Finally, collaborative governance structures should institutionalize localized administrative autonomy and formalize partnerships with external actors, such as education faculty and students serving as instructional mentors, to decrease teacher overload and enhance distributed responsibility.

For future research, the proposed ERT framework should be examined through longitudinal implementations and tested across diverse socio-economic regions of Türkiye to evaluate its adaptability, scalability, and long-term systemic impact. The framework should be interpreted with caution, as it emerged from a single regional context during an exceptional historical period. Its applicability to other settings should therefore be evaluated through context-sensitive adaptation rather than direct transfer.

7. Limitations

Despite its contributions to the conceptualization of a multi-level ERT framework, this current study has some limitations that warrant consideration. First, data were collected exclusively through semi-structured interviews conducted via digital platforms (Zoom). Although this approach ensured continuity of data collection during social distancing measures, it may have inadvertently limited participation among stakeholders with restricted digital literacy or those experiencing screen fatigue. Consequently, certain voices, particularly those most affected by digital exclusion, may be underrepresented in the findings.

Second, the study was conducted with 100 participants (25 per stakeholder group) in a specific regional context. While maximum variation sampling was employed to enhance diversity and transferability, the socio-cultural and infrastructural characteristics of the selected region inevitably shape the nature of the findings. Therefore, the proposed framework should be interpreted as context-sensitive rather than universally generalizable across all K–12 settings in Türkiye.

Finally, as a qualitative case study, the research prioritizes depth of insight over statistical representativeness. The objective was not to produce generalizable prevalence estimates but to construct a theoretically informed, stakeholder-driven framework. Future studies may strengthen the empirical robustness of the model by adopting mixed-method designs that integrate large-scale quantitative validation with qualitative contextual analysis across diverse socio-economic regions and educational levels. Longitudinal investigations would further help assess the sustainability and adaptive capacity of the proposed framework over time.

Acknowledgements

This article is based on the qualitative findings of the first author's **doctoral dissertation**, completed at Atatürk University, Graduate School of Educational Sciences, under the supervision of Prof. Dr. Selçuk Karaman. The authors would like to thank all participants for their time and valuable contributions to this study.

AI Statement: AI tools were used in the preparation of this manuscript. DeepL and ChatGPT were used to support the translation and localization of the manuscript content from Turkish into English, while Grammarly was used to check the grammar, spelling, and stylistic clarity of the English text. All AI-assisted output was subsequently

reviewed, verified, and edited by the human authors to ensure linguistic accuracy, cultural appropriateness, and fidelity to the original meaning. No AI tool was used to generate research ideas, data, analysis, or conclusions.

Ethics Statements: This study was approved by the Bayburt Provincial Directorate of National Education on 9 November 2020, under application number 13616634-605.01.E.16381839. This study was conducted in accordance with the Declaration of Helsinki.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Disclosure statement: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

- Aktan, S. (2021) 'Waking up to the dawn of a new era: Reconceptualization of curriculum post Covid-19'. *Prospects*, 51, pp. 205–217. <https://doi.org/10.1007/s11125-020-09529-3>
- Al Mazrooei, A. K., Almaki, S. H., Gunda, M., Alnoor, A. and Sulaiman, S. M. (2022) 'A systematic review of K–12 education responses to emergency remote teaching during the COVID-19 pandemic'. *International Review of Education*, 68(6), pp. 811–841. <https://doi.org/10.1007/s11159-023-09986-w>
- An, H., Mongillo, G., Sung, W. and Fuentes, D. (2022) 'Factors affecting online learning during the COVID-19 pandemic: The lived experiences of parents, teachers, and administrators in US high-needs K-12 schools'. *Journal of Online Learning Research*, 8(2), pp. 203–234. Available at: <https://www.learntechlib.org/primary/p/220404/> [Accessed 5 September 2022]
- Apaydın, Ç. and Manolova Yalçın, O. (2024) 'The impact of COVID-19 on K-12 school principals' management processes'. *SAGE Open*, 14(3), 21582440241265328. <https://doi.org/10.1177/21582440241265328>
- Avcı, F. and Akdeniz, E. C. (2021) 'Teachers' evaluations of the coronavirus (Covid-19) epidemic and the problems encountered in distance education'. *International Journal of Social Sciences and Education Research*, 3(4), pp. 117–154. Available at: <https://izlik.org/JA49UW88LL> [Accessed 25 May 2022]
- Balcı, A. (2020) 'The Effects of Outbreaks on Education in Specific to Covid-19'. *International Journal of Leadership Studies: Theory and Practice*, 3(3), pp. 75–85. Available at: <https://izlik.org/JA37HH74GR> [Accessed 27 January 2022]
- Ball, M. M. (2022) *K–12 public school teacher perceptions regarding their experiences as instructors who volunteered to teach in the online learning environment in 2020–2021 (COVID-19)*. PhD thesis. Virginia Polytechnic Institute and State University. Available at: <http://hdl.handle.net/10919/109711> [Accessed 28 December 2022]
- Bond, M. (2021) 'Schools and emergency remote education during the COVID-19 pandemic: A living rapid systematic review'. *Asian Journal of Distance Education*, 15(2), pp. 191–247. <https://doi.org/10.5281/zenodo.4425683>
- Bozkurt, A. (2020) 'The coronavirus pandemic process and evaluations on education in the post-pandemic world: New normal and new education paradigm'. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi*, 6(3), pp. 112–142. <https://doi.org/10.29065/usakead.777652>
- Bozkurt, A. and Sharma, R. C. (2020) 'Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic'. *Asian Journal of Distance Education*, 15(1), pp. i–vi <https://doi.org/10.5281/zenodo.3778083>
- Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology'. *Qualitative Research in Psychology*, 3(2), pp. 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N. and Rubin, G. J. (2020) 'The psychological impact of quarantine and how to reduce it: Rapid review of the evidence'. *The Lancet*, 395(10227), pp. 912–920. [https://doi.org/10.1016/S0140-6736\(20\)30460-8](https://doi.org/10.1016/S0140-6736(20)30460-8)
- Cabi, E. (2016) 'Uzaktan eğitimde e-değerlendirme üzerine öğrenci algıları' [Student perceptions on e-assessment in distance education]. *Journal of Higher Education and Science*, 6(1), pp. 94–101. <https://doi.org/10.5961/jhes.2016.146>
- Cavanaugh, C. S. (2001) 'The effectiveness of interactive distance education technologies in K- 12 learning: A meta-analysis'. *International Journal of Educational Telecommunications*, 7(1), pp. 73–88. Available at: <https://www.learntechlib.org/p/8461> [Accessed 20 December 2021]
- Fiş Erümit, S. (2021) 'The distance education process in K–12 schools during the pandemic period: Evaluation of implementations in Turkey from the student perspective'. *Technology, Pedagogy and Education*, 30(1), pp. 75–94. <https://doi.org/10.1080/1475939X.2020.1856178>
- Gall, M. D., Gall, J. P. and Borg, W. R. (2007) *Educational research: An introduction*. 8th edn. Pearson.
- Gören, S. Ç., Gök, F., Yalçın, M., Göregen, F. and Çalışkan, M. (2020) 'Küresel salgın sürecinde uzaktan eğitimin değerlendirilmesi: Ankara örneği' [Evaluation of distance education during the pandemic: The case of Ankara]. *Millî Eğitim Dergisi*, 49(1), pp. 69–94. <https://doi.org/10.37669/milliegitim.784822>
- Guba, E. G. (1981). 'Criteria for assessing the trustworthiness of naturalistic inquiries'. *Educational Communication and Technology Journal*, 29(2), pp. 75–91. <https://doi.org/10.1007/BF02766777>
- Guba, E. G. and Lincoln, Y. S. (1994) 'Competing paradigms in qualitative research.' in N. K. Denzin and Y. S. Lincoln (eds.), *Handbook of qualitative research*. Sage Publications, pp. 105–117.
- Han, F., Demirbilek, N. and Demirtaş, H. (2021) 'Koronavirüs (COVID-19) salgını sürecinde gerçekleştirilen uzaktan eğitim uygulamalarına ilişkin okul yöneticileri ve öğretmenlerin görüşleri' [Opinions of school administrators and teachers on distance education conducted

- during the coronavirus (COVID-19) outbreak]. *Cumhuriyet Journal of International Education*, 10(3), pp. 1168-1193. <https://doi.org/10.30703/cije.819946>
- Hawkins, B. L. (1999) 'Distributed learning and institutional restructuring'. *EDUCAUSE Review*, 34(4), pp. 12–15. 42–44.
- Hodges, C., Moore, S., Lockee, B., Trust, T. and Bond, A. (2020) 'The difference between emergency remote teaching and online learning'. *EDUCAUSE Review*. 27 March. Available at: <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning> [Accessed 22 May 2020]
- Holmberg, B. (1995) 'The evolution of the character and practice of distance education'. *Open Learning: The Journal of Open, Distance and e-Learning*, 10(2), pp. 47–53. <https://doi.org/10.1080/0268051950100207>
- Huber, S. G. and Helm, C. (2020) 'COVID-19 and schooling: Evaluation, assessment and accountability in crises- reacting quickly to explore key issues for policy, practice and research with the school barometer'. *Educational Assessment, Evaluation and Accountability*, 32(2), pp. 237–270. <https://doi.org/10.1007/s11092-020-09322-y>
- Iivari, N., Sharma, S. and Ventä-Olkkonen, L. (2020) 'Digital transformation of everyday life- How COVID-19 pandemic transformed the basic education of the young generation and why information management research should care?'. *International Journal of Information Management*, 55, 102183. <https://doi.org/10.1016/j.ijinfomgt.2020.102183>
- Jackson, B. L. (2021) 'The inequity of educational opportunity during an epidemic'. *Postmodern Openings*, 12(1), pp. 319–327. <https://doi.org/10.18662/po/12.1/263>
- Javorsky, K., Alley, K. M., Miller, N. C., Fondren, K. B., Lemley, S. M. and Hopper, P. F. (2022) *Mississippi's K-12 public education response to the COVID-19 pandemic: From remote learning to recovery and beyond*. Mississippi State University. Available at: <https://scholarjunction.msstate.edu/coe-publications/29/> [Accessed 20 July 2022]
- Karaman, S. and Kurşun, E. (2020) 'Açık ve uzaktan öğrenmeye yönelik kavramsal bir çerçeve önerisi: Bütüncül Analitik Katman (BAK) çerçevesi' [A conceptual framework proposal for open and distance learning: The Holistic Analytical Layer Framework (HALF)]. in S. Karaman and E. Kurşun (eds.), *Uzaktan öğretimde canlı ders uygulama ilkeleri ve örnekleri* [Principles and examples of live online teaching]. Atatürk University Press.
- King, D. L., Delfabbro, P. H., Billieux, J. and Potenza, M. N. (2020) 'Problematic online gaming and the COVID-19 pandemic'. *Journal of Behavioural Addictions*, 9(2), pp. 184–186. <https://doi.org/10.1556/2006.2020.00016>
- Koç, E. S. (2021) 'Nasıl bir uzaktan eğitim? 1 yılın sonunda yapılan çalışmaların değerlendirilmesi' [What kind of distance education? Evaluation of the studies at the end of 1 year]. *International Anatolia Academic Online Journal Social Sciences Journal*, 7(2), pp. 13–26. Available at: <https://dergipark.org.tr/en/pub/iaaoj/issue/60800/892227> [Accessed 20 August 2022]
- Krippendorff, K. (1980) 'Validity in content analysis'. in E. Mochmann (ed.), *Computerstrategien für die kommunikationsanalyse*. Campus, pp. 69-112.
- Kruczek, T., Geesa, R. L., Mayes, R. D. and Odell, K. M. (2022) 'School counselor and administrator perceptions of response to the COVID-19 pandemic'. *Professional School Counseling*, 26(1c), <https://doi.org/10.1177/2156759X221134661>
- Kyllonen, P. C. (2012) 'Measurement of 21st century skills within the common core state standards'. Paper presented at the Invitational Research Symposium on Technology Enhanced Assessments, 7-8 May. Available at: https://oei.org.ar/ibertic/evaluacion/sites/default/files/biblioteca/11_measurement_of_21stcenturyskills.pdf [Accessed 28 June 2022]
- McLoughlin, C. and Lee, M. J. W. (2007) 'Social software and participatory learning: Pedagogical choices with technology affordances in the Web 2.0 era'. in *ICT: Providing Choices for Learners and Learning. Proceedings Ascilite Singapore 2007*. Singapore: Nanyang Technological University, pp. 664–675. Available at: <https://acuresearchbank.acu.edu.au/item/865qz/social-software-and-participatory-learning-pedagogical-choices-with-technology-affordances-in-the-web-2-0-era> [Accessed 25 January 2022]
- McMillan, J. H. and Schumacher, S. (2010) *Research in education: Evidence-based inquiry*. 7th edn. Pearson.
- Merriam, S. B. (1988) *Case study research in education: A qualitative approach*. Jossey-Bass.
- Ministry of National Education (MoNE) (2020) 'Bakan Selçuk, uzaktan eğitim ile yeni eğitim öğretim yılı hazırlıklarını değerlendirdi' [Minister Selçuk evaluated distance education studies and preparations for the new academic year]. Available at: <https://www.meb.gov.tr/bakan-selcuk-uzaktan-egitim-ile-yeni-egitim-ogretim-yili-hazirliklarini-degerlendirdi/haber/21523/tr> [Accessed 30 July 2022]
- Moore, M. G. and Kearsley, G. (2011) *Distance education: A systems view of online learning*. 3rd edn. Cengage Learning.
- Nisiforou, E. A., Kosmas, P. and Vrasidas, C. (2021) 'Emergency remote teaching during COVID-19 pandemic: Lessons learnt from Cyprus'. *Educational Media International*, 58(2), pp. 215-221. <https://doi.org/10.1080/09523987.2021.1930484>
- O'Keefe, L., Dellinger, J. T., Scragg, B., Amelina, N. and Mathes, J. (2020) *The state of online learning in the Kingdom of Saudi Arabia: A COVID-19 impact study for K-12*. Online Learning Consortium.
- Özbay, Ö. (2015) 'Dünyada ve Türkiye'de uzaktan eğitimin güncel durumu' [Current status of distance education in the world and Türkiye]. *Uluslararası Eğitim Bilimleri Dergisi*, 5, pp. 376-394. Available at: <https://dergipark.org.tr/en/pub/inesj/issue/40015/475774> [Accessed 20 September 2022]
- Özdoğan, A. Ç. and Berkant, H. G. (2020) 'COVID-19 pandemi dönemindeki uzaktan eğitime ilişkin paydaş görüşlerinin incelenmesi' [Examining stakeholder views on distance education during the COVID-19 pandemic period]. *Millî Eğitim Dergisi*, 49(1), pp.13–43. <https://doi.org/10.37669/milliegitim.788118>
- Parveen, K., Tran, P. Q. B., Alghamdi, A. A., Namaziandost, E., Aslam, S. and Xiaowei, T. (2022) 'Identifying the leadership challenges of K-12 public schools during COVID-19 disruption: A systematic literature review'. *Frontiers in Psychology*, 13, 875646. <https://doi.org/10.3389/fpsyg.2022.875646>
- Patton, M. Q. (2002) *Qualitative research & evaluation methods*. 3rd edn. Sage Publications.

- Rowntree, D. (1990) *Teaching Through Self-Instruction*. Nicholas Publishing.
- Saroglou, P. and Kostas, A. (2023) ,Emergency remote teaching in K–12 education during the COVID-19 pandemic: Systematic review of technology use, problems and future practices'. *European Journal of Education Studies*, 10(2). <https://doi.org/10.46827/ejes.v10i2.4651>
- Sofianidis, A., Meletiou-Mavrotheris, M., Konstantinou, P., Stylianidou, N. and Katzis, K. (2021) 'Let students talk about emergency remote teaching experience: Secondary students' perceptions on their experience during the COVID-19 pandemic'. *Education Sciences*, 11(6), Article 268. <https://doi.org/10.3390/educsci11060268>
- Suri H. (2011) 'Purposeful sampling in qualitative research synthesis2. *Qualitative Research Journal*, 11(2), pp. 63–75. <https://doi.org/10.3316/QRJ1102063>
- Telli Yamamoto, S. G. and Altun, D. (2020) 'Coronavirüs ve çevrimiçi (online) eğitimin önlenemeyen yükselişi' [The coronavirus and the rising of online education]. *Üniversite Araştırmaları Dergisi*, 3(1), pp. 25–34. <https://doi.org/10.26701/uad.711110>
- Thornton, K. (2021) 'Leading through COVID-19: New Zealand secondary principals describe their reality.' *Educational Management Administration & Leadership*, 49(3), pp. 393–409. <https://doi.org/10.1177/1741143220985110>
- Trust, T. and Whalen, J. (2021) 'K–12 teachers' experiences and challenges with using technology for emergency remote teaching during the COVID-19 pandemic', *Italian Journal of Educational Technology*, 29(2), pp. 10–25. <https://doi.org/10.17471/2499-4324/1192>
- Van Weert, T. J. (2006) 'Education of the twenty-first century: New professionalism in lifelong learning, knowledge development and knowledge sharing'. *Education and Information Technologies*, 11(3-4), pp. 217–237. <https://doi.org/10.1007/s10639-006-9018-0>
- World Health Organization. (2020) *WHO Director-General's opening remarks at the media briefing on COVID-19—11 March 2020*. Available at: <https://www.who.int/director-general/speeches> [Accessed 3 May 2021]
- Yıldırım, A. and Şimşek, H. (2013) *Sosyal bilimlerde nitel araştırma yöntemleri*. [Qualitative research methods in the social sciences]. 9th edn. Seçkin Yayıncılık.
- Yin, R.K. (2018) *Case Study Research and Applications: Design and Methods*. 6th edn. SAGE Publications.
- Yoskovitz, G. and Schechter, C. (2023) ,Agentic leadership: Principals' agency during the COVID-19 crisis.' *Leadership and Policy in Schools*, pp. 1–18. <https://doi.org/10.1080/15700763.2023.2261039>
- Zhang, Y. and Wildemuth, B. M. (2009) 'Qualitative analysis of content'. in B. M. Wildemuth (ed.), *Applications of Social Research Methods to Questions in Information and Library Science*. Libraries Unlimited, pp. 308–319.
- Zhao, Y. (2020) 'Covid-19 as a catalyst for educational change'. *Prospects*, 49(1-2), pp.29–33. <https://doi.org/10.1007/s11125-020-09477-y>

A Comprehensive Psychometric Validation of Student Online Readiness

Hugh Kellam, Jin Liu and Tranell Gilmore

College of Education, University of Texas at Arlington, TX, USA

hugh.kellam@uta.edu (Corresponding Author)

jin.liu@uta.edu

tranell.gilmore@uta.edu

<https://doi.org/10.34190/ejel.24.4.4768>

An open access article under [CC Attribution 4.0](#)

Abstract: As online and hybrid education continue to expand, institutions require valid and equitable tools to identify students who may need additional support before beginning online coursework. Although numerous online readiness assessments have been developed, relatively few have undergone comprehensive psychometric validation, and most have not established measurement invariance across demographic groups, limiting confidence in subgroup comparisons and the use of readiness assessments for equity-informed decision-making. The purpose of this study was to provide a comprehensive psychometric validation of the Online Readiness and Student Success Assessment (ORSSA), a five-factor instrument designed to measure Learning Organization, Learning Preferences, Learning Skills, Technical Readiness Access, and Technical Readiness Skills among incoming online university students. Data were collected from 2,669 undergraduate students enrolled in online degree and certificate programs at a large public R1 university in the southern United States. Confirmatory factor analysis (CFA) using weighted least squares mean and variance adjusted (WLSMV) estimation was conducted to evaluate the proposed factor structure. Additional analyses examined model-based reliability (McDonald's omega), convergent and discriminant validity, measurement invariance across gender, ethnicity, and first-generation status, and Multiple Indicators Multiple Causes (MIMIC) modeling to investigate relationships between demographic characteristics and online readiness factors. Results supported the proposed five-factor model, which demonstrated superior fit relative to alternative models, strong standardized factor loadings, excellent model-based reliability, and satisfactory evidence of convergent and discriminant validity. Measurement invariance analyses demonstrated that the ORSSA functions equivalently across gender, ethnicity, and first-generation status, supporting meaningful and fair comparisons among demographic groups. MIMIC analyses further identified significant demographic differences across several readiness dimensions, particularly in Learning Organization, Learning Skills, and Technical Readiness, while also revealing higher online learning preferences among some historically underrepresented student groups. These findings establish the ORSSA as a psychometrically robust and equitable measure of online readiness that extends beyond traditional readiness assessments by integrating comprehensive validation procedures with demographic fairness testing. The instrument provides higher education institutions with actionable information that can be used during student onboarding to identify readiness strengths and potential barriers, enabling targeted interventions, personalized student support, and data-informed strategies to improve retention and success in online learning environments.

Keywords: Online readiness, Psychometric validation, Confirmatory factor analysis, Measurement invariance, MIMIC modeling, Student success

1. Introduction

Online readiness refers to the extent to which students possess the technological resources, technical competencies, self-regulatory skills, learning preferences, and personal characteristics necessary to successfully engage in online learning environments (Hung et al., 2010; Pillay, Irving and Tones, 2007; Reyes-Millán et al., 2023). As online and hybrid education continue to expand globally, online readiness has emerged as an important construct because it has been associated with student engagement, persistence, academic performance, satisfaction, and successful adaptation to virtual learning environments (Dray et al., 2011; Joosten and Cusatis, 2020; Martin, Stamper and Flowers, 2020). Assessing readiness enables institutions to identify potential barriers to student success before coursework begins and to implement targeted interventions that support equitable participation, retention, and achievement (Wladis and Samuels, 2016; Joosten and Cusatis, 2020). Over the past two decades, numerous instruments have been developed to measure online readiness, which typically assess combinations of self-directed learning, learner control, motivation, communication competencies, technology access, technology skills, self-efficacy, and study habits through self-report survey measures (Yu and Richardson, 2015; Zhong et al., 2023). Despite their contributions, important limitations remain. Many instruments rely primarily on exploratory validation approaches rather than quantitative psychometric methods such as confirmatory factor analysis, model-based reliability estimation, and structural equation modeling (Viladrich, Angulo-Brunet and Doval, 2017; Alanoglu, Karabatak and Yang, 2025).

Furthermore, relatively few readiness instruments have established measurement invariance, or the degree to which an instrument measures the same construct in the same way across different groups, and this limits confidence in subgroup comparisons and raises questions about the fairness of interpreting readiness differences across diverse populations (Van de Schoot, Lugtig and Hox, 2015; Çimşir, Uzunboyulu and Yavuz, 2024). Existing studies have also tended to examine demographic factors independently rather than simultaneously, making it difficult to understand the complex relationships between student characteristics and readiness dimensions (Martin, Stamper and Flowers, 2020). Finally, most readiness instruments focus on describing readiness levels rather than serving as equity-oriented diagnostic tools that inform targeted student support initiatives and institutional decision-making (Barber et al., 2021; Joosten and Cusatis, 2020). These limitations highlight the need for comprehensive readiness assessments that integrate contemporary psychometric validation methods, establish fairness across diverse student populations, and provide actionable information for improving student success in online learning environments (Putnick and Bornstein, 2016; Çimşir, Uzunboyulu and Yavuz, 2024).

The Online Readiness and Student Success Assessment (ORSSA) was developed at a university in the southern United States to address these gaps by integrating comprehensive readiness dimensions with contemporary psychometric validation procedures, including measurement invariance testing and MIMIC modeling, thereby supporting both accurate assessment and equity-informed student support. The ORSSA survey is unique as a data collection tool due to its extensive ability for student demographic data collection and incorporation into a larger student success pathway. The previous paper on the ORSSA using the first 392 participants provided a preliminary validation of the survey scale and questions via principal component analysis. There was a good fit with 35 of the 41 survey questions having values in the component matrix over 0.6, which was excellent for a pilot survey. Overall, 37 questions loaded onto the five factors of learning organization, learning skills, technical readiness access, technical readiness skills, and learning preferences. These encouraging results prompted the more robust validation of the ORSSA structure undertaken in this study. However, the abovementioned gaps have not been addressed in the preliminary study. This study will go beyond this preliminary validation and provide a robust validation of the ORSSA by involving all incoming students who are enrolled in an online degree or certificate program at an R1 institution from January 2024 to May 2025. Multiple psychometric evidence was investigated, including factor structure, concurrent and discriminant validity, reliability, and measurement invariance. Finally, MIMIC modeling was conducted to gain a better understanding of how multiple demographic factors are associated with factors of ORSSA simultaneously. The study was guided by the following research questions:

RQ1: Does the ORSSA demonstrate an empirically supported factor structure consistent with its proposed theoretical framework?

RQ2: To what extent does the ORSSA exhibit adequate reliability, convergent and discriminant validity?

RQ3: Is measurement invariance of the ORSSA supported across gender, ethnicity, and first-generation status, and

RQ4: What relationships between demographic variables and the latent constructs are revealed through MIMIC modeling?

2. Literature Review

2.1 Existing Online Readiness Assessments

Online readiness assessments have evolved substantially alongside the growth of online learning in higher education. Early readiness instruments developed between 1999 and 2007 primarily focused on students' access to technology, basic computer skills, and general attitudes toward distance learning (Mattice and Dixon, 1999; Smith, Murphy and Mahoney, 2003; Watkins, Leigh and Triner, 2004). These early studies relied largely on descriptive statistics and exploratory factor analysis to identify readiness dimensions and establish preliminary evidence of reliability. As online learning became more widespread, readiness frameworks expanded to include self-directed learning, learner control, motivation, communication competencies, and technology self-efficacy, reflecting a growing recognition that success in online learning depends on both technological and academic factors (Dray et al., 2011; Yu and Richardson, 2015). More recent instruments have adopted increasingly sophisticated psychometric approaches, including confirmatory factor analysis, structural equation modeling, convergent and discriminant validity testing, and model-based reliability estimation (Alanoglu, Karabatak and Yang, 2025; Çimşir, Uzunboyulu and Yavuz, 2024). Across this evolution, several dimensions have consistently emerged as core components of online readiness, including technology access, technology skills, self-regulation,

communication competencies, and learner motivation, suggesting broad agreement regarding the foundational characteristics required for online learning success. However, important methodological gaps remain. Most studies continue to rely heavily on Cronbach's alpha as the primary indicator of reliability, relatively few have evaluated measurement invariance across demographic groups, and only a limited number have examined how multiple demographic characteristics simultaneously influence readiness dimensions (Putnick and Bornstein, 2016; Çimşir, Uzunboylu and Yavuz, 2024). Consequently, while online readiness instruments have become more comprehensive and psychometrically sophisticated over time, further work is needed to establish the fairness, validity, and practical utility of these measures across increasingly diverse student populations.

2.2 Demographic Factors and Online Readiness Assessments

Across the online readiness literature, several dimensions consistently emerge as important predictors of student success in online learning environments, including self-regulated learning, time management, learner motivation, communication competencies, technology access, and technology skills (Reyes-Millán et al., 2023). Students who possess strong organizational skills, confidence in their academic abilities, reliable access to technology, and the technical competencies needed to navigate online learning platforms are generally more likely to engage successfully in virtual learning environments and persist in their studies (Dray et al., 2011; Martin, Stamper and Flowers, 2020). These findings suggest that readiness extends beyond technical preparedness and includes the academic and behavioral capacities necessary for effective online learning.

The ORSSA was developed to reflect this multidimensional understanding of readiness through five interconnected dimensions: Learning Organization, Learning Preferences, Learning Skills, Technical Readiness Access, and Technical Readiness Skills. Learning Organization captures students' ability to manage time, prioritize coursework, and maintain productive study habits. Learning Preferences examines students' comfort and engagement with various online instructional and collaborative learning approaches. Learning Skills assess academic competencies such as reading, writing, collaboration, and information literacy. Technical Readiness Access evaluates students' access to reliable technology and internet connectivity, while Technical Readiness Skills measure students' confidence and proficiency in using digital tools required for online learning. Together, these dimensions provide a comprehensive assessment of the factors most associated with online learning success in the literature.

Although these readiness dimensions are important for all learners, research suggests that demographic characteristics may influence how students experience and develop readiness. Studies have shown that race, ethnicity, first-generation status, socioeconomic background, disability status, and prior educational experiences can affect access to technology, self-efficacy, academic preparedness, and opportunities to develop online learning competencies (Joosten and Cusatis, 2020). For example, first-generation and underrepresented minority students experienced disproportionate challenges during the COVID-19 pandemic, including caregiving responsibilities, employment obligations, limited technology access, and financial stressors that may affect readiness for online learning (Barber et al., 2021). Consequently, readiness assessments that incorporate demographic information can help institutions identify students who may benefit from targeted supports and interventions before the start of their programs (Schwam, Greenberg and Li, 2020). However, meaningful comparisons across demographic groups require evidence of measurement invariance—the degree to which an instrument measures the same construct in the same way across groups—so that observed differences can be interpreted as true differences in readiness rather than measurement bias (Putnick and Bornstein, 2016). Establishing measurement invariance is therefore a critical step in using readiness assessments to support equitable and data-informed decision-making.

3. Materials and Methods

3.1 Development of the Online Readiness and Student Success Assessment (ORSSA) Survey

A Likert scale survey design was used to obtain a numeric description of opinions (Creswell and Creswell, 2017) to measure the general trends and attitudes of first-year university students regarding their readiness for online learning. A new assessment tool, 'The Online Readiness and Student Success Assessment (ORSSA)' was developed, revised and validated (Kellam et al., 2025). The scale design involved identifying the five factors from a review of the research literature, developing a set of questions for each factor, refining the survey questions, and piloting the scale. The survey was then distributed in an exploratory study to 392 students who participated in the online orientation modules from January to October 2024, and an initial Principal Component Analysis was performed that identified the potential of the survey and a need for a more rigorous quantitative study with an increased number of participants. The list of five ORSSA factors (Kellam et al., 2025) and their associated

questions can be found in Table 1. In the current study, the survey was examined for validation with psychometric evidence with a large sample (see the information in Data Analysis and Results).

Table 1: Factors and Associated ORSSA Likert Questions

Factor	Likert Questions
1 Learning Organization	LO1 - I am able to prioritize at least 9 hours per week for each online class.
	LO2 - I have a quiet, dedicated place to work on my online classes and studying.
	LO3 - I am usually able to stay on task and avoid distractions while studying (common distractions include texting, social media, web surfing, and multi-tasking).
	LO4 - I consistently finish and submit work on time and complete.
	LO5 - I use software to maintain to-do lists for my class and university projects and deadlines.
	LO6 - I understand the importance of checking school email at least 3 times per week during the semester.
	LO7 - I finish the projects I start.
	LO8 - When I encounter a technical issue while learning online, I persevere until it gets resolved.
	LO9 - I can depend on the support of family and friends while completing my online program.
2 Technical Readiness Skills	TRS1 – I am comfortable conducting internet searches.
	TRS2 - I am comfortable opening files on my computer such as documents, presentations, spreadsheets.
	TRS3 - I am comfortable using videoconferencing software.
	TRS4 - I am comfortable participating in online polls or surveys.
	TRS5 - I am comfortable participating in online discussion forums.
	TRS6 - I am comfortable creating and sharing online videos or presentations.
	TRS7 - I am comfortable using virtual support to resolve a technical issue.
3 Learning Preferences	LP1 - I enjoy learning through online lectures (recorded by the professor).
	LP2 - I enjoy learning through online classes (delivered live via videoconference).
	LP3 - I enjoy learning through videos.
	LP4 - I enjoy learning through podcasts or audio recordings.
	LP5 - I enjoy learning through online readings.
	LP6 - I enjoy learning through discussions with peers.
	LP7 - I enjoy collaborating virtually through videoconference.
	LP8 - I enjoy collaborating virtually through email.
	LP9 - I enjoy collaborating virtually through online whiteboards (like Microsoft Whiteboard).
	LP10 - I enjoy collaborating virtually through online word processing (like MS Word).
4 Technical Readiness Access	TRA1 - I have reliable access to a high-speed Internet connection (recommended minimum download speed of 25 Mbps and upload speed of 3 Mbps).
	TRA2 - I have access to a webcam and microphone for multimedia participation, either on my computer or my smartphone.
	TRA3 - My computer runs reliably on Windows or on Mac OS.
	TRA4 - My browser can play common multimedia (video and audio) formats such as mp3, mp4, mov files.
	TRA5 - I have experience taking online tests using lockdown browsers (extension often used to limit web access during exams).

Factor	Likert Questions
5 Learning Skills	LS1 - When working on a class project independently, I am comfortable performing online academic research utilizing tools such as library databases or Google Scholar.
	LS2 - I am comfortable working on a collaborative online project with classmates.
	LS3 - I am confident in my ability to carefully read written instructions without missing key information.
	LS4 - I am confident in my ability to write at a post-secondary level (university level).
	LS5 - I am confident in my ability to write and respond to emails.
	LS6 - I learn from constructive feedback given by my instructor or my classmates.
	LS7 - I respect academic integrity and the university honor code when completing my online work.

3.2 Sample and Data Collection

This IRB-approved study collected quantitative data using the ORSSA survey at a large public university in the southern United States. All incoming online undergraduate students in all Colleges were required to participate in six online orientation modules created by the Office of Student Success. The modules provided information on campus resources, contacting orientation leaders, academic advising, and completing orientation documents. The fifth module focused on online readiness skills and included the ORSSA survey that began with a consent page where students were provided with details of the research study and could give their consent. Students who refused to consent could still participate in the ORSSA survey, but their data were not collected. Data were collected from incoming online students from January 2, 2024, through January 30, 2025, which included a total of 2884 undergraduate students. A total of 215 students with incomplete responses in the Likert questions were omitted from the data analysis.

In the end, a total of 2,669 undergraduate students were included in the sample. The majority of them were female (83.0%, $n = 2,214$), with males comprising 17.0% ($n = 455$). More than three-quarters of participants were first-generation college students or students whose parents or legal guardians did not complete a four-year bachelor's degree (76.7%, $n = 2,046$) and 23.3% ($n = 623$) were non-first generation. Regarding ethnicity, 33.6% ($n = 898$) identified as Hispanic/Latino, 28.7% ($n = 766$) as White, 22.6% ($n = 604$) as Black/African American, and 15.0% ($n = 401$) as Other.

3.3 Data Analysis

We first examined the factor structure of the ORSSA using confirmatory factor analysis (CFA), allowing correlations among factors. Analyses were conducted in Mplus 8.11 using WLSMV, the default estimation method for ordered categorical (Likert-scale) data (Muthén and Muthén, 1998–2010). Confirmatory factor analysis was conducted to confirm the five correlated factors under the proposed scale: learning organization, technical readiness skills, learning preference, technical readiness access, and learning skills. We also examined the one-factor model as we expected to have a general factor indicating the overall readiness level. As the chi-square statistics were highly sensitive to sample size, model evaluation relied primarily on commonly recommended fit indices: the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). Following Kline (2016), cutoffs for adequate fit were defined as $CFI \geq .90$, $TLI \geq .90$, $RMSEA \leq .08$, and $SRMR \leq .10$. More stringent thresholds ($CFI \geq .95$, $TLI \geq .95$, $RMSEA \leq .06$, $SRMR \leq .08$) were considered to indicate good model good fit (Hu and Bentler, 1999). In addition, factor loadings .50 is preferable for practical significance and .70 or above indicates strong item–factor association (Hair et al., 2019). Only factor loadings over .70 were included in this study.

After the factor structure was confirmed, we evaluated the model-based reliability (i.e., omega values) for each subscale. Following conventional guidelines, values $\geq .70$ were considered acceptable, $\geq .80$ as good, and $\geq .90$ as excellent (Dunn, Baguley, and Brunsden, 2014; Nunnally and Bernstein, 1994). Convergent validity was evaluated using the Average Variance Extracted (AVE), which was the proportion of variance in the observed indicators accounted for by the specific factor. Values higher than .50 are considered adequate to demonstrate convergent validity, while discriminant validity requires that the square root of the AVE for each factor is higher than its factor correlations (Fornell and Larcker, 1981; Hair et al., 2019).

Next, we conducted measurement invariance tests to confirm that the validated factor structure functioned equivalently across demographic groups. Specifically, we evaluated configural invariance (same factor

structure), metric invariance (same structure and factor loadings), and scalar invariance (same structure, loadings, and item thresholds) across gender, ethnicity, and first-generation status. For testing measurement invariance, we compared the configural, metric, and scalar models using scaled chi-square difference tests (via the *DIFFTEST* option in Mplus) to evaluate whether invariance could be supported by progressively adding equality constraints across groups. While a nonsignificant chi-square difference ($p > .05$) traditionally indicates invariance, prior research has noted that this criterion is overly stringent and often unrealistic with large samples (Chen, Sousa and West, 2005). Therefore, we also relied on changes in alternative fit indices—RMSEA, CFI, and SRMR—when evaluating model invariance. In particular, non-invariance is indicated when in CFI is greater than $-.010$, accompanied by an increase of $\geq .015$ in RMSEA or $\geq .010$ in SRMR (Chen, 2007). It was noted that these criteria may be misleading when the estimator was WLSMV and we may evaluate the model fit indices directly following the criteria above (Sass, Schmitt and Marsh, 2014).

Finally, a Multiple Indicators Multiple Causes (MIMIC) model was constructed to include gender, ethnicity, and first generation as predictors of the latent constructs, enabling us to evaluate the extent to which these observed variables influence the underlying readiness factors. MIMIC models embed these covariates directly as predictors of latent variables within a confirmatory factor modeling framework, allowing for simultaneous estimation of measurement and structural components (Wang, 2022). This approach offers a parsimonious alternative to multiple-group CFA, especially when including numerous covariates simultaneously. The same model fit criteria of CFA were applied to evaluate the quality of the MIMIC modeling. We focused on explaining the relationships between demographic variables and latent factors. The following groups were considered the reference group in results interpretation: first generation students, white and female.

4. Results

The first set of analysis was conducted to find the optimal factor structure of the ORSSA. The one factor CFA demonstrated poor model fit: ($\chi^2_{(702)} = 39222.699$, RMSEA = .143, 90% CI .142 - .145, CFI = .783, TLI = .771, SRMR = .126). The five-factor CFA model showed acceptable to good fit with indices above the cut-off values ($\chi^2_{(692)} = 10592.094$, RMSEA = .073, 90% CI .072 - .074, CFI = .944, TLI = .940, SRMR = .048). Therefore, the five-factor model was selected as the optimal model due to the better model fit and alignment with the theoretical framework. The completely standardized loadings of the final solution are displayed in Table 2. Item descriptive statistics can be found in Table 3. All items were significantly loaded on the principal factor, and all factor loadings were higher than 0.71. Factor correlations ranged from .391 to .896. Omega estimates ranged from .911 to .932, suggesting high internal consistencies of all subscales. Evidence for convergent validity was supported, as AVE values were above .628 for all subscales. Regarding discriminant validity, three subscales had square root AVE values greater than their inter-factor correlations. However, high correlations between Learning Skills and Learning Organizations resulted in slightly lower discriminant validity for the five-factor models. After further evaluating the four- and five-factor alternatives, we selected the five-factor model due to its conceptual coherence, better model fit and its ability to reveal unique associations with demographic predictors in subsequent MIMIC analyses (see below).

Measurement invariance was examined across gender, ethnicity, and first-generation status (Table 4). While the scaled chi-square tests showed significant results at each step given their sensitivity to a large sample size, all testing models showed good model fit in terms of CFI, TLI, and RMSEA. Moreover, CFI and RMSEA values improved slightly from the configural to the metric and scalar models, and SRMR values decreased while remaining within the acceptable range. Taking together, these results provide strong evidence for measurement invariance across demographic groups. Then the demographic variables were added to the model. The final model showed acceptable to good fit: ($\chi^2_{(862)} = 10765.891$, RMSEA = .066, 90% CI .065 - .067, CFI = .949, TLI = .944, SRMR = .046). The MIMIC model included significant regression coefficients from each covariate considered in the model and the identified factors (Figure 1).

Gender was significantly associated with learning organization (standardized $b = -.075$, $p < .001$), indicating that female reported higher learning organization levels than male students. First-generation status was significantly related to three of the five subscales: non-first-generation students demonstrated higher Learning Skills ($\beta = .043$, $p < .001$), Technical Readiness Access ($\beta = .065$, $p < .001$), and Technical Readiness Skills ($\beta = .048$, $p < .001$) compared to first-generation students.

Using White students as the reference group, Hispanic students reported lower levels of Learning Organization ($\beta = -.056$, $p = .030$), Learning Skills (standardized $b = -.052$, $p = .040$), Technical Readiness Access ($\beta = -.063$, $p = .017$) and Technical Readiness Skills ($\beta = -.051$, $p = .049$). A similar pattern was observed for students classified as "Other" compared to White students, with significantly lower scores on Learning Organization ($\beta = -.090$, $p < .001$).

.001), Learning Skills ($\beta = -.087$, $p < .001$), Technical Readiness Access ($\beta = -.085$, $p = .001$), and Technical Readiness Skills ($\beta = -.081$, $p = .001$).

Interestingly, both Hispanic ($\beta = .076$, $p = .001$) and Black students ($\beta = .151$, $p < .001$) reported higher levels of Online Learning Preference compared to White students. Finally, Black students also reported significantly lower Technical Readiness Access than White students ($\beta = -.096$, $p < .001$) while there were no significant differences for three other subscales including Learning Organization, Learning Skills and Technical Readiness Skills.

Table 2: Standardized Factor Loadings, Omega, AVE and Factor Correlations

Items	LO	LP	LS	TRA	TRS
1	0.833	0.807	0.807	0.914	0.888
2	0.847	0.720	0.745	0.941	0.913
3	0.774	0.846	0.878	0.935	0.885
4	0.873	0.727	0.831	0.945	0.892
5	0.728	0.737	0.917	0.827	0.899
6	0.889	0.714	0.882		0.853
7	0.896	0.785	0.794		0.874
8	0.834	0.797	0.868		
9	0.743	0.818			
10		0.853			
Omega	0.913	0.917	0.911	0.923	0.932
AVE	0.709	0.628	0.736	0.858	0.790
Square Root AVE	0.842	0.792	0.858	0.926	0.889
Factor Correlations					
LO	1	0.708	0.896	0.557	0.672
LP	0.708	1	0.678	0.391	0.558
LS	0.896	0.678	1	0.522	0.719
TRA	0.557	0.391	0.522	1	0.631
TRS	0.672	0.558	0.719	0.631	1

Table 3: Item Descriptive Statistics

Item	Mean	SD	Skewness	Kurtosis
LO1R	4.361	0.737	-1.342	2.808
LO2R	4.433	0.705	-1.443	3.235
LO3R	4.136	0.858	-1.004	1.048
LO4R	4.455	0.666	-1.292	2.773
LO5R	4.040	0.956	-0.839	0.140
LO6R	4.546	0.619	-1.553	4.258
LO7R	4.486	0.630	-1.254	2.865
LO8R	4.381	0.713	-1.268	2.606
LO9R	4.369	0.810	-1.456	2.534
LP1R	4.254	0.825	-1.191	1.778
LP2R	3.907	0.982	-0.717	0.095
LP3R	4.303	0.754	-1.193	2.273
LP4R	3.976	0.966	-0.843	0.344
LP5R	3.871	0.997	-0.695	0.005

Item	Mean	SD	Skewness	Kurtosis
LP6R	3.814	0.995	-0.630	0.005
LP7R	3.701	1.040	-0.541	-0.206
LP8R	3.851	0.952	-0.652	0.210
LP9R	3.738	0.931	-0.321	-0.167
LP10R	3.889	0.900	-0.509	-0.005
LS1R	4.312	0.773	-1.147	1.598
LS2R	4.067	0.903	-0.961	0.859
LS3R	4.383	0.674	-1.049	1.918
LS4R	4.244	0.776	-0.982	1.212
LS5R	4.508	0.621	-1.293	2.973
LS6R	4.471	0.613	-1.067	2.336
LS7R	4.250	0.813	-1.097	1.342
LS8R	4.629	0.580	-1.827	5.605
TRA1R	4.413	0.953	-2.229	5.133
TRA2R	4.450	0.942	-2.338	5.645
TRA3R	4.426	0.967	-2.221	4.917
TRA4R	4.393	0.956	-2.133	4.725
TRA5R	4.214	1.127	-1.575	1.676
TRS1R	4.572	0.727	-2.392	7.691
TRS2R	4.555	0.729	-2.242	6.784
TRS3R	4.302	0.868	-1.375	2.060
TRS4R	4.468	0.767	-1.858	4.638
TRS5R	4.414	0.810	-1.713	3.660
TRS6R	4.118	0.982	-1.038	0.548
TRS7R	4.280	0.850	-1.294	1.944

Table 4: Test of Measurement Invariance across Gender, First Generation Status and Ethnicity

Models	χ^2 (df)	RMSEA	Diff χ^2 (df) p-value	CFI	TLI	SRMR
Gender						
Configural Model	9815.745 (1384)	.068		.956	.953	.050
Metric Model	6378.693 (1423)	.074	93.847 (39) <.001	.974	.973	.051
Scalar Model	6004.885 (1540)	.047	267.742(117) <.001	.977	.978	.051
First Generation						
Configural Model	10382.020 (1384)	.070		.949	.946	.050
Metric Model	6494.247 (1423)	.052	60.413(39) =.001	.971	.970	.051
Scalar Model	5946.287 (1540)	.046	194.441(117) <.001	.975	.976	.051
Ethnicity						
Configural Model	11157.934 (2768)	.067		.953	.950	.052

Models	χ^2 (df)	RMSEA	Diff χ^2 (df) p-value	CFI	TLI	SRMR
Metric Model	7647.934 (2885)	.050	255.444(117) < .001	.973	.973	.055
Scalar Model	7921.370 (3236)	.047	948.915(351) <.001	.974	.976	.056

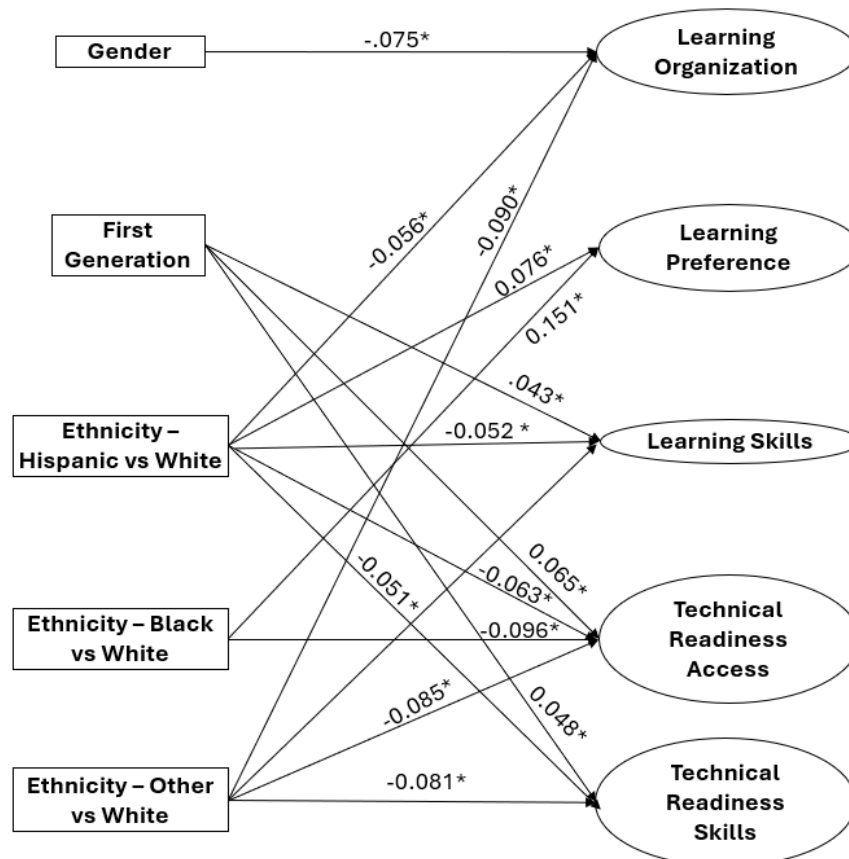


Figure 1: MIMIC Modeling Results

5. Discussion

This study investigated the underlying factor structure, model-based reliability, discriminant and convergent validity, and measurement invariance of gender, ethnicity, and first generation of the ORSSA for undergraduate students who enroll in online courses. After measurement invariance was established, we further investigated the complex relationships between the demographic factors and five subscales of the ORSSA: Learning Organization, Learning Preferences, Learning Skills, Technical Readiness Access, and Technical Readiness Skills.

To evaluate the most appropriate latent structure, both four-factor and five-factor models were estimated and compared using global fit indices. Although both models demonstrated acceptable fit, the five-factor model consistently outperformed the four-factor model across all major fit indices. Specifically, the five-factor model demonstrated lower RMSEA (.073 vs. .077) and SRMR (.040 vs. .051), along with higher CFI (.944 vs. .939) and TLI (.940 vs. .935), indicating improved model fit. Relative to the four-factor model, the five-factor solution resulted in improvements of $\Delta CFI = +.005$, $\Delta TLI = +.005$, $\Delta RMSEA = -.004$, and $\Delta SRMR = -.011$. Collectively, these findings suggest that distinguishing Learning Organization and Learning Skills as separate but related constructs provides a more accurate representation of online readiness than combining them within a four-factor structure. Therefore, the five-factor model was retained for all subsequent analyses because it demonstrated superior empirical fit while maintaining greater theoretical interpretability. The five-factor solution was consistent with the principal component analysis results reported by Kellam et al. (2025), supporting the stability of the factor

structure. Unlike PCA, the present study employed confirmatory factor analysis (CFA), which provided formal model-fit indices to evaluate competing models. CFA allows researchers to test the adequacy of theoretically specified models and assess model fit using multiple statistical criteria, thereby providing stronger support for construct validity and model selection (Brown, 2015). This approach enhances the robustness of the findings and establishes a solid foundation for future comprehensive investigations of online learning readiness.

Although the standardized coefficients were relatively small ($\beta \approx .04-.15$), their consistency across readiness dimensions suggests meaningful patterns at the population level. The finding that Hispanic and Black students reported stronger preferences for online learning while simultaneously reporting lower levels of technical access aligns with broader digital divide research, indicating that enthusiasm for online learning does not necessarily correspond with equitable access to technological resources. These findings suggest that readiness interventions should distinguish between students' willingness to engage in online learning and their ability to access the technological infrastructure required for success. Institutions may therefore benefit from targeted technology-access supports while simultaneously leveraging students' positive attitudes toward online learning.

Factor correlations ranged from low to moderate, indicating that the dimensions capture distinct aspects of the construct. However, the discriminant validity between the two highly correlated factors—Learning Skills and Learning Organization—was limited. Conceptually, these domains overlap, suggesting that a four-factor alternative model, combining these two dimensions, may provide a more parsimonious representation. While we selected the five-factor model based on the theoretical framework and model fit indices, future research should test and cross-validate both the five-factor and four-factor models across different settings. Model-based reliability estimates were provided to indicate the internal consistency of all subscales. Unlike coefficient alpha, model-based reliability incorporates factor loadings and measurement errors from the CFA model, providing a more accurate estimate of internal consistency (Raykov, 1997). Next, as most prior studies have not examined measurement invariance across subgroups, the study findings contribute important evidence to measurement of online readiness levels. Although the importance of measurement invariance is widely recognized, it is often overlooked in applied measurement research (Putnick and Bornstein, 2016). Measurement invariance was tested across students' gender, ethnicity, and first-generation status to ensure meaningful group comparisons. The decent model fit after imposing equality constraints supported the presence of measurement invariance across subgroups despite the significant chi-square difference tests due to the large sample size.

5.1 Practical Implications

This study identified the value of administering an online readiness instrument at orientation that includes the five ORSSA factors, and the importance of collecting disaggregated race/ethnicity data to identify equity gaps at the beginning of college/university studies. It is also important to use a validated tool to perform invariance checks to make group comparisons. Few peer-reviewed higher-education studies report ethnic technology access results directly and more work should present race-specific (not just the broad underrepresented minority) estimates for each factor of the ORSSA.

Universities should treat computer access/reliability and broadband availability as first-order academic needs, as these issues persist for undergraduate students. Also, online-learning skills should be embedded in orientation and gateway courses. Short, graded micro-modules on LMS navigation, time management, help-seeking, and environmental structuring, with practice tasks and nudges, benefit all students and correlate with motivation (Michikyan et al., 2025). Universities must partner with libraries to assess and teach online skills and literacy in first-year seminars/comp courses, and target FGCS sections for extra support (Lemire, 2025). Professors can provide quick checks on technology reliability and skills for their first-year classes, and shouldn't assume national patterns exactly fit their campus, as large studies show variability (Michikyan et al., 2025). A final reminder is that FGCS often have less exposure to technology during their high school years due to socioeconomic factors, leading to a lack of foundational skills in using online educational tools, which can compound their challenges in online learning environments (Wang, 2022).

A notable finding was the strong correlation between Learning Organization and Learning Skills ($r = .896$), indicating substantial shared variance (approximately 80%). Although the two constructs were theoretically specified as distinct dimensions, the magnitude of the relationship suggests that they may represent closely related aspects of academic self-regulation in online learning environments. From a psychometric perspective, this finding raises the possibility of a higher-order readiness factor or a more parsimonious four-factor structure. Nevertheless, the five-factor solution demonstrated superior conceptual interpretability and allowed the identification of unique demographic relationships in subsequent MIMIC analyses. Future research should

directly compare four-factor, five-factor, higher-order, and bifactor representations across independent samples to determine the most appropriate structural model.

The current study focused on correlated factor models based on the theoretical framework. Higher-order or bifactor models were not examined because they address a different question: whether factor correlations can be explained by a general higher-order factor or whether item responses are explained by a general factor. The intended interpretation of the scale emphasizes the sub-dimensional factors rather than a single general factor. In addition, the factor correlations were not uniformly high, ranging from .391 to .896. Most factor correlations were low to moderate. This pattern suggests that the dimensions are related but not consistently reducible to a general/higher order factor.

6. Limitations and Future Research Directions

Several limitations should be considered when interpreting the findings. First, the sample was drawn from a single R1 institution in the southwestern United States, which may limit generalizability to other institutional contexts. Second, the sample was predominantly female (83%), potentially affecting the stability of gender-based comparisons despite evidence of measurement invariance. Third, students who declined research consent were excluded from the analytic sample, introducing the possibility of consent bias if consenting and non-consenting students differed systematically. Fourth, the study focused on psychometric validation and did not evaluate predictive validity; future research should examine the extent to which ORSSA scores predict academic performance, retention, and online course success. Finally, this study provided a broad understanding on online readiness across gender, ethnicity and FGCS, a contextual, qualitative study to comprehend specific challenges across the five ORSSA factors is warranted.

7. Conclusion

The ORSSA meets core standards for psychometric quality and fairness, providing a defensible mechanism for interpreting online readiness across diverse student groups. Its validated five-factor structure, strong reliability, and demonstrated measurement invariance enable institutions to compare groups responsibly and to identify where student resources and support programs will have the greatest impact. The demographic patterns observed through the MIMIC modeling point to practical priorities: address technology access and skill gaps as first-order academic needs; embed short, just-in-time modules on time management, LMS navigation, help-seeking, and study environments in orientation and gateway courses; partner with libraries for foundational digital literacy; and provide additional scaffolds for first-generation students. Used diagnostically rather than as a gatekeeper, the ORSSA can equip universities to deliver targeted interventions that improve students' readiness for and success in online learning. While the ORSSA demonstrated strong psychometric properties, the high correlation between Learning Organization and Learning Skills warrants continued investigation. Furthermore, although demographic effects were statistically significant, their practical magnitude was generally small, suggesting that readiness differences across groups should be interpreted cautiously. Replication across institutions and examination of predictive validity will be important next steps in establishing the broader utility of the instrument.

Ethics Statement: The study protocol was reviewed and approved by the university's Institutional Review Board. All participants provided written informed consent prior to their inclusion in the study, and the research was conducted in accordance with the ethical standards of the 1964 Declaration of Helsinki and its later amendments.

AI Statement: Artificial intelligence was not used in any way in the preparation of this paper.

References

- Alanoglu, M., Karabatak, S. and Yang, H. (2025). Understanding university students' self-directed online learning in the context of emergency remote teaching: the role of online learning readiness and digital literacy, *Journal of Computing in Higher Education*, 38(2), pp. 651–677. doi:10.1007/s12528-025-09458-0
- Barber, P.H., Shapiro, C., Jacobs, M.S., Avilez, L., Brenner, K.I., Cabral, C., Cebreneros, M., Cosentino, E., Cross, C., Gonzalez, M.L., Lumada, K.T., Menjivar, A.T., Narvaez, J., Olmeda, B., Phelan, R., Purdy, D., Salam, S., Serrano, L., Velasco, M.J., Marin, E.Z. and Levis-Fitzgerald, M. (2021). Disparities in remote learning faced by first-generation and underrepresented minority students during COVID-19: Insights and opportunities from a remote research experience. *Journal of Microbiology & Biology Education*, 22(1). doi:10.1128/jmbe.v22i1.2457.
- Chen, F. F., Sousa, K. H. and West, S. G. (2005). Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, 12(3), pp. 471-492. doi: 10.1207/s15328007sem1203_7

- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), pp. 464–504. doi:10.1080/10705510701301834
- Çimşir, E., Uzunboyulu, H. and Yavuz, M. (2024). Development and validation of an online learning readiness self-check instrument. *Education and Information Technologies*, 29(4), pp. 4815–4838.
- Creswell, J. W., and Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.
- Dray, B. J., Lowenthal, P. R., Miszkiewicz, M. J., Ruiz-Primo, M. A. and Marczyński, K. (2011). Developing an instrument to assess student readiness for online learning: A validation study. *Distance Education*, 32(1), pp. 29–47. doi:10.1080/01587919.2011.565496
- Dunn, T. J., Baguley, T. and Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), pp. 399–412. doi:10.1111/bjop.12046
- Fornell, C. and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), pp. 39–50. doi:10.1177/002224378101800104
- Hair, J. F., Hult, G. T. M., Ringle, C. M. and Sarstedt, M. (2019). A primer on partial least squares structural equation modeling (PLS-SEM) (2nd ed.). Sage.
- Hu, L. and 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), pp. 1–55. doi:10.1080/10705519909540118
- Hung, M. L., Chou, C., Chen, C. H. and Own, Z. Y. (2010). Learner readiness for online learning: Scale development and student perceptions. *Computers & Education*, 55(3), pp. 1080–1090. doi:10.1016/j.compedu.2010.05.004
- Joosten, T. and Cusatis, R. (2020). Online learning readiness. *American Journal of Distance Education*, 34(3), pp. 180–193. doi:10.1080/08923647.2020.1726167
- Kellam, H., Jennings, L., Magruder, K., Smallwood, A., Purgason, A., & Richter, L. (2025). An analysis of the online readiness and student success assessment. *American Journal of Distance Education*, pp. 1–21. doi:10.1080/08923647.2025.2555643
- Kline, R. B. (2016). Principles and practice of structural equation modeling (4th ed.). Guilford Press.
- LeMire, S. (2025). Faculty barriers to using open educational resources. *Open Learning: The Journal of Open, Distance and e-Learning*, pp.1-19. doi:10.1080/02680513.2025.2573338
- Martin, F., Stamper, B. and Flowers, C. (2020). Examining student perception of readiness for online learning: Importance and confidence. *Online Learning*, 24(2), pp. 38–58. doi:10.24059/olj.v24i2.2053
- Mattice, N. J. and Dixon, P. S. (1999). Student preparedness for distance education. *Proceedings of the Annual Distance Education Conference*, pp. 211–216.
- Michikyan, M., Subrahmanyam, K., Regan, P., Cano, M.Á., Castillo, L.G., Harkness, A. and Schwartz, S.J. (2025). Differences in and relationship between digital access, use, connectivity, skill, engagement, and motivation: implications for digital inequality in online learning among first-generation and continuing-generation undergraduate students. *Discover Education*, 4(1), pp. 9.
- Muthén, L. and Muthén, B.O. (1998–2010). Mplus users guide (6th ed.). Muthén & Muthén.
- Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). McGraw-Hill.
- Pillay, H., Irving, K. and Tones, M. (2007). Validation of the diagnostic tool for assessing tertiary students' readiness for online learning. *Higher Education Research & Development*, 26(2), pp. 217–234. doi:10.1080/07294360701310821
- Putnick, D. L. and Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41(2), pp. 71–90. doi:10.1016/j.dr.2016.06.004
- Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, 21(2), pp. 173–184. doi:10.1177/01466216970212006
- Reyes-Millán, M., Villareal-Rodríguez, M., Murrieta-Flores, M. E., Bedolla-Cornejo, L., Vázquez-Villegas, P. and Membrillo-Hernández, J. (2023). Evaluation of online learning readiness in the new distance learning normality. *Heliyon*, 9(11). doi:10.1016/j.heliyon.2023.e22070
- Sass, D. A., Schmitt, T. A. and Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling*, 21(2), pp. 167–180. doi:10.1080/10705511.2014.882658
- Schwam, D., Greenberg, D. and Li, H. (2020). Individual differences in self-regulated learning of college students enrolled in online college courses. *American Journal of Distance Education*, 35(2), pp. 133–151. doi:10.1080/08923647.2020.1829255
- Smith, P. J., Murphy, K. L. and Mahoney, S. E. (2003). Towards identifying factors underlying readiness for online learning: An exploratory study. *Distance Education*, 24(1), pp., 57–67. doi:10.1080/0158791030303043
- Van de Schoot, R., Lugtig, P. and Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), pp. 486–492. doi:10.1080/17405629.2012.686740
- Viladrich, C., Angulo-Brunet, A. and Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Anales de Psicología*, 33(3), pp. 755–782. doi:10.6018/analesps.33.3.268401
- Wang, Z. (2022). A comparison of three approaches to covariate effects on latent factors. *Large-scale Assessments in Education*, 10(1). doi:10.1186/s40536-022-00148-2
- Watkins, R., Leigh, D. and Triner, D. (2004). Assessing readiness for e-learning. *Performance Improvement Quarterly*, 17(4), pp. 66–79. doi:10.1111/j.1937-8327.2004.tb00321.x

- Wladis, C. and Samuels, J. (2016). Do online readiness surveys do what they claim? Validity, reliability, and subsequent student enrollment decisions. *Computers & Education*, 98(1), pp. 39–56. doi:10.1016/j.compedu.2016.03.001
- Yu, T. and Richardson, J. C. (2015). An exploratory factor analysis and reliability analysis of the Student Online Learning Readiness (SOLR) instrument. *Online Learning*, 19(5), pp. 120–141. doi:10.24059/olj.v19i5.593
- Zhong, L., Wang, X., Yang, W., & Feng, X. (2023). Reliability and validity assessment of the Chinese version of the online learning readiness scale (OLRS) for nursing students. *Nurse Education Today*, 128. doi:10.1016/j.nedt.2023.105884

Technology-enhanced Simulation for Developing Nursing Competencies and Digital Intelligence

Prakob Koraneekij¹ and Sresuda Wongwiseskul²

¹Center of Excellence in Educational Invention and Innovation, Department of Educational Technology and Communications, Faculty of Education, Chulalongkorn University, Thailand

²Department of Community Nursing, Faculty of Nursing, Suan Dusit University, Thailand

prakob.k@chula.ac.th (Corresponding author)

sresuda_won@dusit.ac.th

<https://doi.org/10.34190/ejel.24.4.4134>

An open access article under [CC Attribution 4.0](#)

Abstract: In the context of digital disruption, nursing education requires approaches that integrate digital intelligence (DI) with professional competencies while addressing limitations of traditional simulation-based learning (SBL), which is often resource-intensive and constrained in large classes. This study developed and evaluated “Nurse Sims,” a progressive web application that integrates with an Internet of Things (IoT)-enabled humanoid manikin (ESP32) to support real-time clinical training and automated feedback for nursing students. A four-phase research and development design was employed. Phase 1 assessed learning needs among 462 nursing students in Thailand using questionnaires and PNIModified analysis. Phase 2 involved the development of the SBL model and technological system. Phase 3 evaluated the intervention with 31 nursing students and six instructors using repeated-measures analysis of variance across three assessment points (Situations 1, 3, and 5). Phase 4 involved expert validation of the innovation. Results show a high need for SBL (PNIModified = 0.796). Significant improvements were observed in professional competencies ($\eta^2 = .911$) and DI ($\eta^2 = .933$) ($p < .001$). Students reported high acceptance of the system ($M = 4.52$, $SD = 0.49$), while experts rated the innovation as highly appropriate ($M = 4.76$, $SD = 0.20$). Qualitative feedback indicated improved digital self-efficacy and engagement with technology-enhanced learning. The findings suggest that IoT-integrated simulation systems can enhance nursing competencies, and DI and may serve as a scalable model for nursing education in resource-constrained settings.

Keywords: Learning innovation, Simulation-based learning, Disruptive educational technology, Nursing competencies, Digital intelligence

1. Introduction

Nursing education plays a critical role in preparing healthcare professionals to deliver safe, effective, and compassionate patient care (Pleshkan, 2025; Hou et al., 2025; Laugaland et al., 2025; Li & Cheung, 2024). As healthcare systems become increasingly complex and digitally driven, nursing education must evolve to incorporate technology-enhanced pedagogies that strengthen professional competencies and align with the demands of contemporary clinical practice (Hou et al., 2025; Hamilton, 2025; Jallad et al., 2024).

Technological advancement has expanded opportunities for innovation in nursing education. In particular, nurses are now expected to demonstrate digital intelligence (DI), which encompasses the competencies required to effectively use digital tools, interpret data, and navigate technology-enabled healthcare environments. DI is increasingly recognized as essential for ensuring safe and high-quality patient care. Emerging educational technologies such as mobile-assisted learning (Lim et al., 2024), generative artificial intelligence (AI) applications, virtual reality environments (Hou et al., 2025), and simulation robotics (Sighinolfi et al., 2023) have therefore gained attention as pedagogical tools. Among these approaches, simulation-based learning (SBL) has emerged as a particularly effective method for developing nursing competencies. SBL provides a safe, controlled, and experiential learning environment in which students can develop clinical knowledge, psychomotor skills, critical thinking, and decision-making abilities essential for high-quality patient care (Ylönen et al., 2025; Roettger et al., 2025; Thompson et al., 2024). In addition, SBL supports the development of DI by enabling self-directed, technology-mediated learning that can occur across time and space (Rungrangtanapol & Khlaisang, 2021).

DI is increasingly recognized as a key factor influencing educational outcomes in nursing. It has been shown to enhance academic motivation, lifelong learning capacity, and creative thinking, particularly in contexts involving AI and immersive digital environments such as the metaverse (Agaoglu et al., 2025; Amin et al., 2025).

Despite these benefits, two major challenges persist. First, SBL remains resource-intensive, placing significant physical and cognitive demands on instructors and limiting scalability (Saito and Dal Sasso, 2023). Furthermore, large class sizes often reduce student engagement, with many learners adopting passive roles during simulation activities (Guinea et al., 2019). Second, there remains a persistent gap between nursing curricula and the digital competencies required in contemporary healthcare practice, particularly in relation to DI development (Agaoglu et al., 2025; Amin et al., 2025; Stunden et al., 2026). These challenges highlight an urgent need for disruptive educational approaches that reduce instructional burden, increase active learner participation, and provide real-time feedback while simultaneously strengthening both clinical competencies and DI. In response, this study explores the integration of SBL with disruptive technologies to enhance nursing education effectiveness. These approaches can enhance the effectiveness and efficiency of nursing education and better prepare nurses to meet patients' complex and evolving healthcare needs through DI in an era of digital disruption. More broadly, this study contributes to the evolving field of digital education by demonstrating how immersive simulation environments with advanced digital systems can support the development of 21st-century skills. By bridging simulation pedagogy and disruptive technology, the study proposes a scalable framework for cultivating nursing competencies, DI, and adaptive capacity in an increasingly globalized and technology-driven healthcare context.

1.1 Research Objectives

To examine the contextual conditions and learning needs within nursing education that inform the development of SBL innovation grounded in active and collaborative learning principles, aimed at addressing gaps in nursing students' professional competencies and DI in response to disruptive technological change.

- To develop a needs-based and theory-informed SBL innovation model integrating simulation pedagogy, Internet of Things (IoT)-enabled humanoid simulators, and web-based learning to support authentic clinical practice, real-time feedback, and collaborative knowledge construction.
- To investigate the effects of the developed SBL innovation model on nursing students' professional competencies and DI, with particular attention to the mechanisms through which technology-enhanced simulation supports skill acquisition, self-directed learning, and clinical readiness.
- To disseminate and validate the practical applicability of the SBL innovation model within nursing education contexts requiring feedback-rich, technology-supported, and resource-efficient approaches to professional competency development.

2. Literature Review

2.1 Disruptive Technology

Disruptive technology refers to sociotechnical processes through which digital innovations transform the delivery, access, and structure of education (Huang, Burton-Jones and Xu, 2024). Such technologies challenge established educational systems while simultaneously generating new opportunities and constraints for teaching and learning (Flavin, 2012; Huapaya et al., 2025). Disruptive technology is characterized by interrelated dimensions: 1) problem-solving capability, 2) system-level transformation, and 3) behavioral change in educational practice (Mills, 2018).

A range of disruptive technologies illustrates the transformation of education systems. Cloud computing enables scalable and distributed learning infrastructures that support collaboration and continuity of instruction, including Massive Open Online Courses (MOOCs) and remote learning environments (Al-Samarraie & Saeed, 2018; James & Weber, 2016; Sharma, Singh and Daim, 2023). Wearable Biosensor Technology enables real-time monitoring of cognitive and physiological states, supporting adaptive instructional responses based on learner engagement and stress levels (Hernández-Mustieles et al., 2024). The IoT connects physical learning environments with real-time data systems, contributing to the development of smart educational ecosystems (Paliktzoglou & Vlachopoulou, 2025). Digital twins and immersive technologies, including virtual reality (VR), augmented reality, and mixed reality, enable medical students to perform risk-free virtual surgeries, transform static textbooks by overlaying 3D holograms, and manipulate virtual objects as though they were tangible (Pandey et al., 2024). MOOCs expand access to education by reducing geographic and financial barriers (Sultan, 2019). Collectively, these disruptive technologies are reshaping education from a content-delivery model into a data-driven and interconnected learning infrastructure.

2.2 Learning Innovation for Nursing Education

Learning innovation involves the integration of emerging technologies, pedagogical strategies, and instructional designs to enhance learning effectiveness and engagement (Leslie, 2015; Redding, Twyman and Murphy, 2013;

Burke et al., 2025). In nursing education, it refers to the systematic use of digital and simulation-based tools to bridge theoretical knowledge and clinical practice. Technological integration in nursing education has led to multiple pedagogical innovations. Mobile Academic Electronic Medical Record (AEMR) systems enable students to practice clinical documentation and strengthen informatics competencies in simulated environments (Choi, Lee and Park, 2018). Cloud-based AEMR systems further enhance clinical reasoning by supporting complex patient data synthesis during simulation (Hong et al., 2022). Immersive tools further bridge the gap between theory and practice, including virtual reality mobile game-based applications (VR-MGBA), which improve long-term retention in disaster scenario training (Hu et al., 2022), while robotic patient simulators enhance procedural skill development (Huang et al., 2017). Smart glasses support real-time procedural guidance and reflective learning through video-assisted debriefing (Kim, Seo and Shin, 2024; Kopetz, Wessel and Jochems, 2020). Mobile applications such as E-MunDiabetes support evidence-based decision-making (Da Silva Negreiros et al., 2022), while clinical education platforms, such as the Clinical Education Management Information System, streamline the evaluation process between students and preceptors (Oxyandi et al., 2024). Exposure to sophisticated robotic surgery simulators further increases students' interest in modern surgical disciplines (Sighinolfi et al., 2023). Collectively, these technologies extend learning beyond traditional instruction and strengthen clinical preparedness.

2.3 Simulation-based Learning (SBL)

SBL is a pedagogical approach that provides a structured and safe environment for nursing students to develop clinical reasoning, decision-making, and procedural competence in nursing education (Holmes, 2024; Skedsmo et al., 2023; Thompson et al., 2024). It shifts learning from passive knowledge acquisition to active experiential engagement within controlled clinical scenarios. The four instructional models outline the essential steps for designing and implementing SBL as follows: 1) preparation, 2) presentation of objectives and guidelines, 3) assignment of learner roles and responsibilities, 4) execution of the simulation task, and 5) debriefing and discussion. This pedagogical approach transforms nursing education from a passive, theoretical exercise into a dynamic, psychologically safe, and structurally standardized learning framework. Instructional models for SBL, such as Tag Team Patient Safety Simulation (Guinea et al., 2019), have been introduced to empower students to challenge unsafe practices and advocate for patient safety within hierarchical clinical environments. High-fidelity SBL (HF-SBL; Lee et al., 2019) promotes learning through practice, enabling students to enhance self-directed learning, critical thinking, and confidence. Immersive Repeated Standard Simulation (iRS-Sim) and Team Deliberate Practice (iSim-TDP; Platt et al., 2024) evaluate repeated scenario engagement and progressively develop clinical self-efficacy without risk to patients. The Electronic Workflow for High-Fidelity Clinical Simulation (EWH-FCSN; Saito & Dal Sasso, 2023) provides a structured digital workflow tool that supports educators in designing and implementing simulations, reducing preparation time, standardizing simulation processes, and facilitating evidence-based practice.

However, several challenges limit the effectiveness of SBL. Early-stage nursing students may experience anxiety, cognitive overload, and emotional stress during high-fidelity simulations, which may negatively affect learning outcomes (Platt et al., 2024). In addition, large class sizes often result in student-based simulations that reduce individual participation and lead to passive observation (Guinea et al., 2019). Furthermore, SBL implementation places significant resource and instructional demands on educators, particularly those with limited experience in simulation pedagogy (Guinea et al., 2019; Saito & Dal Sasso, 2023).

2.4 Digital Intelligence (DI)

DI is a multidimensional competency framework encompassing technical, cognitive, metacognitive, and socioemotional skills grounded in ethical principles for effective participation in digital environments (DQ Institute, 2024). It includes competencies such as digital identity, use, safety, security, emotional intelligence, communication, literacy, and rights (DQ Institute, 2024; Terry et al., 2019). Recent research identifies DI as a key determinant of academic and clinical success in nursing education. Digital literacy mediates the relationship between AI use and creative thinking, particularly when supported by strong academic environments (Agaoglu et al., 2025). Immersive technologies such as metaverse environments enhance digital literacy and student motivation (Amin et al., 2025). However, significant gaps persist between technological advancement and institutional readiness. Nurse educators often lack sufficient digital deficiencies (Li et al., 2025), while nursing curricula remain misaligned with contemporary digital clinical requirements (Stunden et al., 2026). These findings highlight the urgent need for structured digital literacy development within nursing education systems.

2.5 Nursing Competencies

Thailand Nursing and Midwifery Council (2010) defines the core competency framework for registered nurses and midwives across eight key operational domains: 1) ethics, professional conduct, and legal practice; 2) core nursing and midwifery practice; 3) professional attributes; 4) leadership, management, and quality improvement; 5) academic and research competence; 6) communication and relationship management; 7) information technology competence; and 8) social competence. These competencies emphasize patient-centered care, communication, collaboration, and professional accountability, and collectively support the development of nurses capable of advancing healthcare quality and nursing practice.

3. Research Methods

This study adopted a research and development design aimed at developing an SBL innovation within the context of disruptive technology to enhance nursing students' professional competencies and DI. The research was conducted in four phases. Ethical approval was obtained from the Research Ethics Review Committee for Research Involving Human Research Participants, Chulalongkorn University. All participants provided informed consent, and confidentiality was strictly maintained. Data were reported in aggregate form only, ensuring that no participant could be identified.

Ethical approval was granted by Chulalongkorn University Research Ethics Committee with two Certificate of Research Approval numbers (COA No.) granted for each phase: COA No. 038/2564 for Phases 1–2 and COA No. 038/2564 for Phases 3–4. Participant confidentiality was ensured throughout all phases, and all findings were reported in aggregated form for educational and research purposes only.

3.1 Phase 1: Needs Analysis for SBL Innovation

Phase 1 involved two sample groups. The first group consisted of 12 key informants for focus group interviews, including instructors, administrators from the Faculty of Nursing, and educational technology specialists. Participants were selected through purposive sampling based on their relevant expertise and direct experience in nursing education, SBL, and educational technology. The second group comprised survey respondents. Based on Cochran's formula, a minimum sample size of 384 participants was required to ensure statistical adequacy; however, 462 complete responses were obtained and included in the analysis. These nursing students were recruited from the four regions of Thailand (north, northeast, central, and south) to enhance geographical representation.

Both sample groups were drawn from 89 nursing institutes certified by the Nursing Council of Thailand.

Data collection included focus group interviews and structured questionnaires. Instruments were developed based on a literature review and theoretical frameworks, then validated using the Index of Item Objective Congruence by three experts. All items met acceptable validity thresholds (Polit & Beck, 2004) and were refined before data collection. Data were analyzed using content analysis, descriptive statistics, and the Priority Needs Index Modified (PNI_{Modified}).

3.2 Phase 2: Development of SBL Innovation

Phase 2 involved the development of the SBL model and associated technological components based on findings from Phase 1 and relevant literature. The process comprised three stages.

First, an initial SBL model was developed and evaluated by seven experts in nursing education and educational technology, focusing on communication, content validity, usability, and system applicability. Five additional experts assessed content validity in terms of model appropriateness and workflow coherence.

Second, a web-based application integrated with simulation robotics and sensor-controlled systems was developed in alignment with the SBL model. The system was optimized for multi-device compatibility, including desktops, laptops, tablets, and smartphones. Three experts evaluated usability, functionality, and design suitability.

Finally, a focus group with nine experts, including nursing instructors and nursing institute administrators, was conducted to refine the SBL model and system. Feedback from this session informed iterative improvements to both pedagogical design and technological integration.

3.3 Phase 3: Evaluation of SBL Innovation

Phase 3 employed purposive sampling of voluntary participants, comprising 6 nursing instructors and 31 nursing students. Data were collected using four instruments: 1) SBL model, 2) web-based application, 3) sensor-controlled simulation robots, and 4) instructional manual, which was developed during Phase 2. A structured workshop was conducted using the developed SBL system. Students' competencies were assessed across three time points following simulation sessions: Session 1 (Respiratory System Conditions), Session 3 (Subcutaneous Injection), and Session 5 (Intravenous Injection). Descriptive statistics and repeated-measures ANOVA were used to examine changes in students' competencies and digital intelligence across three assessment points. Effect sizes and F value were also calculated to evaluate the magnitude of changes and consistent of improvement over time. Competency assessment instruments included 1) nursing professional competency rubrics, 2) DI rubrics, and 3) a questionnaire evaluating SBL usage. The nursing students' professional competency rubric assessed eight dimensions: ethics, professional conduct, and legal practice; core nursing and midwifery practice; professional attributes; leadership, management, and quality improvement; academic and research competence; communication and relationship management; information technology competence; and social competency. The DI rubric assessed eight dimensions: digital identity, use, safety, security, emotional intelligence, communication, literacy, and rights. Both rubrics used a three-level self-assessment scale: 1) Needs Improvement, 2) Fair, and 3) Good. All instruments were validated by three experts in nursing education, educational technology and communications, and educational measurement and evaluation. The researchers incorporated expert feedback and revised both rubric assessments to ensure their suitability for actual use. For the third instrument, a questionnaire assessing the use of SBL innovation in the era of disruptive technology to enhance nursing students' professional competencies and DI was designed using a five-point Likert scale. The instrument was reviewed by three experts who evaluated its content validity, as well as the completeness and comprehensiveness of its items. Following revision, the questionnaire was piloted with nursing students to assess reliability and ensure instrument quality before full implementation. Reliability analysis confirmed strong internal consistency, with a Cronbach's alpha coefficient of 0.981.

3.4 Phase 4: Diffusion and Validation

Phase 4 focused on refining and disseminating the SBL innovation based on findings from Phase 3. Iterative improvements were made to the learning model and technological system before expert validation. Seven experts in nursing, educational evaluation, and educational technology participated in the validation process. Qualitative feedback from experts was analyzed using content analysis to inform further system refinement. Quantitative evaluation of innovation acceptance was conducted using descriptive statistics based on a five-point Likert scale.

4. Findings

4.1 Research Results of Phase 1: Conditions and Needs of SBL Innovation

Most respondents reported prior exposure to SBL ($n = 400$, 80.6%) and expressed strong interest in SBL participation ($n = 462$, 93.1%). Participants indicated that technology-enhanced SBL would support flexible learning, real-time feedback, independent information retrieval, post-learning review, pre-class preparation, and improved accessibility during disruptions such as the COVID-19 pandemic.

In terms of learning needs, students prioritized simulation-based clinical scenarios as their preferred instructional approach, followed by blended online learning activities and structured practical assessments. Additionally, students expressed interest in collaborative learning through peer discussion and knowledge exchange.

Instructors reported that the proposed innovation would enhance nursing students' professional competencies and DI while also providing institutions with accessible and cost-effective tools for monitoring student performance in virtual learning environments in real-time.

4.2 Research Results of Phase 2: Development of SBL Innovation

The results of the literature synthesis on SBL, together with the analysis of Phase 1 needs assessment, informed the development of the learning model components and instructional processes. The model comprises five components: objectives, scenario content, participant roles, learning activities using the Nurse Sims system, and assessment. The instructional sequence follows the 5Ps framework: 1) preparation of equipment and the Nurse Sims applications, 2) presentation of simulations using Nurse Sims, 3) practice of simulations using Nurse Sims, 4) project discussion and debriefing, and 5) performing the assessment in Figure 1.



Figure 1: The components and 5Ps steps of the learning model

The SBL innovation comprises two integrated systems: the Nurse Sims web application and the Nurse Sims simulation equipment, which function together as shown in **Figure 2** and **Figure 3**. The Nurse Sims web application is accessible via <https://nursing-sim.edii.in.th> and is designed as a progressive web application with a responsive interface compatible across devices. The system architecture is presented in **Figure 5**, while the entity-relationship structure is shown in **Figure 6**.

Instructional manuals for both students and instructors are provided in **Figure 4**.

The hardware system includes 1) a vascular simulation arm model, 2) an IoT-enabled sensor system comprising an ESP32 microcontroller and total dissolved solids sensor, 3) respiratory and cardiac sound simulation modules, and 4) a web-based monitoring interface for data visualization and recording (**Figure 7**). The integrated system supports five simulations: 1) respiratory assessment, 2) cardiovascular assessment, 3) intramuscular injection, 4) intravenous injection, and 5) intravenous fluid administration (**Figure 8**).

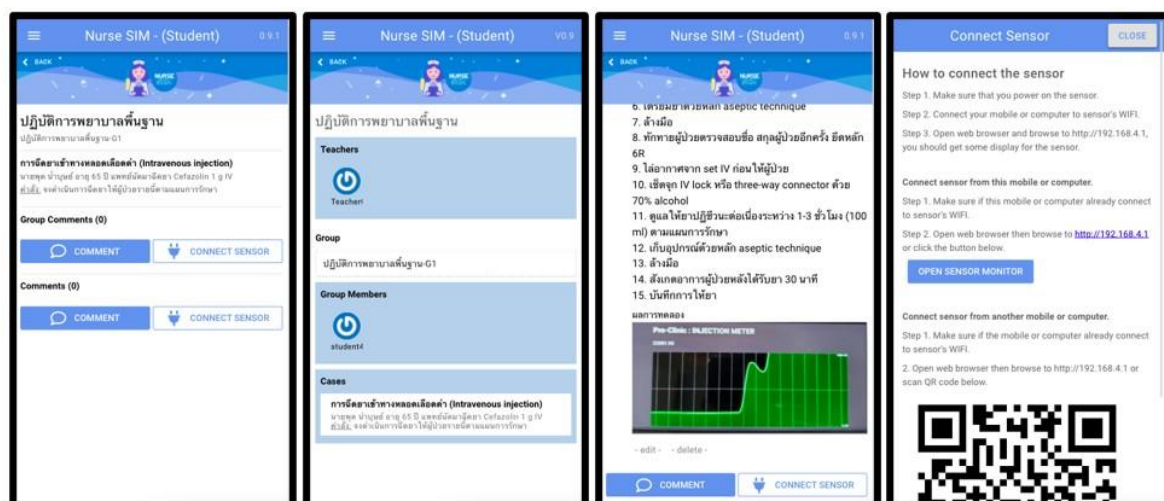


Figure 2: Nurse Sims application (students' view)

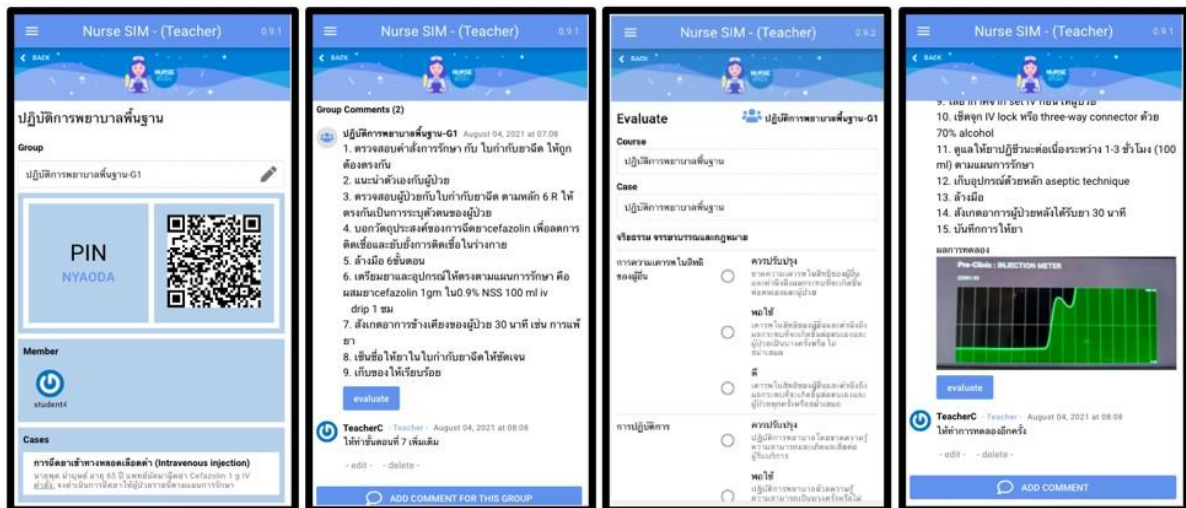


Figure 3: Nurse Sims application (instructors' view)



Figure 4: Manual for the Nurse Sims application for students (left) and instructors (right)

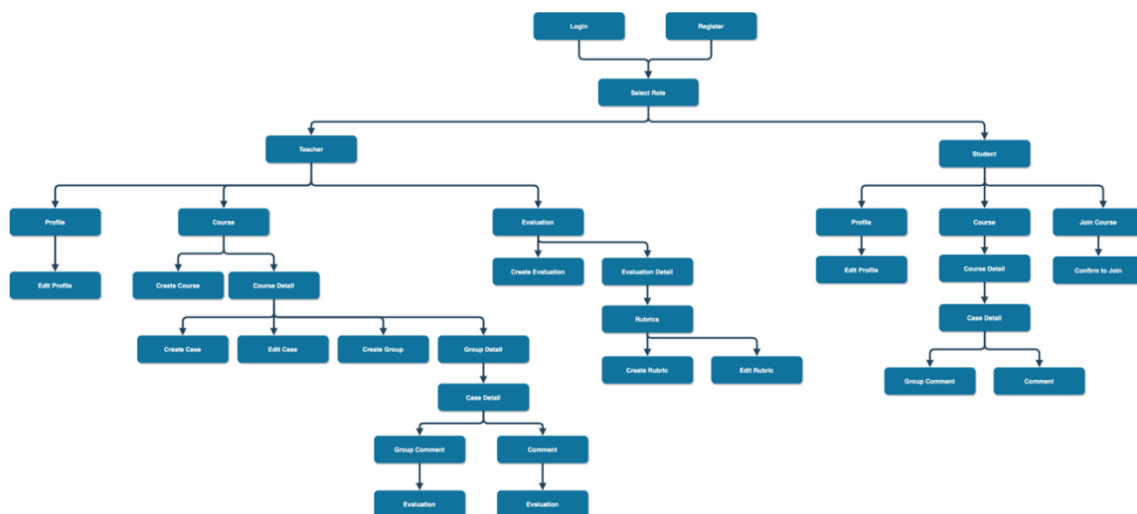


Figure 5: The structure of the Nurse Sims application

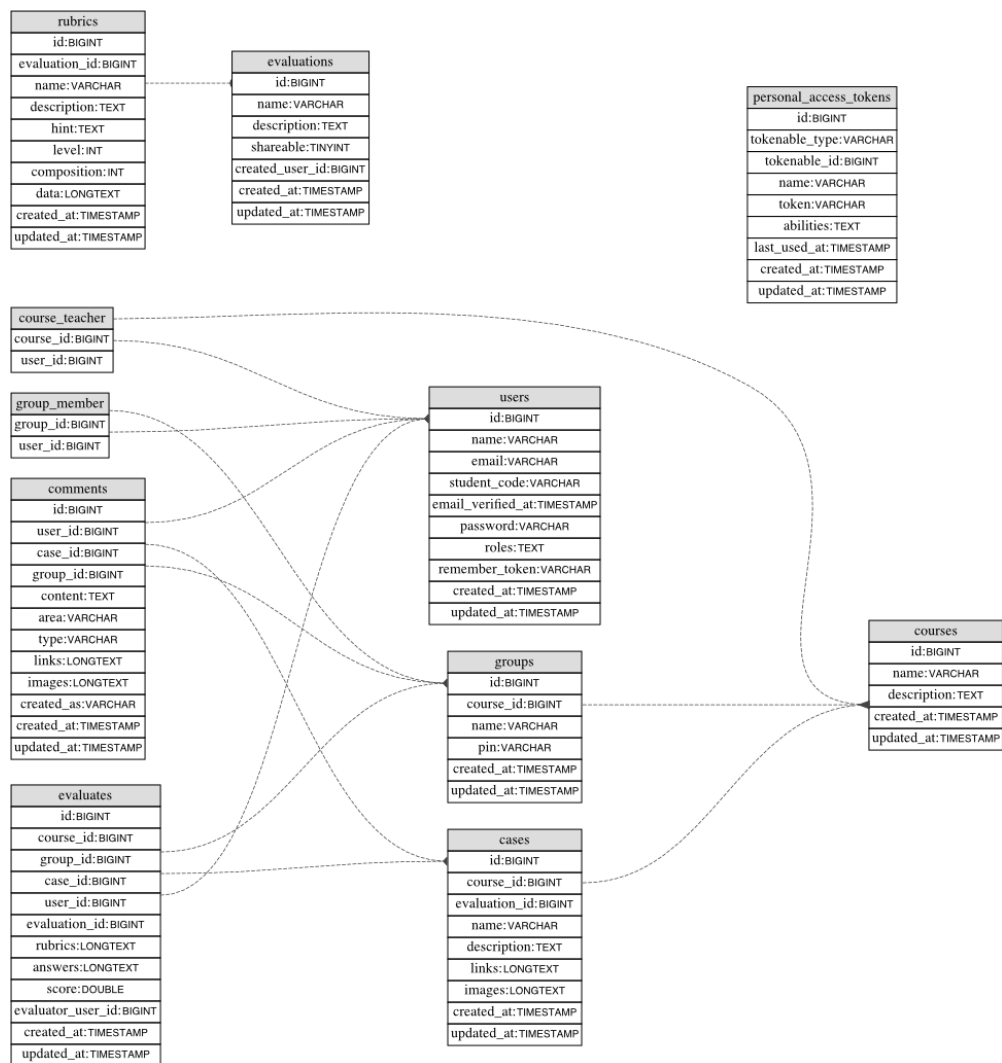


Figure 6: The entity-relationship of the Nurse Sims application

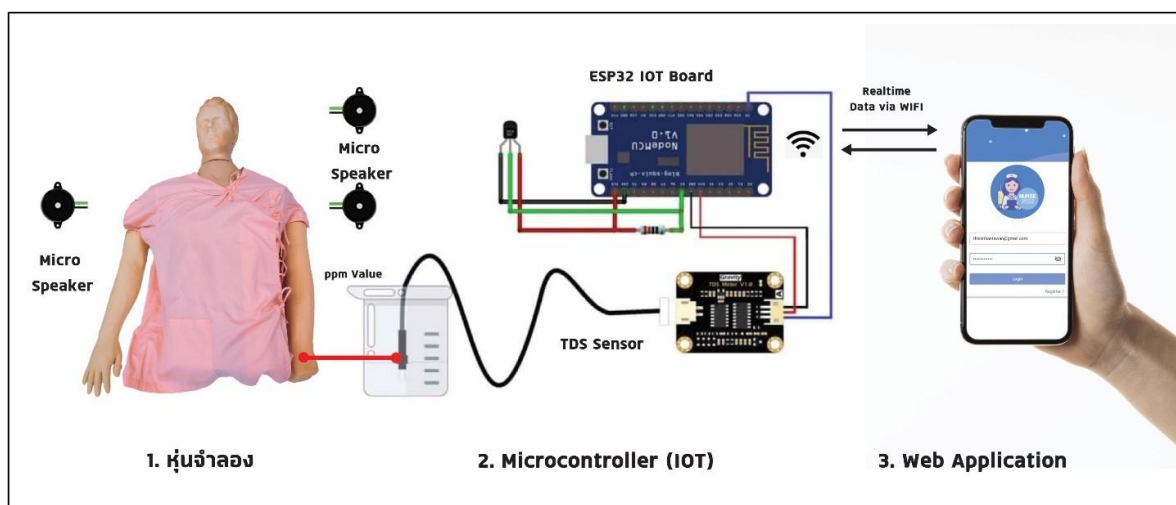


Figure 7: Technical components of the Nurse Sims application



Figure 8: Five simulations of the Nurse Sims application

4.3 Research Results of Phase 3: Effects of SBL Innovation

A repeated-measures ANOVA was conducted to examine changes in nursing students' professional competency scores across three assessment points. Data were collected using Likert scale assessments administered after Situations 1, 3, and 5 (descriptive analysis show in **Table 1**). These situations represented sequential simulation assessment points completed by the same participants during the intervention and corresponded to the early, intermediate, and final stages of learning progression. The results from Table 2 Situations 1, 3, and 5 indicated a statistically significant difference in the overall mean scores of nursing students at the .05 significance level ($F = 1,293.430, p < .001$). The highest mean score was recorded in Situation 5 ($M = 2.84, SD = 0.14$), followed by Situation 3 ($M = 2.44, SD = 0.19$) and Situation 1 ($M = 1.42, SD = 0.24$). Overall, nursing students demonstrated the highest competency level at the final assessment point (Situation 5; post-intervention), followed by the intermediate (Situation 3) and baseline (Situation 1; pre-intervention) assessments. The effect size analysis using partial eta squared (η^2) from the Tests of Within-Subjects Effects indicated a very large effect of Nurse Sims on nursing students' professional competencies. The η^2 value of 0.911 suggests that a substantial proportion of variance in competency scores across assessment points was attributable to the intervention, indicating both statistical and practical significance. These results indicate that nursing competencies improved following exposure to Nurse Sims.

Table 1: Descriptive analysis of nursing competency scores across three situations

Nursing competencies	Situation 1		Situation 3		Situation 5	
	Mean	SD	Mean	SD	Mean	SD
Ethics, code of conduct, and the law	1.71	0.36	2.82	0.24	1.71	0.36
Core nursing and midwifery practices	1.13	0.34	2.32	0.48	1.13	0.34
Professional characteristics	1.31	0.31	2.36	0.25	1.31	0.31
Leadership, management, and quality improvement	1.16	0.37	2.16	0.37	1.16	0.37
Academics and research competencies	1.10	0.20	2.19	0.25	1.10	0.20
Communication and relationship	1.57	0.38	2.32	0.24	1.57	0.38
Information technology	1.71	0.53	2.71	0.46	1.71	0.53
Social competency	1.61	0.50	2.68	0.48	1.61	0.50
Overall	1.42	0.24	2.44	Overall	1.42	0.24

Table 2: Repeated-measures ANOVA of nursing competency scores across three situations

Nursing competencies	Variance	SS	df	MS	F	sig	Results
Ethics, code of conduct, and the law	Estimate	23.536	1	23.536	419.322	.001	5>3>1
	Error	1.684	30	0.056			
Core nursing and midwifery practices	Estimate	35.629	1	35.629	276.125	.001	5>3>1
	Error	3.871	30	0.129			
Professional characteristics	Estimate	33.981	1	33.981	561.916	.001	5>3>1
	Error	1.814	30	0.06			
Leadership, management, and quality improvement	Estimate	40.323	1	40.323	328.947	.001	5>3>1
	Error	3.677	30	0.123			
Academics and research competencies	Estimate	42.778	1	42.778	803.712	.001	5>3>1
	Error	1.597	30	0.053			
Communication and relationship	Estimate	23.907	1	23.907	241.671	.001	5>3>1
	Error	2.968	30	0.099			
Information technology	Estimate	25.806	1	25.806	184.615	.001	5>3>1
	Error	4.194	30	0.140			
Social competency	Estimate	29.823	1	29.823	243.289	.001	5>3>1
	Error	3.677	30	0.123			
Overall	Estimate	31.084	1	31.084	1293.430	.001	5>3>1
	Error	0.721	30	0.024			

The repeated-measures ANOVA examined differences in nursing students' DI scores across three assessment points (Situations 1, 3, and 5) (descriptive analysis show in **Table 3**). The results indicated a statistically significant difference in overall DI scores at the 0.05 level ($F = 608.285, p < .001$). Mean DI scores increased from Situation 1 ($M = 1.43, SD = 0.30$) to Situation 3 ($M = 2.51, SD = 0.35$), and reached the highest level in Situation 5 ($M = 2.93, SD = 0.10$). The implementation of Nurse Sims was associated with a marked improvement in DI outcomes. The effect size ($\eta^2 = .933$) indicated a very large effect, suggesting that the intervention accounted for a substantial proportion of variance in DI scores across time points. These findings demonstrate both statistical and practical significance in the development of DI through SBL. Consistent with the competency findings, the significant F-values (**Table 4**) confirm that DI improved significantly following the intervention.

Table 3: Descriptive analysis of DI scores across three situations

Digital intelligence	Situation 1		Situation 3		Situation 5	
	Mean	SD	Mean	SD	Mean	SD
Digital identity	1.71	0.46	2.61	0.50	1.71	0.46
Digital use	1.39	0.50	2.29	0.46	1.39	0.50
Digital safety	1.58	0.50	2.58	0.50	1.58	0.50
Digital security	1.55	0.51	2.55	0.51	1.55	0.51
Digital emotional intelligence	1.29	0.46	2.45	0.51	1.29	0.46
Digital communication	1.23	0.43	2.45	0.51	1.23	0.43
Digital literacy	1.23	0.43	2.48	0.51	1.23	0.43
Digital rights	1.48	0.51	2.48	0.51	1.48	0.51
Overall	1.43	0.30	2.51	Overall	1.43	0.30

Table 4: Repeated-measures ANOVA of DI scores across three situations

Digital intelligence	Variance	SS	df	MS	F	sig	Results
Digital identity	Estimate	25.806	1	25.806	242.424	.001	5>3>1
	Error	3.194	30	0.106			
Digital use	Estimate	28.452	1	28.452	240.545	.001	5>3>1
	Error	3.548	30	0.118			
Digital safety	Estimate	29.823	1	29.823	243.289	.001	5>3>1
	Error	3.677	30	0.123			
Digital security	Estimate	28.452	1	28.452	130.345	.001	5>3>1
	Error	6.548	30	0.218			
Digital emotional intelligence	Estimate	38.726	1	38.726	307.821	.001	5>3>1
	Error	3.774	30	0.126			
Digital communication	Estimate	45.306	1	45.306	425.606	.001	5>3>1
	Error	3.194	30	0.106			
Digital literacy	Estimate	45.306	1	45.306	324.115	.001	5>3>1
	Error	4.194	30	0.140			
Digital rights	Estimate	32.661	1	32.661	255.252	.001	5>3>1
	Error	3.839	30	0.128			
Overall	Estimate	34.875	1	34.875	608.285	.001	5>3>1
	Error	1.720	30	0.057			

Overall, nursing students reported a very high level of acceptance toward SBL innovation in a changing technological context for enhancing professional competencies and DI ($M = 4.52$, $SD = 0.49$). Instructors demonstrated a very positive attitude toward the use of SBL innovation for enhancing nursing students' professional competencies and DI. The overall level of agreement was high ($M = 4.69$, $SD = 0.49$).

4.4 Research Results of Phase 4: The Diffusion of SBL Innovation in the era of Disruptive Technology to Enhance Nursing Students' Professional Competencies and DI

In Phase 4, experts evaluated the SBL innovation and rated it as highly appropriate ($M = 4.76$, $SD = 0.20$). Experts further confirmed that the model had well-structured components, clearly defined learning steps, and a transparent implementation process. Additional feedback highlighted the practicality of the system in terms of usability, guidelines, and assessment, supporting its feasibility for real-world implementation.

5. Discussion

The effectiveness of Nurse Sims as an SBL innovation can be explained by its needs-driven and theory-informed design. The system was developed to address nursing students' demand for flexible simulation environments that provide real-time feedback, opportunities for post-learning reflection, and pre-class preparation. Nurse Sims integrates IoT-enabled sensor technology and a web-based platform, enabling students to engage in realistic, technology-enhanced clinical practice in a controlled environment. The intervention was delivered through a 5P learning process grounded in active and collaborative learning principles. Through structured activities including brainstorming, discussion, peer interaction, and hands-on practice, students co-constructed knowledge while developing both professional competencies and DI. The significant improvements observed in outcomes, together with expert validation of feasibility, indicate that Nurse Sims provides a practical model for integrating SBL, humanoid technology, and collaborative learning in nursing education.

The findings are consistent with previous studies demonstrating the effectiveness of SBL in nursing education. Kuesakul et al. (2024) and Platt et al. (2024) reported that SBL improves students' satisfaction, attitudes toward simulation design, knowledge acquisition, and self-efficacy by enabling safe practice environments and immediate feedback. In addition, Saito and Dal Sasso (2023) highlighted that technology-supported simulation systems enhance instructional efficiency by standardizing simulation design and reducing educator workload. Collectively, these findings suggest that technology-enhanced SBL improves both learning and teaching management efficiency.

Evidence also supports the effectiveness of manikin-based and high-fidelity simulation in nursing education. Li et al. (2024) found improvements in cooperative ability, learning motivation, and self-management among nursing students. Simulation-based environments also enhance critical thinking, problem-solving, and situational awareness by exposing students to complex clinical scenarios requiring team coordination (Hill et al., 2023). Furthermore, Brohard et al. (2021) demonstrated that simulation allows safe practice of emotionally challenging clinical communication without risk to patients. The integration of mobile and web-based learning tools in this study aligns with previous findings. Raurell-Torredà et al. (2025) reported that mobile simulation platforms significantly increased student satisfaction and self-confidence compared to traditional methods. Similarly, Khoshbakht-Pishkhani et al. (2024) found improvements in cognitive skills following mobile-assisted learning interventions. Carvalho and Sousa (2024), in interviews with 10 nursing students in Portugal, further demonstrated that mobile applications can enhance student engagement and understanding of patient self-practices in chronic disease management.

Additionally, the results align with studies on digital competence in nursing education. Abou Hashish and Alnajjar (2024) reported positive student attitudes toward digital transformation and confidence in using AI-based tools. Jallad et al. (2024) found that contemporary nursing students frequently use digital tools, including mobile applications, simulation platforms, and generative AI, to support learning. Reed and Dodson (2024) further emphasized that digital storytelling and AI-based visualization tools improve engagement and comprehension among Generation Z learners.

6. Limitations

Despite the promising findings of this study, several limitations should be acknowledged. First, the absence of a control or comparison group limits the ability to draw causal inferences regarding the relative effectiveness of the Nurse Sims intervention compared with traditional teaching methods. Second, the relatively small sample size in the implementation phase ($n = 31$) may restrict the generalizability of the findings to broader nursing student populations. Future research should employ larger samples and randomized controlled designs to strengthen internal validation and further examine the long-term impact of simulation-integrated technologies in nursing education.

7. Recommendations

7.1 Recommendations for the Application of Findings

To enhance the effectiveness of using simulation robots, along with controlling equipment and sensors, it is recommended that these tools be used in conjunction with the Nurse Sims application. The platform supports efficient scenario design, activity organization, and structured facilitation by instructors, while enabling students to engage meaningfully with simulated clinical situations.

When using the robotic simulator with integrated sensor-controlled systems alongside the Nurse Sims application, adherence to the 5Ps instructional framework is recommended. This structured process supports the systematic development of nursing students' professional competencies and DI.

7.2 Recommendations for Future Research

Future research should explore the development of simulation-based innovations using pediatric patient simulations. Such expansion would better prepare nursing students for specialized clinical contexts and improve readiness for pediatric care delivery.

Future research should enhance the realism and complexity of simulation interactions by incorporating dynamic patient behaviors, including noncooperative patients and spontaneous learner–patient dialogue. This would broaden scenario variability and strengthen clinical decision-making skills.

Future research is encouraged to adopt more robust methodological designs, including longitudinal studies and cross-disciplinary replications, to validate findings and strengthen the generalizability of the Nurse Sims innovation across diverse educational and clinical contexts.

AI Statement: The authors declare that they have used AI tools to enhance the spelling and readability of this manuscript.

Ethics Statement: This study involved human participants. Ethical approval was obtained from the Chulalongkorn University Research Ethics Committee (Multi-Institutional Group 2: Social Sciences, Humanities, and Fine Arts; reference numbers 038/2021 and 080/2022). Written informed consent was obtained from all

participants for participation and publication, with assurances of anonymity. Throughout the study, strict adherence to privacy, data protection, and confidentiality protocols was maintained. All participant data were anonymized and securely managed, and no identifying information was disclosed in this publication.

References

- Abou Hashish, E. A. and Alnajjar, H., 2024. Digital proficiency: assessing knowledge, attitudes, and skills in digital transformation, health literacy, and artificial intelligence among university nursing students. *BMC Medical Education*, 24.
- Agaoglu, F. O., Bas, M., Tarsuslu, S., Ekinci, L. O. and Agaoglu, N. B., 2025. The mediating digital literacy and the moderating role of academic support in the relationship between artificial intelligence usage and creative thinking in nursing students. *BMC Nursing*, 24.
- Al-Samarraie, H. and Saeed, N., 2018. A systematic review of cloud computing tools for collaborative learning: Opportunities and challenges to the blended-learning environment. *Computers & Education*, 124, pp 77-91.
- Amin, S. M., El-Fattah Mahgoub, S. A., Tawfik, A. F., Khalil, D. E., El-Sayed, A. A. I., Atta, M. H. R., Albzia, A. and Morsy Mohamed, S. R., 2025. Nursing education in the digital era: the role of digital competence in enhancing academic motivation and lifelong learning among nursing students. *BMC Nursing*, 24.
- Bayram, S. B. and Caliskan, N., 2019. Effect of a game-based virtual reality phone application on tracheostomy care education for nursing students: A randomized controlled trial. *Nurse Education Today*, 79, pp 25-31.
- Brohard, C., Moreland, E., Shammas, S. and Tonsul, R., 2021. From Development to Implementation: The Effectiveness of a High-Fidelity Simulation for Advance Care Planning for Undergraduate Nursing Students. *Journal of Hospice and Palliative Nursing*, 23, pp 504-511.
- Burke, P. F., Schuck, S., Burden, K. and Kearney, M., 2025. Mediating learning with mobile devices through pedagogical innovation: Teachers' perceptions of K-12 students' learning experiences. *Computers & Education*, 227, 105226.
- Carvalho, S. M. and Sousa, M. R. M. G. C., 2024. Nurses' Perspectives on the Use of Mobile Applications for Self-Care in Chronic Illnesses. *Texto e Contexto Enfermagem*, 33.
- Chen, Y., Fan, N. and Wu, H., 2025. AI-Enabled Metaverse for Education: Challenges and Opportunities. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pp 103-113.
- Choi, M., Lee, H. and Park, J. H., 2018. Effects of using mobile device-based academic electronic medical records for clinical practicum by undergraduate nursing students: A quasi-experimental study. *Nurse Educ Today*, 61, pp 112-119.
- Cochran, W. G., 1977. *Sampling Techniques*, New York: John Wiley & Sons.
- Da Silva Negreiros, F. D., Flor, A. C., De Araújo, A. L., Cestari, V. R. F., Florêncio, R. S., Moreira, T. R., Araújo, S. T., Da Silva, L. M. S. and Moreira, T. M. M., 2022. E-MunDiabetes: A Mobile Application for Nursing Students on Diabetes Education During the COVID-19 Pandemic. *CIN - Computers Informatics Nursing*, 40, pp 325-334.
- DQ Institute. 2024. *DQ Institute Launches AI-Powered Certification System for Digital Literacy Programs* [Online]. DQ Institute. Available: <https://www.dqinstitute.org/news-post/dq-institute-launches-ai-powered-certification-system-for-digital-literacy-programs/>.
- European Commission, 2013. Staff working document on incro-credentials for lifelong learning and employability. European Commission.
- Flavin, M., 2012. Disruptive technologies in higher education. *Research in Learning Technology*, 20.
- Guinea, S., Andersen, P., Reid-Searl, K., Levett-Jones, T., Dwyer, T., Heaton, L., Flenady, T., Applegarth, J. and Bickell, P., 2019. Simulation-based learning for patient safety: The development of the Tag Team Patient Safety Simulation methodology for nursing education. *Collegian*, 26, pp 392-398.
- Gupta, P. and Verma, M., 2024. Metaverse and industry 4.0: An effective integration for disruption in varied sectors and business services. *New Business Frontiers in the Metaverse*.
- Hamilton, T. 2025. Preserving Creativity in Nurse Education. *Nursing Science Quarterly*, 38, pp 21-24.
- Hernández-Mustieles, M. A., Lima-Carmona, Y. E., Pacheco-Ramírez, M. A., Mendoza-Armenta, A. A., Romero-Gómez, J. E., Cruz-Gómez, C. F., Rodríguez-Alvarado, D. C., Arceo, A., Cruz-Garza, J. G., Ramírez-Moreno, M. A. and Lozoya-Santos, J. D. J., 2024. Wearable Biosensor Technology in Education: A Systematic Review. *Sensors*, 24.
- Hill, K., Schumann, M., Farren, L. and Clerkin, R., 2023. An Evaluation of the Use of Low-Fidelity and High-Fidelity Mannequins in Clinical Simulations in a Module Preparing Final Year Children's and General Nursing Students for Internship Placement. *Comprehensive Child and Adolescent Nursing*, 46, pp 295-308.
- Holmes, K. 2024. The Future of Simulation-Based Learning in Acute Pediatric Nursing Education: A Conceptual Review. *Clinical Simulation in Nursing*, 87, 101491.
- Hong, S., Cho, I., Park, M., Lee, J. Y., Lee, J. and Choi, M. 2022. Simulation Education Incorporating Academic Electronic Medical Records for Undergraduate Nursing Students: A Pilot Study. *Healthcare Informatics Research*, 28, pp 376-386.
- Hou, M., Lu, Y., Liu, Y., Luo, W., Xue, H., Yuan, H. and Zhang, X. 2025. The impact of virtual reality applied in nurse education on students' clinical thinking: A systematic review. *Nurse Education Today*, 146, 106553.
- Hu, H., Lai, X., Li, H. and Nyland, J. 2022. Teaching disaster evacuation management education to nursing students using virtual reality mobile game-based learning. *CIN - Computers Informatics Nursing*, 40, pp 705-710.

- Huang, S., Burton-Jones, A. and Xu, D. A., 2024. Configurational theory of digital disruption. *Information Systems Journal*, pp 1737-1786.
- Huang, Z., Lin, C., Kanai-Pak, M., Maeda, J., Kitajima, Y., Nakamura, M., Kuwahara, N., Ogata, T. and Ota, J., 2017. Impact of Using a Robot Patient for Nursing Skill Training in Patient Transfer. *IEEE Transactions on Learning Technologies*, 10, pp 355-366.
- Huapaya, E. S. R., Chucos, G. L., Sosa, E. P. and Meza, M. I., 2025. Disruptive technologies in the university curriculum: use of artificial intelligence. *International Journal of Evaluation and Research in Education*, 14, pp 671-681.
- Jallad, S. T., Alsager, K., Albadareen, B. I. and Al-Maghaireh, D., 2024. Artificial intelligence tools utilized in nursing education: Incidence and associated factors. *Nurse Education Today*, 142, 106355.
- James, C. N. and Weber, J., 2016. Chapter 7 - Cloud Computing in Education. In: VANCE, T. C., MERATI, N., YANG, C. & YUAN, M. (eds.) *Cloud Computing in Ocean and Atmospheric Sciences*. Academic Press.
- Kandil, O., Rosillo, R., Abd El Aziz, R. and De La Fuente, D., 2025. Investigating the impact of the Internet of Things on higher education: a systematic literature review. *Journal of Applied Research in Higher Education*, 17, pp 254-273.
- Khoshbakht-Pishkhani, M., Javadi-Pashaki, N., Esfandi, N. A., Koodakani, M. B., Maroufizadeh, S. and Madani, A. H., 2024. The effect of educational application in nursing internship clinical training on cognitive and functional skills and students' satisfaction. *BMC Nursing*, 23.
- Kim, J., Seo, M. and Shin, H., 2024. Learners' Experiences With First-Person Perspective Video Debriefing Using Smart Glasses in Nursing Simulation Education. *Clinical Simulation in Nursing*, 94.
- Kopetz, J. P., Wessel, D. and Jochems, N., 2020. User-centered development of smart glasses support for skills training in nursing education. *i-com*, 18, pp 287-299.
- Kuesakul, K., Nuampa, S., Pungbangkadee, R., Ramjan, L. and Ratinthorn, A., 2024. Evaluation of antenatal simulation-based learning on satisfaction and self-confidence levels among Thai undergraduate nursing students during the COVID-19 pandemic: a mixed-method study. *BMC Nursing*, 23.
- Laugaland, K. A., Klippen, M. I. F., Ravik, M., Gonzalez, M. T. and Akerjordet, K., 2025. Exploring co-creation with student nurses in contemporary nursing education: A systematic scoping review. *Nurse Education Today*, 146, 106515.
- Lee, J. J., Yeung, K. C. Y., Clarke, C. L. and Yoo, J., 2019. Nursing Students' Learning Dynamics and Perception of High-Fidelity Simulation-Based Learning. *Clinical Simulation in Nursing*, 33, pp 7-16.
- Leslie, D., 2015. *Innovative Learning Case: HAMK university of applied sciences*. Bachelor's Thesis].
- Li, P., Tan, R., Yang, T. and Meng, L., 2025. Current status and associated factors of digital literacy among academic nurse educators: a cross-sectional study. *BMC Medical Education*, 25.
- Li, Q. and Cheung, D. S. T., 2024. Nurse education in Asia: A microcosm of the changing world. *Nurse Education Today*, 106551.
- Li, Y., Lv, Y., Dorol, R. D., Wu, J., Ma, A., Liu, Q. and Zhang, J., 2024. Integrative virtual nursing simulation in teaching cardiopulmonary resuscitation: A blended learning approach. *Australasian Emergency Care*, 27, pp 37-41.
- Lim, S. H., Ang, S. Y., Lim, Y. Y. A., Leow, W. X. B., Binte Sunari, R. N., Foo, X. A. and Aloweni, F., 2024. Effectiveness of spaced education pedagogy in enhancing Nurses' knowledge on emotional resilience—A quasi-experimental trial. *Journal of Nursing Scholarship*.
- Mills, K., 2018. *Three Characteristics of Digital Disruption* [Online]. Available: <https://www.linkedin.com/pulse/three-characteristics-digital-disruption-kerri-kalfai/>.
- Oladokun, B. D., Ajani, Y. A., Tom-George, N. W. and Okeke, O. C., 2024. From metaverse to meta AI: a dynamic disruption in libraries in higher education institutions. *Library Hi Tech News*, 41, pp 12-14.
- Oxyandi, M., Panduragan, S. L., Said, F. M. and Rivanica, R., 2024. Development of Mobile Learning Application: Clinical Education Management Information System (CEMIS) Among Nursing Students in Indonesia. *Malaysian Journal of Medicine and Health Sciences*, 20, pp 217-224.
- Paliktzoglou, V. and Vlachopoulou, O., 2025. Exploring the Educational Transformations: A Systematic Literature Review on the Influence of the Internet of Things in Higher Education. *EAI Endorsed Transactions on Internet of Things*, 11.
- Pandey, P. K., Pandey, P. K., Mahajan, S., Paul, J. and Iyer, S., 2024. Digital Twin and Virtual Reality, Augmented Reality, and Mixed Reality. *Digital Twins for Smart Cities and Villages*.
- Platt, A., Allan, J., Leader, C., Prescott-Clements, L. and Mcmeekin, P., 2024. Preparing for practice, the effects of repeated immersive simulation on the knowledge and self-efficacy of undergraduate nursing students: A mixed methods study. *Nurse Education in Practice*, 74, 103866.
- Pleshkan, V., 2025. Standardizing nurse practitioner clinical education: The Cognitive Preceptorship Model. *Journal of Professional Nursing*, 57, pp 26-42.
- Polit, D. F. and Beck, C. T., 2004. *Nursing research: Principles and methods*, Philadelphia: Lippincott Williams & Wilkins.
- Raurell-Torredà, M., Obejo-Alamillos, C., Turégano-Duaso, C., Esteban-Sepúlveda, S., Adalid-Villaronga, R., Estrada-Masllorens, J. M., Sarria-Guerrero, J. A., Basco-Prado, L., Zaragoza-García, I. And Fernández-Cervilla, A. B., 2025. SimCapture for Skills®: Mobile-based web skills training for undergraduate nursing students: A randomized clinical trial. *Clinical Simulation in Nursing*, 99.
- Redding, S., Twyman, J. S. and Murphy, M., 2013. *What is an innovation in learning?* [Online]. Available: http://www.centeril.org/handbook/resources/fullchapter/What_is_an_Innovation_SA.pdf.
- Reed, J. M. and Dodson, T. M., 2024. Generative AI Backstories for Simulation Preparation. *Nurse Educator*, 49.
- Roettger, L., Tobias, J., Didonato, S., Wenocur, K., Page, A. and Wycoff, K., 2025. Interprofessional Simulation Learning Using Pediatric Nurse Practitioner and Counseling Students. *The Journal for Nurse Practitioners*, 21, 105269.

- Roh, Y. S., Kim, S. S., Park, S. and Ahn, J. W., 2020. Effects of a simulation with team-based learning on knowledge, team performance, and teamwork for nursing students. *CIN - Computers Informatics Nursing*, 38, pp 367-372.
- Rungrangtanapol, N. and Khlaisang, J., 2021. Development of a Teaching Model in Virtual Learning Environment to Enhance Computational Competencies in the 21st Century. *International Journal of Interactive Mobile Technologies*, 15, pp 93-107.
- Saito, C. M. and Dal Sasso, G. T. M., 2023. Electronic Workflow for High-Fidelity Clinical Simulation in Nursing. *CIN - Computers Informatics Nursing*, 41, pp 338-345.
- Sharma, M., Singh, A. and Daim, T., 2023. Exploring cloud computing adoption: COVID era in academic institutions. *Technological Forecasting and Social Change*, 193.
- Sighinolfi, M. C., Terzoni, S., Scanferla, E., Bianchi, P. P., Formisano, G., Gaia, G., Marconi, A. M., Chiumello, D., Patel, V., Moschovas, M. C., Turri, F., Dell'orto, P., Maruccia, S., Grasso, A., Sangalli, M., Centanni, S., Stocco, M., Assumma, S., Sarchi, L., Calcagnile, T., Panio, E., Bozzini, G. and Rocco, B., 2023. Impact of hands-on practice with HugoRAS and Versius System simulators on the attractiveness of robotic surgery among medical and nurse undergraduate students. *Journal of Robotic Surgery*, 17, pp 1471-1476.
- Skedsmo, K., Nes, A. A. G., Stenseth, H. V., Hofsø, K., Larsen, M. H., Hilderson, D., Smis, D., Hagelin, C. L., Olaussen, C., Solberg, M. T., Bingen, H. M., Ølnes, M. A. and Steindal, S. A., 2023. Simulation-based learning in palliative care in postgraduate nursing education: a scoping review. *BMC Palliative Care*, 22.
- Stunden, A., Ginige, A., O'reilly, R., Sanagavarapu, P. and Heaton, L., 2026. Exploring educational barriers to pre-registration nursing students digital literacy competency in the clinical setting: A descriptive qualitative design. *Nurse Education Today*, 163, 107091.
- Sultan, N., 2019. The Disruption and Global Implications of Massive Open Online Courses (MOOCs) for Higher Education. *Contributions to Economics*.
- Terry, J., Davies, A., Williams, C., Tait, S. and Condon, L., 2019. Improving the digital literacy competence of nursing and midwifery students: A qualitative study of the experiences of NICE student champions. *Nurse Educ Pract*, 34, pp 192-198.
- Thailand Nursing and Midwifery Council, 2010. Competency in nursing and Midwifery Professionals.
- Thompson, A. J., Boswell, A. C., Evans, T. and Clarke, B., 2024. Clinical emergencies in inpatient hospice: simulation-based training to improve nursing confidence. *BMJ Supportive and Palliative Care*, 14, e2483-e2487.
- Ylönen, M., Paajanen, P., Kukkonen, T., Torkki, P., Paloneva, J. and Rosqvist, E., 2025. From Emergency Room to Operating Room: Multidisciplinary Simulation Training in Emergency Laparotomy for Ruptured Abdominal Aortic Aneurysm – Learning Outcomes and Costs. *Annals of Vascular Surgery*, 111, pp 341-350.

Ubuntu-informed Critical Thinking in First-year Online Accounting: Academics' Perceptions

Rika Dry¹, Annelien van Rooyen¹ and Geesje van den Berg²

¹Department of Financial accounting, University of South Africa, South Africa

²Department of Curriculum and Instructional studies, University of South Africa, South Africa

vzylhj@unisa.ac.za

vrooyaa@unisa.ac.za

vdberg@unisa.ac.za

<https://doi.org/10.34190/ejel.24.4.4478>

An open access article under [CC Attribution 4.0](#)

Abstract: This study investigates how the African philosophy of *ubuntu* can be harnessed to foster critical thinking - particularly good thinking habits - among first-year financial accounting students in an online learning environment. As critical thinking becomes an essential skill for accounting graduates navigating a rapidly evolving profession shaped by digital automation and complexity, developing these competencies in an Open Distance e-Learning (ODEL) context presents unique pedagogical challenges. Drawing on the constructivist research paradigm, the study examines accounting academics' perceptions of how *ubuntu* principles - such as respect, care, interconnectedness, and shared responsibility - can complement social constructivist approaches to enable deeper learning and support the development of the affective domain of critical thinking. A qualitative case study design was employed, with semi-structured face-to-face interviews conducted with 10 academics at a South African ODEL institution. Participants were purposively selected based on their expertise in first-year modules, decision-making roles, and insight into exit-level competencies. Data were analysed using a codes-to-themes-to-theory model, supported by ATLAS.ti™ 8. The findings revealed that *ubuntu* values align closely with constructivist principles such as social interaction, contextualised learning, and the centrality of student agency. When embedded in learning design and assessment, these values help cultivate critical thinking habits such as inquisitiveness, perseverance, honest reflection, adaptability, and an appreciation for diverse perspectives. The study concludes that *ubuntu*-informed pedagogy has the potential to humanise digital learning spaces, ground education in culturally resonant values, and support the development of empathetic and ethically aware graduates. The findings offer practical implications for educators and curriculum developers seeking to design inclusive and high-impact learning experiences in ODEL contexts. Furthermore, the study contributes to advancing indigenous knowledge systems in education and to the broader global discourse on fostering critical thinking in digitally mediated learning environments.

Keywords: Accounting education, African philosophy, Good thinking habits, Online learning, Value system

1. Introduction

Critical thinking has become an essential skill in accounting education (Terblanche, Van Rooyen and Enwereji, 2025), especially as the profession advances in response to rapid technological change and the increasing complexity of business environments (SAICA, 2023). First-year students studying financial accounting at online universities face a unique challenge as they must grasp technical content and learn to analyse, question, and apply that content in practical, often unpredictable scenarios. These skills cannot be developed solely through memorisation or passive learning (Paul and Elder, 2005). Instead, students need exposure to active learning strategies, problem-solving activities, and assessments that support deeper engagement. In an open distance e-learning (ODEL) environment, where students learn independently, fostering critical thinking becomes even more essential. The design of learning materials and assessments must stimulate higher-order thinking and equip students for real-world demands in the accounting profession (Bates, 2015).

The African worldview, particularly the philosophy of *ubuntu*, offers rich opportunities to enhance critical thinking in higher education. *Ubuntu*, which emphasises interconnectedness, shared humanity, and collective learning, aligns with social constructivist theories of learning that prioritise collaboration and dialogue (Letseka, 2016; Venter, 2004). In many African contexts, knowledge is constructed not in isolation but through communal experiences and reflective engagement. These practices closely mirror the pedagogical approaches of theorists such as Vygotsky, who argued that learning occurs most effectively through interaction with others (Harasim, 2012). *Ubuntu* has been increasingly embraced as a guiding philosophy at the ODEL institution for teaching, learning, student support, and assessment practices, making it an important foundation for curriculum development (Heydenrych, 2010). Integrating *ubuntu* principles into the online learning environment may not

only enhance critical thinking but also root learning in values that resonate with African students' cultural and social realities.

There is limited research on how an accounting curriculum develops critical thinking, especially in the early years of study, as much of the existing curriculum remains focused on technical knowledge rather than analytical and evaluative thinking (Rossouw and Steenkamp, 2025). Furthermore, the potential of indigenous knowledge systems and philosophies, such as *ubuntu*, to enrich the development of critical thinking in online learning remains largely untapped. The study examines accounting academics' perceptions of how the affective domain of critical thinking can be developed in an ODeL setting, and how the African philosophy of *ubuntu* influences critical thinking. The research question is:

How do accounting academics perceive the application of ubuntu principles in supporting the development of the affective domain of critical thinking?

The main research objective of this study is to explore accounting academics' perceptions of how the affective domain of critical thinking can be developed in an ODeL setting, and how the principles of the African philosophy of *ubuntu* - such as respect, care, interconnectedness, and shared responsibility - can complement social constructivist approaches to foster deeper learning. The study is motivated by the need to address the limited attention given to the affective dimension of critical thinking, particularly within online and distance education contexts, as well as the underutilisation of African philosophical perspectives in accounting education. The study seeks to enhance teaching and learning practices by providing insights into how relational and value-driven approaches can support student engagement, reflection, and ethical awareness in ODeL environments. In doing so, the study aims to contribute to the development of more holistic and contextually relevant teaching approaches that promote deeper learning. The recommendations of this study may inform curriculum design, academic development, and institutional strategies to strengthen critical thinking and student support in ODeL contexts, both within South Africa and in similar higher education settings globally.

2. Literature Review

The literature review of this study will expound on the definitions of critical thinking, good thinking habits (dispositions), critical thinking and accountancy, African philosophy, *ubuntu*, the context of study, and the concepts of constructivist learning theory.

2.1 Defining Critical Thinking

Critical thinking is recognised as essential for navigating the complexities of modern society (Anggraeni, Prahani, Suprpto, Shofiyah and Jatmiko, 2023) and has been widely studied across philosophy, psychology, and education (Lai, 2011; Lewis and Smith, 1993; Sternberg, 1986). Philosophical literature focuses on the qualities and characteristics that constitute the 'perfect' critical thinker (Lai, 2011), linked to critical thinking attitudes (Atabaki, Keshtiaray and Yarmohammadian, 2015), or good thinking habits. The psychology literature focuses on cognitive and metacognitive skills (Hepner, 2015), emphasising the cognitive processes underlying critical thinking and the development of thinking skills over time (Atabaki *et al.*, 2015; Lewis and Smith, 1993). Educational literature draws on insights from both philosophy and psychology to inform teaching and learning strategies (Terblanche, 2019).

A panel of experts (from numerous disciplines) conducted the APA Delphi study. It defined critical thinking as "purposeful, self-regulatory judgement which results in interpretation, analysis, evaluation and inference, as well as an explanation of the evidential, conceptual, methodological, criteriological or contextual considerations upon which that judgement is based" (Facione, 1990, p.2). Furthermore, the ideal critical thinker was described as –

[Someone who is] habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgements, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit (Facione, 1990, p.2).

The APA Delphi study pinpointed six fundamental cognitive skills for critical thinking (Facione, 1990). Each fundamental skill was further broken down into 16 non-hierarchical subskills, with certain competencies overlapping (Facione, 1990; Paul and Elder, 2005). Additionally, one should reflect on the quality of thinking, and thus, standards are required to measure the level of thinking. The universal standards include clarity, accuracy, precision, relevance, depth, breadth, logic, significance, and fairness (Elder and Paul, 2006). Within the

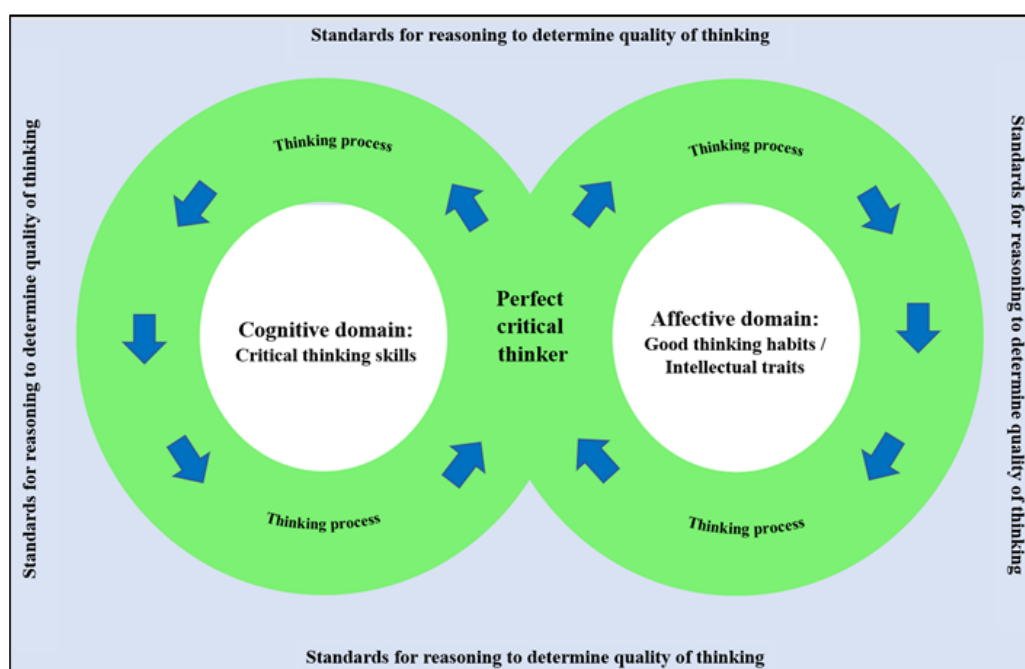
accounting profession, the South African Institute of Chartered Accountants (SAICA) states that critical thinking is: “evidenced by clarity, accuracy, precision, consistency, relevance, sound evidence, good reasoning, depth and breadth” (SAICA, 2017).

Since critical thinking encompasses more than just applying critical thinking skills (Halpern, 1998; Yang and Chou, 2008), the affective domain must also be considered, as cognitive skills are closely tied to it (Anderson and Krathwohl, 2001). Ricketts and Rudd (2004) found a positive correlation between analysis and good thinking habits, such as inquisitiveness, being well-informed, and diligence in seeking information. While cognitive skills can be developed, lacking the motivation or habits to use those skills will prevent one from becoming a true critical thinker (Facione, 2015; Halpern, 1998). Therefore, these dispositions must also be considered in critical thinking research.

2.2 Good Thinking Habits (Dispositions)

Critical thinking attitudes, often referred to in the literature as “dispositions”, “habits of the mind”, “traits of the mind” or “personal traits” (Facione, 1990, p. 11), involve testing beliefs, developing open-mindedness when drawing evidence-based conclusions, and applying methods of inquiry and reasoning for problem solving (Dewey, 1910). Good thinking habits consist of a combination of attitudes, intentions, values, and beliefs woven into a person’s character (Facione, 2000). Good thinking habits are defined as individuals’ willingness or motivation to engage in critical thinking; therefore, developing good thinking habits is a prerequisite for critical thinking to occur (Aybek, 2018; Facione, 2015; Halpern, 1998).

Figure 1 illustrates the “perfect” critical thinker, who integrates both cognitive and affective domains and evaluates thinking processes against universal standards for reasoning (see section 2.1), which then inform those processes.



Source: Authors’ own compilation

Figure 1: Critical thinking dimensions

Krathwohl, Bloom and Masia (1964) developed a framework for the affective domain, as depicted in Table 1 (Anderson and Krathwohl, 2001, p.xxvii), which identified receiving, responding, valuing, organisation, and characterisation as the important good thinking habits. Buissink-Smith, Mann and Shephard (2011) then identified the like-minded characteristics of each of the good thinking habits identified by Krathwohl *et al.* (1964). A total of 19 good thinking habits were identified in the APA Delphi study (Facione, 1990). The characteristics described in the Buissink-Smith *et al.* (2011) study were used to link the framework created by Krathwohl *et al.* (1964) to the good thinking habits identified in the APA Delphi study (Facione, 1990), as seen in Table 1 below:

Table 1: Comparison between the affective domain expressed in taxonomies and critical thinking attitudes or dispositions

Krathwohl <i>et al.</i> (1964)	Buissink-Smith, Mann and Shephard (2011) (adapted) Characteristics of attributes identified by Krathwohl <i>et al.</i> (1964)	Facione (1990) (adapted) Attitudes or dispositions expected from a critical thinker
Receiving	Students are willing to engage with new concepts and experiences and are willing to listen. They can ask questions, listen, participate, discuss, and acknowledge what they have received.	• Inquisitiveness – showing an interest in learning more about a wide array of concepts and/or issues. • Lifelong learning – demonstrating interest in becoming and remaining generally well-informed.
Responding	Students are active participants in the classroom. They can interpret information, clarify ideas, and question them.	• Able to recognise opportunities to make use of critical thinking skills. • Placing trust in the processes of reasoned inquiry. • Showing confidence in one's own reasoning skills. • Able to state the question or concern regarding information and/or opinions succinctly.
Valuing	Students value their own opinions and can express them. They can debate, evaluate critically, refute, and justify arguments.	• Being open-minded when confronted with antithetical worldviews. • Demonstrating flexibility through consideration given to alternative solutions and opinions. • Showing an understanding of the opinions of other people when, for example, evaluating arguments. • Searching assiduously for relevant information.
Organisation	Students develop a value system. Students can formulate arguments, defend their opinions, recognise relationships between sets of information, prioritise, contrast, and compare.	• Being fair-minded when judging reasoning. • Reflecting honestly upon one's own biases, prejudices, stereotypes, egocentric or sociocentric tendencies, one can determine the appropriate action to take. • Showing caution and/or care when delaying, making, or changing judgements. • Methodical when dealing with complex tasks. • Demonstrating good discernment during the selection and application of criteria. • Being conscientious when focusing on any task or problem. • Able to persevere when complexities or difficulties are encountered.
Characterisation	Students adopt a value system and consistently act in accordance with it. They can act in a certain manner and influence others.	• Willingly adapting one's own view after honest reflection had indicated that change is necessary. • Demonstrating meticulous accuracy to the degree that would be expected based on the subject and the circumstances.

Source: Dry (2021), as adapted

The good thinking habits identified by Facione (1990) (see Table 1) were used to craft the interview questions in this study to detect further good thinking habits and *ubuntu* for integration into the first-year financial accounting curriculum.

2.3 Critical Thinking and Accountancy

In accounting education, critical thinking is increasingly viewed as important (Wolcott and Sargent, 2021) due to the digitalisation and automation of routine tasks (Terblanche and De Clercq, 2021), and the evolving role of accountants in society (Powell and McGuigan, 2020). In line with the requirements of professional accounting bodies, accounting academics have traditionally focused on technical knowledge (Dyki, Singorahardjo and Cotronei-Baird, 2021) but have often neglected the development of critical thinking skills. However, professional accounting bodies, such as the Certified Public Accountants (CPA), CPA Canada, the International Accounting Education Standards Board (IAESB) and SAICA, now include critical thinking in their competency frameworks,

expecting students to develop these skills before entering the profession (American Institute of Certified Public Accountants [AICPA], 2023; CPA Canada, 2024; IAESB, 2019; SAICA, 2023; Wolcott and Sargent, 2021). Despite recognition of its importance, the question remains: How can accounting academics effectively cultivate critical thinking skills?

Several models and frameworks for critical thinking have been developed in the field of accountancy. Kimmel (1995) applied a critical thinking framework to confirm the importance of critical thinking within financial accounting, while Baril, Cunningham, Fordham, Gardner and Wolcott (1998) identified relevant critical thinking skills and dispositions for accounting professionals. More recently, Terblanche and De Clercq (2021) crafted a critical thinking competency framework for accounting students. All three studies highlighted the importance of both the cognitive (skills) and dispositional (good thinking habits) dimensions.

However, most accounting education research has focused on *defining* critical thinking and *identifying* aspects relevant to the discipline, rather than on *how* to implement and develop critical thinking in accounting courses (Terblanche and De Clercq, 2020; Wolcott and Sargent, 2021). Case studies dominate the limited research in this area (see de Souza Costa, de Sousa Gomes, Braunbeck & Eduarda Gomes Santana, 2018; Gunderson, 2013; Samkin and Keevy, 2019), with other interventions such as e-portfolios (Bodle, Malin and Wynhoven, 2016; Bradley and Kininmonth, 2025), gamification (Rosli, Khairudin and Saat, 2019), experiential learning (Butler, Church and Spence, 2019) and problem-based assessment tasks (Wilkin, 2016) receiving less attention. Papageorgiou (2023) provided students with explicit instruction in critical thinking, whereas Terblanche, Joint, Stiglingh, Du Preez and Diedericks (2025) combined explicit instruction with practical accounting activities. Scholars have predominantly explored the development of skills within the cognitive dimension of critical thinking, using Bloom's taxonomy (de Souza Costa et al., 2018) and analytical tasks (Gunderson, 2013). Turner and Tyler (2023) addressed both the cognitive and affective domains through integrated interventions across teaching, assessments and collaborative activities.

2.4 African Philosophy, *Ubuntu* and the Context of Study

Educational institutions around the world continue to explore diverse delivery modes to accommodate students' changing needs. These include face-to-face (on-campus), distance education, fully online, or blended learning approaches (Gosper, Ifenthaler and Kretzscheider, 2018). Operating as Africa's largest ODeL institution, the institution's educational model integrates the principles of distance education, open learning, and digital technologies to provide flexible access to higher education. Its strategic mission includes removing barriers to education, particularly for students from historically disadvantaged backgrounds, rural communities, and working professionals (Enwereji, Van Rooyen, Ngcobo and Musundwa, 2024). The ODeL institution aims to increase enrolment and to ensure meaningful learning by adopting pedagogical practices that support independent learning, critical engagement, and digital inclusion (Maboe, 2019).

The social constructivist learning theory underpins much of the pedagogical approach in the ODeL environment. This theory asserts that knowledge is actively constructed by students through meaningful interactions with peers, instructors, and learning resources (Anderson, 2010; Bates, 2015). Learning is not seen as a passive transfer of information but as a socially mediated process where knowledge emerges through dialogue, shared experiences, and collaborative problem-solving. This theory is particularly suitable for online learning environments, where physical presence is replaced by virtual engagement, requiring educators to design activities that foster collaboration, reflection, and interaction. Within this theoretical lens, students are encouraged to think critically, co-create knowledge, and apply what they learn in real-life situations; objectives that align closely with the vision of preparing future-ready graduates in fields such as accounting, business, and education.

This collaborative, relational view of learning aligns strongly with African philosophical traditions, such as the philosophy of *ubuntu*. Rooted in communalism, *ubuntu* emphasises interconnectedness, mutual respect, and the belief that individual identity and knowledge are shaped through relationships with others (Venter, 2004; Ukhun, 2009). As articulated by the African philosopher John Mbiti, "I am because we are, and since we are, therefore I am" (Goodreads, 2024). This worldview centres on the idea that knowledge is not individually owned but collectively developed to serve the broader community. *Ubuntu* goes beyond being a cultural philosophy; it functions as a value system that promotes dignity, compassion, and reciprocity in all human interactions (Kamwangamalu, 1999; Letseka, 2016; Msila, 2008). These principles are especially relevant in educational contexts where students come from diverse backgrounds and where collaboration and respect are crucial for success.

In an ODeL institution in South Africa, *ubuntu* can serve as both a philosophical foundation and a practical pedagogical tool. When embedded into online learning design, it fosters a “family-like” atmosphere where students feel seen, heard, and respected (Akinola and Uzodike, 2018). *Ubuntu* encourages learning spaces that are inclusive, dialogic, and emotionally safe, which mirrors the ideals of social constructivism. Knowledge emerges not only from cognitive engagement but also from ethical and relational interaction (Adyanga and Romm, 2022). As shown in Figure 2, *ubuntu* nurtures a culture in which students are encouraged to share ideas freely, listen to others with respect, and respond with kindness (Enslin and Horsthemke, 2016). These values are not just philosophical; they translate into practical approaches to student support, assessment, and curriculum design in online learning environments. When paired with constructivist learning principles, *ubuntu* creates a culturally grounded, human-centred approach that can transform how critical thinking is developed in first-year students, especially in disciplines like accounting, where ethical reasoning and collaboration are key.



Figure 2: *Ubuntu's* core values of humanness (Dry, 2021)

2.5 Constructivist Learning Theory

Constructivist learning theory holds that students actively construct knowledge through experience and reflection, rather than passively absorbing information. According to Piaget (1972), knowledge is not simply transmitted from teacher to student but actively built by the student as they engage with the world. Vygotsky (1978) expanded this notion by emphasising the importance of social interaction in cognitive development and introducing the concept of the Zone of Proximal Development (ZPD), in which students achieve more with the guidance of others than alone. Fosnot and Perry (2005) describe constructivism as a theory grounded in the belief that learning is a process of constructing meaning rather than merely acquiring facts. In the context of higher education, this approach encourages students to engage critically, explore ideas, ask questions, and co-create understanding through dialogue and reflection.

Key principles of constructivist learning include active learning, social interaction, contextualised experiences, and the importance of prior knowledge (Jonassen and Rohrer-Murphy, 1999). Students are seen as central agents in the learning process, with teachers acting as facilitators rather than mere transmitters of information. Learning is viewed as a continuous, developmental process that builds on the student's existing cognitive structures. Constructivist pedagogy also values collaboration and encourages students to work in groups, share perspectives, and engage meaningful, relevant tasks (Bruner, 2004). Assessment under constructivism shifts away from standardised testing toward authentic, formative evaluations that reflect students' understanding and problem-solving capabilities in real-world contexts.

The principles of constructivist learning theory closely align with the African philosophy of *ubuntu*, which promotes collective learning, mutual respect, and human dignity. *Ubuntu*, often summarised as "I am because

we are," fosters a sense of interconnectedness and community in which learning is seen as a shared journey rather than an individual pursuit (Letseka, 2016). Just as constructivism emphasises social collaboration and co-construction of knowledge, *ubuntu* stresses the value of communal wisdom, shared experience, and relational accountability in the learning process. In combining these frameworks, educators can create culturally relevant, inclusive learning environments that honour both cognitive development and the communal spirit of African students. Thus, applying constructivist principles through an *ubuntu* lens offers a powerful, contextually meaningful approach to fostering critical thinking and lifelong learning, particularly in the South African ODeL context.

3. Research Method

This study adopted the constructivist research paradigm because the researchers aimed to understand and interpret how academics currently support the development of critical thinking to improve practice (Lincoln, Lynham & Guba, 2011). As a case study allows an in-depth focus on a single unit of analysis (Creswell and Poth, 2018), a qualitative case study with semi-structured interviews with academics teaching financial accounting was chosen. The selection of a qualitative research approach enabled in-depth insights into academics' experiences, perceptions, and the influence of *ubuntu* philosophy on the development of critical thinking in an online financial accounting environment.

Since this research project involved research participants from the ODeL institution, ethical clearance was obtained from the Research Ethics Review Committee. After obtaining ethical approval and an ethics number, the researchers sought participants' consent before they participated in the study. In this case, participation was entirely voluntary, and individuals were free to withdraw at any point without consequence. To ensure protection against harm, interviews were conducted face-to-face to provide a safe, comfortable setting, with researchers prioritising participants' emotional and psychological well-being. Confidentiality and anonymity were maintained by securing raw data with passwords and using pseudonyms (e.g., P1, P2) during reporting. Lastly, all sources used in the study were properly acknowledged to avoid plagiarism.

Semi-structured, in-depth interviews were used in this study to gain rich, participant-driven insights into the research topic, allowing for flexibility through follow-up questions based on individual responses. Rooted in a conversational, interactive approach, the interviews aimed to understand the topic from participants' perspectives. To ensure participant comfort and reduce potential discomfort, individuals were allowed to choose the interview date, time, and location. Interviews ranged from 12 minutes to 1 hour, depending on how much participants wished to share.

Two first-year financial accounting modules were chosen. Module 1 focuses mainly on basic accounting principles, concepts and procedures and Module 2 on basic financial accounting and reporting. These modules were chosen because they feed into the BCompt degree accredited by SAICA and must implement the CA competency framework (CA of the Future), which includes the development of critical thinking skills (SAICA, 2023). As depicted in Figure 3, five academics were purposively selected based on their in-depth knowledge of these two modules; a further three were selected based on their decision-making roles as leaders; and a further two academics from second-year modules were selected for their insight into the exit-level skills required during the first year. The 10 participants, therefore, provide us with a composite view.

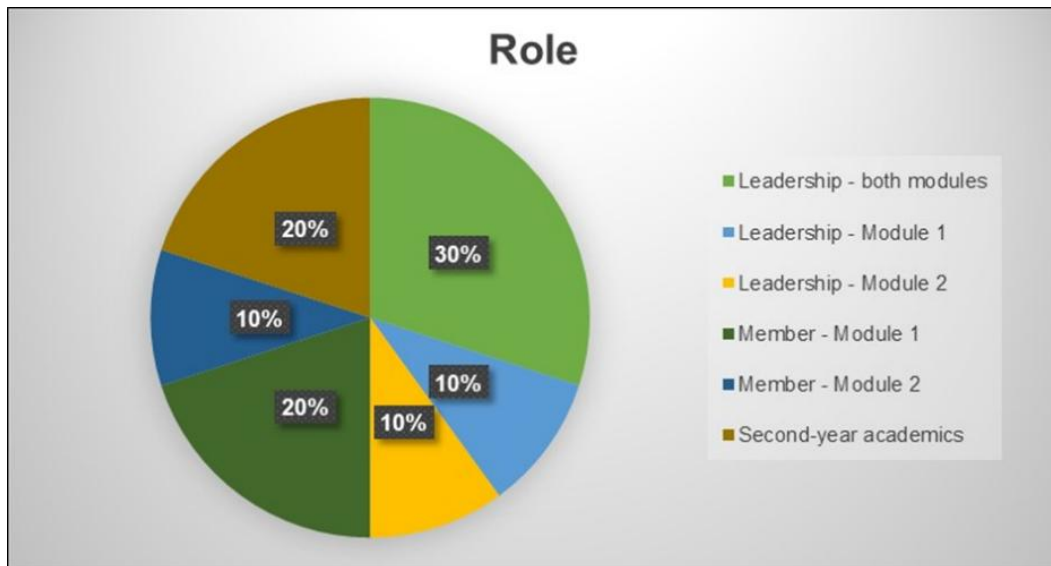


Figure 3: Roles of participants (Dry, 2021)

The interview protocol was developed with the research question in mind: namely, *how the principles of ubuntu could be applied to assist in the development of the affective domain of critical thinking*. To explore this, participants were asked the following interview questions:

1. Which thinking habits should be developed in first-year financial accounting to enable students to become good critical thinkers?
2. How is the teaching and learning environment contributing towards the development of good thinking habits in first-year financial accounting students?
3. How are the principles of the African philosophy of *ubuntu* applied to assist with the development of the thinking habits you identified earlier?

The study ensured trustworthiness (Lincoln and Guba, 1985) by maintaining participant anonymity, using a pre-determined interview schedule, and utilising sound research methods. Ethical clearance was obtained at both the institutional and the college level. Interviews were electronically recorded in quiet offices (Creswell and Poth, 2018) and transcribed using Trint.com (see: <https://trint.com/?cn-reloaded=1>). The main researcher reviewed each transcript whilst listening to the recordings for accuracy, followed by member checking with participants (Harper and Cole, 2012). An independent coder (IC) conducted parallel, yet autonomous, coding of the transcripts.

Saldaña's (2016) codes-to-themes-to-theory model was used to analyse the empirical data. A codebook (see Saldaña, 2016) was developed from the literature review, with additional codes added after reviewing the transcripts. This direct content analysis (Hsieh and Shannon, 2005) allowed for theory-concept-driven coding. Prefix coding was used to categorise the data, followed by specific analytical phrases to reveal participants' perspectives. The researchers followed the codebook for initial coding, then applied focused, pattern coding. From the final cycle of codes, categories were theoretically identified, leading to themes (Saldaña, 2016). The IC was also asked to cross-check the coding using the detailed codebook (Barbour, 2001).

The IC used qualitative data analysis software (QDAS) ATLAS.ti™ 8. During the first cycle, deductive coding based on the researchers' codebook was applied, with larger quotations lumped to ensure confirmability and meaningful analysis and to add inductive value (Saldaña, 2016). In later cycles, the IC refined and stabilised the codes. The researchers compared the IC's codes, integrating them where relevant to confirm trustworthiness. The researchers then used a combination of coding sources to create categories (as an intermediate step and not covered in this article) and to develop theoretically driven themes. An example of the trajectory is in Figure 4, which will be discussed in the next section.

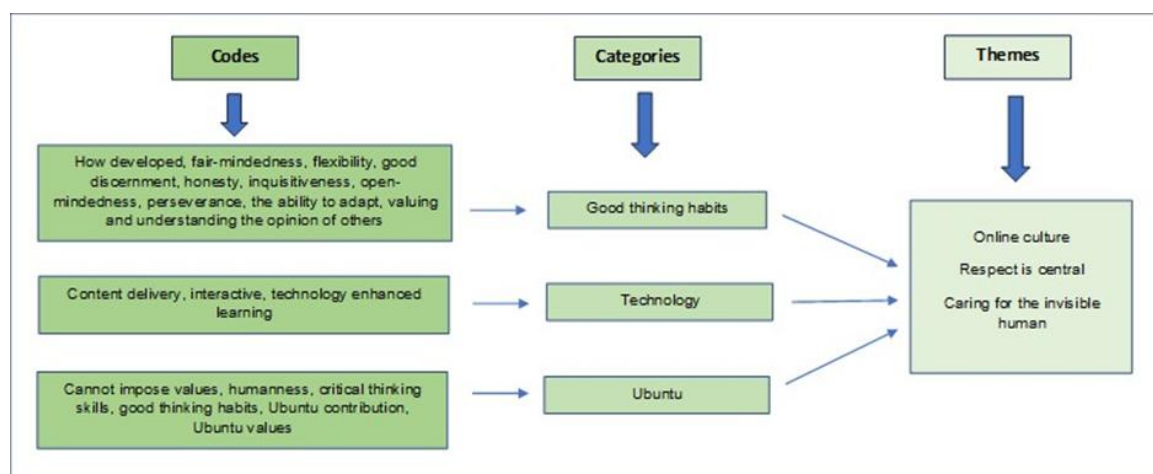


Figure 4: Example of trajectory from codes to categories and then themes

4. Findings

Three themes were generated. These are 1) the fostering of an online culture; 2) the centrality of respect within collaborative online spaces; and 3) acknowledging the responsibility to care for other human beings, even though we cannot always see them when using technology. In the next sections, the discussion will focus on the contribution of *ubuntu* to create an online culture and the values of *ubuntu*, respect and care, which were linked to the following good thinking habits: inquisitiveness, open-mindedness, adaptability, perseverance, and valuing and understanding the opinions of others.

4.1 Theme 1: Online Culture

Ubuntu is an African philosophy that is deeply ingrained in African culture (Letseka, 2016). According to P9, to understand and appreciate *ubuntu*, “[One should] expose yourself to that culture”. Although *ubuntu* is not a financial system but rather a value system (Kamwangamalu, 1999; Letseka, 2016; Msila, 2008), there is an opportunity to apply its principles within the financial accounting domain and in online classrooms.

Participant 1 cautioned that one cannot impose values and beliefs on people:

[Y]our value system is very closely related to what you believe or how you believe or if you even believe, but that will also guide you in what your values are, and that will be what you then carry forward into an online [...] People can tell you and try to police you, but a value system is something that comes from within. So even though the university can say, “we accept this as our values”, it is still something that every individual must make their own, and they have a choice because they could, or they could not.

However, P8 argued that *ubuntu* is not an African concept, but rather a universal human concept that spans across cultures and religions. According to P9, *ubuntu* is the obligation to treat people as human beings with compassion and care. As the principles of *ubuntu* are not unfamiliar to other cultures and beliefs, there is potential to find common ground by applying them across the diversity of online classrooms.

Ubuntu, which underpins the concept of humanness (P8), is supported by prior research (Ekoh and Warriar, 2023; Kamwangamalu, 1999; Letseka, 2016; Msila, 2008). Therefore, it can be argued that students, as human beings, can relate to these principles. *Ubuntu*’s emphasis on respect is linked to the cultivation of good thinking habits.

4.2 Theme 2: Respect is Central

In the online environment, the students must be made aware of the culture expected from them (P6). The online learning communities we create must be facilitated to instil respect, as intolerance and even hate speech are present (P8). Participant 9 made the following profound statement about technology causing people to forget that the interaction taking place is still with a human being:

[When you] go to technology and you start using technology, and with the anonymity of technology, people tend to misbehave a lot. [...] That, [...] principle there, [...], that, [...], people are gonna [...] respect each other when they're using the technology, even though people can't see you. [...] and just know that there's actually a human being behind.

The learning communities we create should, therefore, be built on a foundation of respect, where all questions are welcomed (P2 and P8). Inquisitiveness has been identified by Turner and Tyler (2023) as one of the good thinking habits that could be cultivated in first-year financial accounting. Participant 3 mentioned that respect fosters *inquisitiveness*, as students are more inclined to ask questions to seek clarity. When feeling respected, students will feel comfortable with asking questions, causing students to

[come] back to ask more questions, which automatically links to perseverance and reading other people's responses and not being so critical about how they're responding, but learning to take the information, you know, self-assess it, put something back into the online platform where everybody else is not criticizing your information because the lecturer would be facilitating [...] (P8)

Students regularly contact academics privately, stating, "I'm struggling with this" (P5). The role of academics guiding students through content and creating a safe environment where students feel respected is crucial for fostering collaboration and perseverance. Feeling respected gives students the confidence to engage, even when they face difficulties. Perseverance – or in Facione's (1990) words "persistence through difficulties are encountered" – is the good thinking habit students need during times of struggle. First-year academics should support students by setting goals, providing timely feedback, and motivating them to "keep on going and inspire them to carry on" (P2). Participant 8 described perseverance as a "huge one". The reason perseverance is so important is that

[P]erseverance is going, I'm wrong, but I've got to get up and keep going, which is the nature of business. So, if you're not going to learn perseverance, you're not going to get through this, the grief, for starters, because there's a lot of knocking down that here. And even in the assessment methods [...] If you get it wrong, it comes back, and you get it wrong, and it keeps coming back. So, your ability to get up and do it again and again and again till you get it right.

Respecting the diverse cultures within the virtual space's vast diversity is very important, as others bring unique perspectives that can enrich the learning experience for all students (P7). One way to show respect to others is to listen (or read feedback on an online discussion tool) their opinions (P4). "[I]f you are a good listener [...] then you can also accept [...], constructive criticism and then [...] you are able to apply your mind constructively" (P4). It is not surprising that two participants (P1 and P6) specifically linked respect to the development of the good thinking habit of *valuing and understanding the opinions of others*.

Participant 6 continued the argument and stated that, through respect, one would value others' opinions, which would then make it much easier to *adapt* one's own thinking with the opinion of others. Financial accounting students need to be equipped to operate amid a rapidly changing, technologically advanced knowledge society (Boyce, 2014; Cunningham, 2014; SAICA, 2023). The financial accounting curriculum is regularly updated to reflect financial accounting standards that are continually evolving to keep pace with the changing world, leaving students confronting regular changes in their knowledge base (P6). The ability to respect this and adapt is a non-negotiable skill forced upon all students by the way our universities function (P8), for example, adapting to how examinations are structured rather than how we teach (P5). When the need arises to adapt, it is important to honestly face one's own biases, prejudices, and stereotypes (Facione, 1990) in a respectful way. Participant 9 highlighted the need for "*honest* evaluation of knowledge and thought processes from the very beginning".

4.3 Theme 3: Caring for the "Invisible" Human

The theory of social constructivism proposes that knowledge is constructed through reflection and social engagement, in which different views are shared and validated through collaborative endeavours (Bates, 2015). The advancement of critical thinking, particularly in gaining understanding, is fostered by *ubuntu* during social interaction. The essence of *ubuntu* enriches this process by establishing a familial ambiance (Akinola and Uzodike, 2018), thereby creating a safe social environment where students can feel free to engage with one another.

"Do not leave any student behind" (P3). Caring involves responding to students' questions, as academics want to see them succeed (P5). Participant 3 emphasised fostering the *ubuntu* principles to cultivate a culture of care in online collaboration spaces. Though emotional responses may be invisible in virtual collaboration, they still exist, making it essential to create a safe environment where academics facilitate care. This enables students to feel secure in asking questions and sharing diverse perspectives. Participant 1 suggested using discussion forums to provoke students' *inquisitiveness*. However, cultivating an *ubuntu* culture takes time; attempting to do so is the first step towards fostering collaboration rather than being fixated solely on correct and incorrect answers (P8).

When academics and students start sharing their knowledge generously, it demonstrates “caring in such a way [that] they are able to ask each other questions which help them to reason better” (P3). It is through caring that we can unite different cultures and foster the sharing of differing viewpoints (P7), from which certain good thinking habits, such as inquisitiveness, valuing and understanding the opinions of others, honest reflection, adaptability, and perseverance, can be cultivated.

In summary, even though *ubuntu* is deeply rooted in African philosophy, its value system transcends cultural boundaries and offers valuable insights beyond its origins. To establish an online environment conducive to collaborative learning, respect is essential. Instilling these values within technology-mediated environments will help maintain participants’ recognition as human beings. Emphasising the importance of caring for one another fosters an inclusive learning space where students are supported and empowered to engage with diverse perspectives. By embracing the values of *ubuntu*, a culture of empathy, resilience, and inclusivity can be fostered, ultimately enriching the learning journey for all students.

5. Discussion of Results

The first theme highlights the importance of cultivating a values-based online learning culture rooted in *ubuntu* philosophy. Participants noted that while *ubuntu* is not a financial or academic system, its core values such as humanness, compassion, and collective responsibility (Venter, 2004) can be meaningfully applied to virtual learning spaces, particularly in first-year financial accounting. *Ubuntu* offers a culturally grounded framework that supports the development of a supportive, inclusive online culture where students feel connected despite physical separation (Ukhun, 2009). Participants emphasised that adopting such values must come from within and cannot be imposed; nonetheless, the universal nature of *ubuntu* provides common ground across diverse student populations. This theme also suggests that integrating *ubuntu* into the online learning environment can help students see beyond rigid content and focus on the humanity behind digital interactions. It calls on academics to intentionally create a learning atmosphere that embraces diversity, encourages shared responsibility, and builds a sense of belonging on critical elements for students steering the transition into complex disciplines like accounting.

Respect emerged as a foundational value in online learning communities, with participants emphasising its central role in fostering inquisitiveness, perseverance, and adaptability which are the core habits of critical thinkers. In virtual spaces where anonymity can lead to disengagement or incivility, maintaining a culture of respect ensures that students feel safe to participate, ask questions, and admit when they do not understand a concept (Cress, Collier and Reitenauer, 2023; Liang *et al.*, 2020). The interviews revealed that respect enhances student confidence and promotes meaningful dialogue, peer learning, and self-reflection. Students who feel respected are more likely to persist through challenges and value the perspectives of others. Moreover, respect was linked to adaptability, especially in a field like financial accounting that requires constant updating of knowledge in response to changing standards and practices. The ability to accept feedback, adapt to new information, and navigate diverse viewpoints is vital in such contexts. This is similar to the findings of Turner and Tyler (2023), who found that inquisitiveness and openness to divergent worldviews are important habits for first-year accounting students. Therefore, this theme highlights that respect is not merely a social nicety but a strategic tool for cultivating higher-order thinking and resilience in online learning.

The third theme centers on the ethic of care, particularly in relation to students’ often “invisible” emotional experiences in online environments. Participants spoke passionately about the need to foster an *ubuntu*-informed culture of care where every student is seen, supported, and encouraged, even in the absence of face-to-face interaction. The findings suggest that caring is expressed through timely academic support, responsiveness to questions, and the creation of collaborative spaces where students feel safe to share and explore different perspectives. This emotional and intellectual care, according to Terblanche *et al.* (2025), enhances engagement and promotes the critical thinking habit of inquisitiveness. Educators can foster a climate that bridges the emotional distance often associated with online learning by facilitating reflective discussions and encouraging students to ask questions without fear of judgment. Participants also emphasised that true caring involves guiding students through difficulties and instilling the perseverance needed to succeed. In this sense, caring is not just emotional support; it is a strategic act that empowers students to think critically, challenge assumptions, and continue growing within a nurturing academic community (Adeoye, 2024).

Underlying all three themes is the recognition that both *ubuntu* and the constructivist paradigm offer complementary foundations for fostering critical thinking in online financial accounting education. Constructivism holds that knowledge is actively constructed through experience, interaction, and reflection (Bada and Olusegun, 2015), principles that closely align with *ubuntu*’s emphasis on communal engagement,

empathy, and mutual respect. In this study, the constructivist lens enabled an understanding of how students' thinking habits are shaped by content delivery and by their lived experiences within digitally mediated learning communities. *Ubuntu* deepens this framework by adding a moral and cultural dimension, reminding educators that students are not isolated data receivers, but interconnected human beings whose growth depends on emotional safety, inclusivity, and relational accountability. Together, these two frameworks fostered online environments where students co-constructed knowledge, practiced reflective reasoning, and developed dispositions such as perseverance, adaptability, and inquisitiveness. This integrated approach demonstrates how culturally rooted values and learning theories can work together to humanise online education and build resilient, thoughtful students.

6. Contribution of the Study

The contributions of the study are outlined below:

6.1 Theoretical Contribution

The study extends the existing critical thinking literature by foregrounding the *affective domain*, which is often underexplored relative to cognitive skills. By integrating *ubuntu* with social constructivist perspectives, the study offers a more holistic and relational understanding of how online learning environments may support the development of critical thinking dispositions.

6.2 Contextual (African-centred) Contribution

The study makes an African-centred contribution by bringing an African philosophical lens (*ubuntu*) into accounting education research. Although the study acknowledges shared values of humanness across cultures, it demonstrates how culturally specific philosophies can provide contextually relevant insight for higher education. In this way, the study contributes to the contextualisation of teaching and learning and to the broader conversations about decolonisation.

6.3 Pedagogical Contribution

Insights are provided into how values such as respect, care, interconnectedness, and shared responsibility can be embedded in teaching practices to support deeper learning, particularly in online and distance environments.

6.4 Practical Implications for ODeL

The study offers guidance for lecturers in ODeL settings by highlighting the importance of respectful interaction, timely support and a sense of human connection in physically distanced learning environments. These insights are particularly relevant in online contexts, where developing the affective domain may be more difficult because students can easily feel isolated.

6.5 Empirical Contribution

By drawing on accounting academics' perceptions, the study contributes discipline-specific insights to the broader discourse on critical thinking and teaching practices. It adds to the understanding of how academics view the role of *ubuntu*-informed values in fostering critical thinking dispositions in financial accounting education.

6.6 Contribution to Professional Formation

The study suggests that fostering the affective domain through *ubuntu*-informed values may contribute to the development of ethically aware, reflective and socially responsible accounting graduates. In particular, the emphasis on respect, care, adaptability, and perseverance points to qualities that are important not only for academic success but also for professional success.

7. Recommendations

The recommendations of the study are outlined below:

7.1 Recommendations for Higher Education Institutions

Higher education institutions should intentionally incorporate culturally grounded philosophies, such as *ubuntu*, into their teaching and learning frameworks to create more inclusive online environments. Academics should be encouraged and supported through professional development programmes and workshops to understand better how to embed *ubuntu* values, such as respect, interdependence, and care, into their teaching practices.

Doing so will promote a sense of belonging among students from diverse backgrounds and also affirm their identities in the learning process. Inclusive pedagogy fosters a classroom culture where students feel acknowledged and safe, which is vital for nurturing good thinking habits and critical engagement, particularly in the early stages of their academic journeys. Further empirical research should focus on determining which ubuntu-informed teaching practices most effectively foster students' sense of belonging, affirm their identities and promote online engagement. Longitudinal studies are needed to determine whether ubuntu-based pedagogical interventions produce sustained improvement in critical engagement across diverse student groups.

7.2 Recommendations for Educators

Educators should revise curriculum content and teaching strategies to reflect the communal and collaborative ethos of *ubuntu*. This can be operationalised through group-based projects, peer learning sessions, collaborative assessments, and online discussion forums that emphasise mutual respect and shared responsibility. Such practices promote collective reasoning, openness to diverse perspectives, and a deeper understanding of subject matter, which are essential components of critical thinking. Embedding these values into course design will encourage students to go beyond individual performance and instead contribute to a supportive learning community where academic and ethical growth intertwine. Future research could investigate the impact of ubuntu-aligned curricula and assessment designs on the quality of collaborative learning, critical thinking and academic performance. Furthermore, it should be investigated whether this approach offers demonstrable advantages over established collaborative learning frameworks.

7.3 Recommendations for Online Learning Environments

To effectively nurture critical thinking among first-year students, educators must create online learning environments that encourage open dialogue, curiosity, and resilience. This includes designing opportunities for students to engage in reflective journaling, brainstorming activities, and low-stakes formative assessments where they can make mistakes and learn from them without fear. Academics should prioritise prompt feedback, encourage student questions, and cultivate a culture in which perseverance through academic struggle is normalised and supported. Educators can help students build the confidence and cognitive flexibility needed to thrive in academia by aligning the online learning environment with *ubuntu* values such as care and respect. Further studies are needed to explore the effects of ubuntu-informed online learning environments on psychological safety, intellectual risk-taking and the development of good thinking habits.

8. Conclusion

This study set out to understand how *ubuntu* values can help foster critical thinking among first-year financial accounting students in an online environment. The findings showed that values like respect, care, and collective responsibility help students feel more connected and supported in virtual spaces. When students feel safe and respected, they are more likely to ask questions, reflect deeply, and engage with different viewpoints. The use of the constructivist paradigm helped uncover how students learn through interaction, shared experiences, and reflection. *Ubuntu* strengthened this process by adding a cultural and human touch to learning. Together, these approaches created an environment where students could grow intellectually and emotionally. Although the study focused on one group of students, the insights are relevant beyond this setting. *Ubuntu* offers a powerful reminder that education should not lose its humanity, especially in online spaces. Teaching with compassion, care, and respect helps shape students who think critically and act ethically.

AI Statement: The authors declare that they have not used generative or assisted artificial intelligence tools for content creation at any stage of the paper's conception and revision. All content presented is the result of the authors' independent work, which guarantees originality, integrity, and compliance with ethical and scientific principles. AI was solely used to improve language and grammar.

Ethics Statement: Prior to data collection, ethical approval was obtained from the Ethics Committee under the reference number 2019_RPSC_050.

References

- Adeoye, M.A., 2024. Education in flux: nurturing minds for the future. In *Proceeding of International Conference on Education, Society and Humanity*, 2(1), pp. 73–82.
- Adyanga, F.A.A. and Romm, N.R.A., 2022. Reflections upon our way of invoking an indigenous paradigm to co-explore community mobilization against irresponsible practices of foreign-owned companies in Nwoya district, Uganda. *The Qualitative Report*, 27(7), pp.1359–1389. doi: 10.46743/2160-3715/ 2022.5147.

- Akinola, A.O. and Uzodike, U.O., 2018. Ubuntu and the quest for conflict resolution in Africa. *Journal of Black Studies*, 49(2), pp.91–113. doi: 10.1177/0021934717736186.
- American Institute of Certified Public Accountants (AICPA)., 2023. Uniform CPA Examination Blueprints. Effective date: January 1, 2025. Available at <<https://future.aicpa.org/resources/article/learn-what-is-tested-on-the-cpa-exam>> [Accessed 14 November 2024].
- Anderson, L. and Krathwohl, D. (Eds.), 2001. *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.
- Anderson, T., 2010. Theories for learning with emerging technologies. In G. Veletsianos (ed.). *Emerging technologies in distance education*. Athabasca: Athabasca University Press, pp.23–39.
- Anggraeni, D.M., Prahani, B.K., Suprpto, N., Shofiyah, N. and Jatmiko, B., 2023. Systematic review of problem based learning research in fostering critical thinking skills. *Thinking Skills and Creativity*, 49(101334). doi: 10.1016/j.tsc.2023.101334.
- Atabaki, A.M.S., Keshtiaray, N. and Yarmohammadian, M.H., 2015. Scrutiny of critical thinking concept. *International Education Studies*, 8(3), pp.93–102. doi: 10.5539/ies.v8n3p93.
- Aybek, B., 2018. An examination of the relationship between the critical thinking dispositions of prospective teachers and their attitudes toward multicultural education. *Yükseköğretim Dergisi*, 8(3), pp.282–292. doi:10.2399/yod.18.012.
- Bada, S.O. and Olusegun, S., 2015. Constructivism learning theory: A paradigm for teaching and learning. *Journal of Research and Method in Education*, 5(6), pp.66–70. doi: 10.9790/7388-05616670.
- Barbour, R.S., 2001. Checklists for improving rigour in qualitative research: A case study of the tail wagging the dog?, *British Medical Journal*, 322(7294), pp.1115–1117.
- Baril, C.P., Cunningham, B.M., Fordham, D.R., Gardner, R.L. and Wolcott, S.K., 1998. Critical thinking in the public accounting profession: Aptitudes and attitudes. *Journal of Accounting Education*, 16(4), pp.381–406. doi: 10.1016/S0748-5751(98)00023-2.
- Bates, A.W., 2015. *Teaching in a digital age*. BCCampus Open Education. Available at <<https://opentextbc.ca/teachinginadigitalage/open/download?filename=Teaching-in-a-Digital-Age-1455136052andtype=pdf>> [Accessed 14 November 2024].
- Bodle, K.A., Malin, M. and Wynhoven, A., 2017. Students' experience toward ePortfolios as a reflective assessment tool in a dual mode indigenous business course. *Accounting Research Journal*, 30(3):333–350. doi: 10.1108/ARJ-06-2015-0089
- Boyce, G., 2014. Ethics and accounting education. In M.S. Wilson (ed.). *The Routledge companion to accounting education*. Abingdon: Routledge, pp.533–558.
- Bradley, L. and Kininmonth, K.W., 2025. Using an eportfolio to help undergraduate accountancy students develop deep-learning and critical thinking skills. *Accounting Education*, pp.1-32. doi: 10.1080/09639284.2025.2586573
- Bruner, J., 2004. James's varieties and the "new" constructivism. In *William James and a Science of Religions: Reexperiencing the varieties of religious experience* (pp. 73-85). Columbia University Press.
- Buissink-Smith, N., Mann, S. and Shephard, K., 2011. How do we measure affective learning in higher education? *Journal of Education for Sustainable Development*, 5(1), pp.101–114. doi: 10.1177/09734082100050011
- Butler, M.G., Church, K.S., and Spencer, A.W. 2019. Do, reflect, think, apply: Experiential education in accounting. *Journal of accounting education*, 48, 12–21.
- CPA Canada, 2024. The Chartered Professional Accountant competency map: Understanding the competencies a candidate must demonstrate to become a CPA. Available at <<https://www.cpacanada.ca/en/become-a-cpa/why-become-a-cpa/the-cpa-certification-program/the-cpa-competency-map>> [Accessed 14 November 2024].
- Cress, C.M., Collier, P.J. and Reitenauer, V.L., 2023. *Learning through serving: A student guidebook for service-learning and civic engagement across academic disciplines and cultural communities*. Taylor and Francis.
- Creswell, J.W. and Poth, C.N., 2018. *Qualitative inquiry and research design: Choosing among five approaches*. London: Sage.
- Cunningham, B.M., 2014. Developing critical thinking in accounting education. In M.S. Wilson (ed.). *The Routledge companion to accounting education*. Abingdon: Routledge, pp.399–419.
- De Souza Costa, P. de Sousa Gomes, G. Braunbeck, G.O. and Eduarda Gomes Santana, M. 2018. A Safari in Brazil: evidence regarding the framework-based approach to teaching. *Revista Contabilidade & Finanças*, 29(76), 129–147. doi: 10.1590/1808-057x201804760
- Dewey, J., 1910. *How we think*. Boston, MA: Heath.
- Dry, H.J., 2021. The development of critical thinking in the first-year financial accounting curriculum at an open distance and e-learning institution in South Africa. Unpublished Master dissertation. Pretoria: University of South Africa.
- Dyki, M., Singorahardjo, M. and Conronei-Baird, V.S., 2021. Preparing graduates with the employability skills for the unknown future: reflection on assessment practice during Covid-19. *Accounting Research Journal*, 34(2), pp.229–245. doi: 10.1108/ARJ-09-2020-0285.
- Ekoh, P.C. and Warri A., 2023. Ubuntu and forced migration: reflections for social work engagement with older refugees in Canada. *African Journal of Social Work*, 13(1), pp.1–11. <https://dx.doi.org/10.4314/ajsw.v13i1.1>.
- Elder, L., and Paul, R., 2006. *Critical thinking: Learn the tools the best thinkers use*. Maryland: Rowman and Littlefield.
- Enslin, P. and Horsthemke, K., 2016. Philosophy of education: Becoming less Western, more African? *Journal of Philosophy of Education*, 50(2), pp.177–190. doi: 10.1111/1467-9752.12199.

- Enwereji, P.C., Van Rooyen, A.A., Ngcobo, R.N. and Musundwa, F.S., 2024. Empowering e-Tutors for Effective Online Teaching and Learning: E-tutors' Challenges and Perceptions. *International Journal of Learning, Teaching and Educational Research*, 23(1), pp.374–399. <https://doi.org/10.26803/ijlter.23.1.19>.
- Facione, P.A., 1990. *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Executive summary of "The Delphi report"*. Fullerton, CA: California Academic Press.
- Facione, P.A., 2000. The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal Logic*, 20(1), pp.61–84. <https://doi.org/10.22329/il.v20i1.2254>.
- Facione, P.A., 2015. *Critical thinking: What it is and why it counts*. Insight Assessment. Available at <https://www.researchgate.net/profile/Peter_Facione/publication/251303244_Critical_Thinking_What_It_Is_and_Why_It_Counts/links/5849b49608aed5252bcbe531/Critical-Thinking-What-It-Is-and-Why-It-Counts.pdf> [Accessed 14 November 2024].
- Fosnot, C.T. & Perry, R.S., 2005. Constructivism: A psychological theory of learning. In C.T. Fosnot (Ed.) *Constructivism: Theory, perspectives, and practice (2nd ed.)*, New York: Teachers College Press, pp.8-38.
- Goodreads, 2024. *John S. Mbiti. Quotes*. Retrieved from <https://www.goodreads.com/quotes/111766-i-am-because-we-are-and-since-we-are-therefore> (Accessed 21 April 2024).
- Gosper, M.B.M., Ifenthaler, D. and Kretzscha, C.W.M., 2018. On-campus, distance or online? Influences on student decision-making about study modes at university. *Australian Journal of Educational Technology*, 34(5), pp.72–85. doi: <https://doi.org/10.14742/ajet.3781>.
- Gunderson, K.E. 2013. Distinguishing between liabilities and equity: Two mini-cases for improving students' critical thinking in intermediate financial accounting. *Journal of the International Academy for Case Studies*, 19(3):51–63.
- Halpern, D.F., 1998. Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist*, 53(4), pp.449–455. doi: 10.1037//0003-066x.53.4.449.
- Harasim, L., 2012. *Learning theory and online technologies*. New York, NY: Routledge.
- Harper, M. and Cole, P., 2012. Member checking: Can benefits be gained similar to group therapy? *The Qualitative Report*, 17(2), pp.510–517. doi: 10.46743/2160-3715/2012.2139.
- Hepner, M.R., 2015. The erosion of critical thinking development in post-secondary education: The need to return to liberal education. In S. Wisdom and L. Leavitt (eds.). *Handbook of research on advancing critical thinking in higher education*. Hershey: IGI Global, pp.68–96.
- Heydenrych, J.F., 2010. *Conceptual framework for student support at Unisa*. Unisa. Available at <<http://uir.unisa.ac.za/handle/10500/3292>> [Accessed 24 July 2025].
- Hsieh, H. and Shannon, S.E., 2005. Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), pp.1277–1288. doi: 10.1177/1049732305276687.
- International Accounting Education Standards Board (IAESB), 2019. *Handbook of international education requirements 2019 Edition*, New York: International Federation of Accountants. Available at <<https://www.iaesb.org/publications/2019-handbook-international-education-standards>> [Accessed 28 June 2021].
- Jonassen, D.H. and Rohrer-Murphy, L., 1999. Activity theory as a framework for designing constructivist learning environments. *Educational technology research and development*, 47(1), pp.61-79. <https://doi.org/10.1007/BF02299477>
- Kamwangamalu, N.M., 1999. Ubuntu in South Africa: A sociolinguistic perspective to a pan-African concept. *South-North Cultural and Media Studies*, 13(2), pp.24–41 doi: 10.1080/02560049985310111.
- Kimmel, P., 1995. A framework for incorporating critical thinking into accounting education. *Journal of Accounting Education*, 13(3), pp.299–318. doi: 10.1016/0748-5751(95)00012-B.
- Krathwohl, D.R., Bloom, B.S. and Masia, B.B., 1964. *Taxonomy of educational objectives: The classification of educational goals. Handbook II: Affective domain*. New York, NY: Longman.
- Lai, E.R., 2011. *Critical thinking: A literature review*. Research report. Pearson. Available at <<http://images.pearsonassessments.com/images/tmrs/CriticalThinkingReviewFINAL.pdf>> (Accessed 11 May 2019).
- Letseka, M., 2016. *Open distance learning through the philosophy of ubuntu*. New York, NY: Nova Science.
- Lewis, A. and Smith, D., 1993. Defining higher order thinking. *Theory into Practice*, 32(3), pp.131–137. <http://www.jstor.org/stable/1476693>.
- Liang, C.T., Rocchino, G.H., Gutekunst, M.H., Paulvin, C., Melo Li, K. and Elam-Snowden, T., 2020. Perspectives of respect, teacher–student relationships, and school climate among boys of color: A multifocus group study. *Psychology of Men and Masculinities*, 21(3), p.345. <https://doi.org/10.1037/men0000239>.
- Lincoln, Y.S. and Guba, E.G., 1985. *Naturalistic inquiry*. Newbury Park, CA: Sage.
- Lincoln, Y.S., Lynham, S.A. & Guba, E.G., 2011. Paradigmatic controversies, contradictions, and emerging confluences, revisited. In N.K. Denzin & Y.S. Lincoln (Eds.). *The Sage handbook of qualitative research*. London: Sage.
- Maboe, K.A.M., 2019. Students' support in an ODeL Context: Students in ODeL. In LA. Darinskaia, and GI. Molodtsova, (Eds.), *Modern Technologies for Teaching and Learning in Socio-Humanitarian Disciplines*. IGI global: Hershey USA.
- Msila, V., 2008. Ubuntu and school leadership. *Journal of Education*, 44, pp.67–84.
- Papageorgiou, E., 2023. The inclusion of critical thinking in an accounting curriculum: Students' perceptions. *Journal of Education (University of KwaZulu-Natal)*, (91), pp.132-154. doi: <https://doi.org/10.17159/2520-9868/i91a08>
- Paul, R. and Elder, L., 2005. *Critical thinking competency standards: Standards, principles, performance indicators, and outcomes*. The Foundation for Critical Thinking. Available at <[http://www.criticalthinking.org/resources/PDF/CT-competencies 2005.pdf](http://www.criticalthinking.org/resources/PDF/CT-competencies%2005.pdf)> [Accessed 14 November 2024].

- Piaget, J., 1972. Development and learning. In CS. Lavattelly and F. Stendler (Eds.), *Readings on child behavior and development* (pp. 38–46). Hartcourt Brace Janovich, Inc.
- Powell, L. and McGuigan, N., 2020. Teaching, virtually: a critical reflection. *Accounting Research Journal*. doi: 10.1108/ARJ-09-2020-0307.
- Ricketts, J.C. and Rudd, R., 2004. The relationship between critical thinking dispositions and critical thinking skills of selected youth leaders in the national FFA Organisation. *Journal of Southern Agricultural education research*, 54(1).
- Rosli, K., Khairudin, N., and Saat, R. M. 2019. Gamification in entrepreneurship and accounting education. *Academy of Entrepreneurship Journal*, 25(3):1–6.
- Rossouw, M. and Steenkamp, G., 2025. Developing the critical thinking skills of first year accounting students with an active learning intervention. *The International Journal of Management Education*, 23(1), p.101086. doi: <https://doi.org/10.1016/j.ijme.2024.101086>
- Saldaña, J., 2016. *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage.
- Samkin, G. and Keevy, M. 2019. Using a stakeholder developed case study to develop soft skills. *Meditari Accountancy Research*, 27(6):862–882. doi: 10.1108/MEDAR-01-2018-0260
- South African Institute of Chartered Accountants (SAICA)., 2017. *Competency framework: Detailed guidance for the academic programme*. Available at <<https://www.saica.co.za/Portals/0/LearnersStudents/documents/Competency%20Framework%202019.pdf>> [Accessed 12 December 2018].
- South African Institute of Chartered Accountants (SAICA). (2023), CA(SA) Competency framework “CA of the future”. Available at <<https://saicawebprstorage.blob.core.windows.net/uploads/resources/CA-of-the-Future-Competency-Framework-applicable-to-IAC-January-2025.pdf>> [Accessed 25 July 2025].
- Sternberg, R.J., 1986. *Critical thinking: Its nature, measurement and improvement*. National Institute of Research. Available at <<https://files.eric.ed.gov/fulltext/ED272882.pdf>> [Accessed 14 November 2024].
- Terblanche, E.A.J., 2019. Developing critical thinking in auditing students through technology-based educational interventions: A conceptual framework. Unpublished PhD dissertation. Pretoria: University of South Africa.
- Terblanche, E.A.J. and De Clercq, B., 2020. Factors to consider for effective critical thinking development in auditing students. *South African Journal of Accounting Research*, 4(2), pp.96–114. doi: 10.1080/10291954.2019.1669293.
- Terblanche, E.A.J. and De Clercq, B., 2021. A Critical thinking competency framework for accounting student. *Accounting Education*. doi: 10.1080/09639284.2021.1913614.
- Terblanche, E.A.J., Joynt, C., Stiglingh, M., Du Preez, H. and Diedericks, L., 2025. Fostering critical thinking in first-year accounting students: insights from a semester-long intervention. *Accounting Education*, pp.1-22.
- Terblanche, A, Van Rooyen, A.A, and Enwereji P.C., 2025. When teachers become learners: Challenges with the integration of critical thinking into accounting curricula. *Thinking Skills and Creativity*, 58, doi: 10.1016/j.tsc.2025.101878.
- Turner, M. and Tyler, M., 2023. Demonstrating critical thinking in accounting: applying a competency framework. *Accounting Education*, 32(6), pp.713-734. doi: 10.1080/09639284.2022.2105653
- Ukhun, C.E., 2009. Critical thinking, community elders and Africa’s development. *Journal of Social Development in Africa*, 4(1), pp.119–139. doi: 10.4314/jsda.v24i1.54267.
- Venter, E., 2004. The notion of ubuntu and communalism in African educational discourse. *Studies in Philosophy and Education*, 23, pp.149–160. doi: 10.1023/B:SPED.0000024428.29295.03
- Vygotsky, L. S., 1978. *Mind in society*. Cambridge, MA: Harvard University Press.
- Wilkin, C.L., 2017. Enhancing critical thinking: accounting students’ perceptions. *Education + Training*, 59(1):15–30. doi: 10.1108/ET-01-2015-0007
- Wolcott, S.K. and Sargent, M.J., 2021. Critical thinking in accounting education: Status and call to action. *Journal of Accounting Education*, 56, pp.1–19. doi: <https://doi.org/10.1016/j.jaccedu.2021.100731>.
- Yang, Y.C. and Chou, H., 2008. Beyond critical thinking skills: Investigating the relationship between critical thinking skills and dispositions through different online instructional strategies. *British Journal of Educational Technology*. 39(4), pp.666–684. doi: 10.1111/j.1467-8535.2007.00767.x.

University Students' Use of Generative AI for English Language Learning: AI Readiness, Challenges, and Perceived Impact on Language Skills

Rastislav Metruk^{1,2}

¹Faculty of Education, Trnava University, Slovakia

²Faculty of Education, University of Ostrava, Czech Republic

rastislav.metruk@truni.sk

<https://doi.org/10.34190/ejel.24.4.4931>

An open access article under [CC Attribution 4.0](#)

Abstract: Generative AI (artificial intelligence) tools are now widely used by university students for language learning, yet empirical evidence from Central and Eastern Europe remains scarce, and a number of studies rely on closed-ended instruments that capture attitude strength but not the reasoning behind it. This mixed-methods study examines Slovak university students' use of generative AI for English language learning through three research questions addressing AI readiness, perceived challenges, and the impact on language skills development. A purpose-designed 62-item semi-structured questionnaire was administered to 153 students. The instrument combined 59 closed-ended items with three open-ended questions, creating a within-instrument triangulation design in which qualitative responses corroborated, contextualised, and extended the quantitative findings. An interesting tension emerged from the data: students reported high AI readiness and confidence in responsible use, yet expressed genuine concern about academic integrity and the consequences of AI use. Content reliability was the most prominent challenge, with fabricated references receiving the highest agreement score among the challenge items. The open-ended data reinforced this pattern, with content inaccuracy and over-reliance emerging as the most frequently cited problems. A clear split appeared between text-based and oral skills: students perceived strong AI benefits for vocabulary, grammar, reading, and writing, but reported limited or no perceived benefit for speaking, listening, and pronunciation, with listening falling in the Low range. The qualitative data revealed patterns the Likert scale could not detect: a self-reinforcing cycle of over-reliance and diminished critical thinking, and clear divisions on whether institutions should adopt or limit AI. An institutional gap was also documented: students disagreed that their university had provided adequate AI training, and independently identified AI literacy courses as the most frequently mentioned intervention. This study advances e-learning research by providing one of the first empirical analyses in this specific geographical context. It highlights the importance of including open-ended questions in surveys to reveal contradictions that closed-ended questions might miss. Additionally, it offers concrete recommendations for practice: embedding AI literacy in language curricula, teaching source verification alongside prompting skills, and preserving the dedicated speaking and listening practice for which generative AI is not a substitute.

Keywords: Generative artificial intelligence, English language learning, AI readiness, Higher education, Language skills

1. Introduction

Since the public release of ChatGPT in November 2022, generative artificial intelligence (AI) has moved from novelty to near-ubiquity in higher education. Tools such as ChatGPT, Google Gemini, or Claude now offer students capabilities that range from automated writing feedback and grammar correction to interactive conversational practice (Lee et al., 2026), and empirical research has expanded rapidly in response – though unevenly, with some regions and skill areas far better covered than others (Li, B. et al., 2025).

This expansion of research has also exposed genuine problems. Academic integrity concerns have multiplied as students negotiate unclear boundaries around acceptable AI use (Michel-Villarreal et al., 2023). The accuracy of AI-generated content remains unreliable, with fabricated references and plausible-sounding misinformation a known risk (Chan and Hu, 2023). Institutions face pressure to reform curricula in ways that foster digital literacy and responsible AI engagement (Walczak and Cellary, 2023), and in language education specifically, questions about learner dependency and the potential erosion of language skills have grown more urgent (Xiao and Zhi, 2023).

Two gaps in the existing literature motivate the present study. First, most empirical work has been conducted in East and Southeast Asian educational contexts or in well-resourced Anglophone institutions (Li, B. et al., 2025). Central and Eastern Europe is largely absent from this literature. Pikhart, Klímová and Al-Obaydi (2024), one of the few Central European contributions, examined Czech, Taiwanese, and Iraqi students' preferences for digital language learning tools, including ChatGPT, but their scope covered digital resources broadly rather than

generative AI specifically. Second, although mixed-methods designs are increasingly common in this field (Li, B. et al., 2025), several large-scale surveys of learner perceptions have used closed-ended instruments that capture attitude strength but not the reasoning behind it (e.g. Vo and Nguyen, 2024; Maxwell et al., 2025; Huang and Derakhshan, 2025). Likert scales measure the strength of attitudes but cannot capture the reasoning, contradictions, and contextual detail that open-ended questions elicit.

Against this background, the novelty of the present study lies in the combination of three features that, to our knowledge, prior work has not brought together. First, large-scale studies of perceptions of generative AI in higher education have been conducted in Vietnam (Vo and Nguyen, 2024), the United States (Maxwell et al., 2025), Hong Kong (Chan and Hu, 2023), and internationally across 76 countries (Yusuf, Pervin and Román-González, 2024), yet Central and Eastern Europe remains sparsely represented; the present study reports from a Slovak institution. Second, while Vo and Nguyen (2024) assessed perceived usefulness across the four macro-skills, much of the skill-specific research has focused on individual areas, most often writing (Song and Song, 2023; Guo and Wang, 2024) and less often speaking (Fathi, Rahimi and Derakhshan, 2024; Tai and Chen, 2024). The present study extends this work by comparing perceived impact across seven language domains – vocabulary, grammar, pronunciation, speaking, listening, reading, and writing – within a single instrument. Third, integrating open-ended items with quantitative measures within the same survey instrument makes it possible to identify contradictions between what students report about their own practice and what they perceive among other students, patterns that may remain hidden in closed-ended measures alone.

The present study addresses both gaps by examining Slovak university students' use of generative AI tools for English language learning, guided by three research objectives and corresponding research questions, which are set out in Section 3.

2. Literature Review

2.1 AI Readiness and Technology Acceptance

AI readiness in education refers to students' preparedness and competence to use AI tools productively for learning. Early usability research established a baseline: Shaikh et al. (2023) conducted a small-scale usability evaluation with ten university students and found that participants generally perceived ChatGPT as useful and easy to use for conversation, writing, grammar, and vocabulary tasks, though they concluded that AI should complement rather than replace human instruction. This link between AI literacy and effective engagement has since been confirmed from multiple theoretical angles. Vo and Nguyen (2024) applied a technology acceptance model to 369 Vietnamese English-major students and found a telling split – students considered ChatGPT easy to use but held only neutral views on its educational usefulness, suggesting that ease of access does not automatically translate into perceived learning value. Maxwell et al. (2025) surveyed 482 US students and documented that perceived readiness, benefits, and challenges varied across demographic groups, complicating the notion that students form a homogeneous group of AI adopters. From a motivational perspective, Du and Alm (2024) showed, using 24 New Zealand postgraduates, that ChatGPT supported autonomy and competence through a self-determination theory lens, but the picture was less positive for relatedness. Some students felt that AI use reduced human interaction. Huang and Derakhshan (2025) reinforced this multidimensional view through structural equation modelling with 363 Chinese EFL learners, finding that both learning motivation and digital literacy predicted AI adoption, which suggests that readiness cannot be reduced to a single factor. Taken together, these studies converge on a multidimensional view in which technical familiarity, motivational disposition and critical judgement operate jointly, and in which no single dimension predicts productive use on its own.

Prompting competence, which can be described as the ability to craft queries that produce useful AI responses, has emerged as a specific dimension of readiness. Tour and Zadorozhnyy (2025) proposed a framework grounded in the Four Resources Model, arguing that prompt literacy involves code-breaking, meaning-making, text-using, and text-analysing practices. In practice, learners need to understand how AI interprets input, evaluate its outputs, and refine their prompts iteratively – skills that overlap with broader language and digital literacies. Xiao and Zhi (2023) provided empirical support at a Chinese university, where EFL students demonstrated critical judgement and iterative prompt modification during AI-assisted tasks. Pikhart, Klímová and Al-Obaydi (2024), in one of the few Central European studies, explored Czech, Taiwanese, and Iraqi students' preferences for digital language learning tools, including ChatGPT, but their scope covered digital resources broadly rather than AI readiness as a distinct construct. Where universities do not offer structured guidance, students develop AI competence informally, risking uneven skill levels and inconsistent ethical practices.

2.2 Challenges of Using Generative AI in Language Learning

The challenges of generative AI in education cluster around several recurring themes, though the research increasingly shows that these are context-dependent rather than universal. Academic integrity, assessment design, and the need for new pedagogical approaches have been identified as the principal institutional-level concerns (Michel-Villarreal et al., 2023). Walczak and Cellary (2023) have argued that addressing them requires genuine curricular reform: not merely adding AI policies to existing syllabi, but rethinking how digital literacy is taught alongside disciplinary knowledge. The scale of the integrity concern appears to vary with context. Yusuf, Pervin and Román-González (2024), surveying 1,217 participants across 76 countries, found that perceptions of AI as an integrity threat depended heavily on cultural norms, institutional expectations, and whether ethical guidelines were available at all. Chan and Hu (2023) documented a related tension at the student level: learners simultaneously appreciated AI's utility and worried about the ethics of using it – an ambivalence that surfaces repeatedly across studies from different regions and educational systems.

Content reliability poses a distinct kind of challenge, particularly for language learners. Generative AI produces fluent, confident text that may be factually wrong, including fabricated references, a phenomenon commonly termed hallucination (Goh and Aryadoust, 2025). For students working in a foreign language, the problem is compounded: they may lack the domain expertise or linguistic confidence to recognise when an AI output is plausible yet incorrect.

Beyond accuracy, affective and social dimensions have begun to receive attention. Huang and Zou (2024) found that AI chatbots could enhance enjoyment and willingness to communicate, but outcomes depended on the specific tool and task. Rezai, Soyoo and Reynolds (2024) showed that emotion regulation mediated the relationship between ChatGPT use and well-being among EFL learners, and Wang, Y. (2025) noted that communication benefits depended on pedagogical integration and students' awareness of AI's limitations.

Two concerns that have received less empirical attention but may prove the most consequential are dependency and skill atrophy. Lee et al. (2026), in their systematic review, flagged both as under-researched, warning that students are increasingly delegating to AI the kinds of tasks, such as drafting arguments, evaluating sources, and structuring ideas, that traditionally built independent cognitive skills. Derakhshan and Taghizadeh (2026) explored this through a phenomenological study of L2 learners' lived experiences and concluded that AI acts as a double-edged sword for higher-order thinking: it can scaffold analytical reasoning by providing diverse perspectives and models; however, it can also undermine that reasoning if learners passively accept outputs and cease engaging with the intellectual work themselves. The challenge, in other words, is not just whether AI content is accurate, but whether the process of using AI leaves students more or less capable than they were before.

2.3 Impact of Generative AI on Language Skills Development

Research regarding AI's influence on particular language skills has progressed inconsistently, with writing receiving the most attention. Song and Song (2023) found significant gains in organisation, coherence, grammar, and vocabulary among Chinese EFL students receiving AI-assisted instruction, though students voiced concerns about dependency. Baskara (2023) highlighted ChatGPT's potential for personalised writing feedback while flagging unresolved ethical challenges. Guo and Wang (2024) compared ChatGPT's feedback with instructor feedback on fifty argumentative essays and found that ChatGPT produced significantly more feedback overall and distributed it relatively equally across content, organisation, and language, whereas teachers concentrated on content and language issues. The two sources also differed in feedback type, suggesting that AI and human feedback may complement rather than replace each other.

The literature on speaking presents a more positive perspective, but with a caveat. Experimental studies have reported improved speaking and willingness to communicate (Fathi, Rahimi and Derakhshan, 2024), reduced anxiety (Zhang, Meng and Ma, 2024), and differential effects by chatbot type (Wang, C. et al., 2024). Tai and Chen (2024) tested individual versus paired interaction modes, and Wang, Li and Zou (2025) examined AI speaking practice through positive psychology. Lyu, Lai and Guo's (2025) meta-analysis of 31 studies confirmed a medium effect ($g = 0.608$). The caveat is that virtually all this evidence comes from structured, researcher-designed interventions. Whether the benefits transfer to informal, self-directed AI use – how most students actually interact with generative AI – has not been tested.

Receptive skills represent the largest gap in the evidence base. Vocabulary and grammar have been documented as areas where AI can help (Song and Song, 2023; Shaikh et al., 2023), but evidence on reading comprehension, listening, and pronunciation is scarce. Çelik, Yangın Ersanlı and Arslanbay (2024) offer one of the few data points:

Turkish EFL undergraduates who read blog texts simplified by ChatGPT scored higher on comprehension tasks, though reading anxiety was unaffected. As Li, B. et al. (2025) observe in their review of 144 empirical studies, writing dominates the research agenda (42.4% of studies), while speaking, listening and reading have received comparatively less attention – an imbalance that may reflect real limitations in what text-based AI tools can do for non-text skills.

Read together, the evidence points to an asymmetry across skill areas: writing is the most extensively studied domain, while speaking, listening, and reading have received comparatively less attention; moreover, the positive evidence for oral skills in the studies reviewed here comes largely from structured chatbot-based interventions rather than informal, self-directed use. The present study contributes to this area by assessing perceived AI impact across seven language domains, while the limited evidence from Central Europe leaves open the question of how far findings generated largely in other educational contexts transfer to this setting.

3. Research Objectives and Research Questions

The present study examines university students' use of generative AI tools for English language learning and academic work in the Slovak higher education context. Three research objectives (ROs) and corresponding research questions (RQs) guide the investigation:

RO1: *To examine the AI readiness of Slovak university students for using generative AI tools in English language learning and academic work.*

RQ1: *How do students assess their own readiness to use generative AI tools for academic purposes, and how does that self-assessment relate to the preparation they report receiving from their institution?*

RO2: *To identify the challenges that students experience when using generative AI tools in their academic work.*

RQ2: *Which challenges do students identify as most significant, and how do the challenges they report in their own practice compare with those they perceive among other students?*

RO3: *To investigate the perceived impact of generative AI tools on the development of English language skills.*

RQ3: *How does perceived impact differ across text-based and oral/aural language skills, and what might explain that difference?*

4. Methodology

4.1 Research Design

This study adopted a mixed-methods design, implemented as a cross-sectional survey in which quantitative and qualitative data were collected concurrently within a single instrument. Descriptive statistics were computed for all Likert-scale items, with results reported by thematic category. Three open-ended questions were included to allow students to elaborate on the challenges they experience when using generative AI tools, and these responses were analysed using thematic content analysis. The strands were given complementary rather than equal weight: the quantitative strand establishes the distribution of attitudes across the sample, while the qualitative strand supplies the reasoning behind them. Following the convergent logic of within-instrument triangulation, the two were analysed separately and then compared, so that qualitative findings could corroborate, contextualise, or challenge the quantitative results.

4.2 Research Sample

The sample consisted of 153 university students specialising in English language studies at a Slovak higher education institution, selected through convenience sampling. The mean age of participants was 21.3 years, and the sample was predominantly female (82.4%), with 15.0% male and 2.6% identifying as non-binary or preferring not to disclose. The majority self-assessed their English proficiency at B2 level (54.2%), followed by C1 (30.7%), B1 (10.5%), and C2 (4.6%). Participants were drawn from all years of study, with a concentration in the earlier undergraduate years: 69 students (45.1%) were in the first year of Bachelor's study, 58 (37.9%) in the second, and 8 (5.2%) in the third, while 13 (8.5%) were in the first and 5 (3.3%) in the second year of Master's study. Bachelor's students therefore accounted for 88.2% of the sample and Master's students for 11.8%. Although the study was conducted at a Slovak institution, not all participants were Slovak nationals: 82.4% reported Slovak nationality, 11.8% Ukrainian, and 5.9% another nationality.

Data were collected online using Microsoft Forms. The questionnaire link was distributed through institutional communication channels. Participation was voluntary, and the form was configured to record responses

anonymously, so no names or email addresses were collected. Respondents gave informed consent on the opening page and could withdraw at any point by closing the form without submitting it.

Two features of the sample should be acknowledged as limitations at the outset. Convenience sampling from a single institution, combined with a predominantly female sample of English-major students, means that the findings describe this population rather than Slovak university students in general; the gender imbalance reflects the composition of English language programmes at the Slovak institution rather than a sampling decision.

4.3 Research Instrument

Data were collected using a purpose-designed, self-administered questionnaire comprising 62 items distributed across five sections. Section 1 (18 items) gathered demographic and background information. Sections 2–4 employed a five-point Likert scale (1 = Strongly disagree, 5 = Strongly agree): Section 2 (7 items) measured AI readiness and prompting competence, Section 3 (25 items) covered experiences with AI use across five thematic areas (academic integrity, dependency and skill atrophy, content reliability, social and psychological impact, and access and privacy), and Section 4 (9 items) assessed the perceived impact of generative AI on English language skills development. Section 5 comprised three open-ended questions inviting students to describe challenges and suggest interventions.

The questionnaire was developed specifically for this study, as no existing validated instrument comprehensively addressed all three research objectives. Item development was informed by the literature reviewed and guided by the three research objectives. The instrument was reviewed for face validity and content clarity prior to administration.

4.4 Reliability

Internal consistency was assessed using Cronbach's alpha. All three scales exceeded the threshold of $\alpha \geq .70$, considered acceptable for newly developed instruments (Nunnally, 1978; George and Mallery, 2003; Taber, 2018). Section 2 (AI Readiness) yielded $\alpha = .758$ (six items, after removing Q22 for low item-total correlation). Section 3 (Experiences with AI Use) yielded $\alpha = .741$ (22 items, after removing Q30, Q37, Q42). Section 4 (Language Skills) yielded $\alpha = .746$ (all nine items). The four removed items are reported separately as standalone items, as they remain substantively informative.

4.5 Data Analysis

Descriptive statistics – means (M) and standard deviations (SD) – were computed for each Likert-scale item. Statistical computations, including Cronbach's alpha coefficients and descriptive statistics, were performed using Claude (Anthropic, 2025), a large language model with code execution capabilities, and independently verified by the author using JASP, an open-source statistics program (JASP Team, 2026). Mean scores were interpreted using the levels-of-agreement framework proposed by Pimentel (2010): 1.00–1.79 = very low, 1.80–2.59 = low, 2.60–3.39 = neutral, 3.40–4.19 = high, 4.20–5.00 = very high. For the three open-ended questions, responses were first filtered to exclude non-substantive entries (e.g., 'No', 'Nothing', 'I don't know'). Recurring patterns were identified across responses and grouped into thematic categories, such as content reliability, over-reliance, and AI literacy needs. Because individual responses frequently addressed multiple themes, each distinct theme within a response was coded separately, and category frequencies reflect the number of mentions rather than the number of respondents.

5. Results

This section presents the descriptive statistics for the Likert-scale items across the three main sections of the questionnaire. Reverse-coded items are marked with an asterisk (*). Four standalone items (Q22, Q30, Q37, Q42) that were removed from their respective scales due to low item-total correlations are reported separately in Table 4.

Almost all respondents reported using generative AI tools in their studies: 19.6% did so daily, 40.5% a few times per week, 27.5% a few times per month, and 10.5% a few times per semester, while three respondents (2.0%) reported using no generative AI tools at all. Regarding the specific tools, versions, and purposes of use, ChatGPT was by far the most commonly used generative AI tool (92.2%), followed by Gemini (53.6%), NotebookLM (18.3%), Perplexity (13.7%), Copilot (13.1%), DeepSeek (8.5%), Meta AI (5.9%), Grok (5.2%), and Claude (4.6%). The vast majority of students (75.2%) used free versions only, with a further 15.0% using mostly free versions with occasional paid access. The most common academic purposes for AI use were language polishing (69.3%), understanding difficult concepts (62.7%), exam preparation (62.1%), and text summarisation (54.9%).

5.1 AI Readiness and Prompting Competence

Table 1 presents the six items comprising the AI Readiness and Prompting Competence scale ($\alpha = .758$). The overall scale mean was 3.85 (SD = 0.54), placing it in the High range on Pimentel's (2010) framework. The highest-rated item was Q25 ("I understand the limitations and potential risks of using generative AI"; $M = 4.33$, Very high), while the lowest was the reverse-coded Q21 ($M = 3.50$), indicating that a proportion of students still experience confusion when using AI tools despite rating their readiness positively overall.

Table 1: AI Readiness and Prompting Competence (N = 153, $\alpha = .758$)

Item	Statement	M	SD	Level
Q25	I understand the limitations and potential risks of using generative AI.	4.33	0.68	Very high
Q24	I feel confident in my ability to use AI tools responsibly.	4.13	0.72	High
Q23	I feel confident in my ability to use AI tools critically.	3.86	0.81	High
Q19	I feel adequately prepared to use generative AI tools for academic purposes.	3.66	0.83	High
Q20	I know how to write effective prompts to get useful and relevant responses from AI.	3.65	0.81	High
Q21*	I often feel lost or confused when trying to use AI tools effectively. (reverse-coded)	3.50	0.92	High

5.2 Experiences with AI Use

Table 2 presents the 22 items comprising the Experiences with AI Use scale ($\alpha = .741$). The overall scale mean was 3.06 (SD = 0.42), placing the average response at the Neutral level. Items are organised into five thematic groups for clarity, though these groups are not treated as independent sub-scales. The highest individual item mean was Q40 ("AI tools sometimes fabricate references"; $M = 4.18$, High), and students also agreed that AI output does not always meet academic standards (Q39; $M = 3.75$) and that paid AI tools create inequity of access (Q50; $M = 3.76$).

At the other end of the scale, students rejected the idea that they accept AI-generated answers without verification (Q35; $M = 2.17$, Low) – the lowest item mean across the entire questionnaire. They also disagreed that AI use had harmed their relationships with classmates or teachers (Q48*; $M = 2.32$) or caused feelings of isolation (Q46; $M = 2.44$). One finding worth noting is Q45 ($M = 3.14$): students neither agreed nor disagreed that they consult teachers less frequently because they can ask AI instead, suggesting an emerging but not yet decisive shift in help-seeking behaviour from human to AI sources.

Table 2: Experiences with AI Use (N = 153, $\alpha = .741$)

Item	Statement	M	SD	Level
Academic Integrity and Ethical Concerns				
Q31	I worry about the consequences if my teacher discovers that I have used AI in my work.	3.44	1.03	High
Q27	I worry about presenting work that was created with AI assistance as my own.	3.35	1.13	Neutral
Q28	I am uncertain about how to properly acknowledge or cite AI-generated content.	3.27	1.05	Neutral
Q29	I feel guilty when I use AI to complete parts of my assignments.	3.20	1.11	Neutral
Q26	I find it difficult to determine how much AI assistance is acceptable in my academic work.	2.80	1.04	Neutral
Dependency and Skill Atrophy				
Q34	Using AI makes me less likely to think critically about texts and tasks.	3.07	1.07	Neutral
Q36	I feel that AI does the intellectual work that I should be doing myself.	3.07	1.09	Neutral

Item	Statement	M	SD	Level
Q32	I am concerned that using AI reduces my motivation to develop my own English language skills.	3.03	1.24	Neutral
Q33	I notice that my own writing ability has declined since I started using AI regularly.	2.65	1.18	Neutral
Q35	I tend to accept AI-generated answers without verifying or questioning them.	2.17	1.01	Low
Content Reliability				
Q40	AI tools sometimes fabricate references or sources that do not actually exist.	4.18	0.90	High
Q39	AI-generated content sometimes does not meet the standards or expectations of academic writing in my field.	3.75	0.98	High
Q38	I find it challenging to deal with inaccurate or misleading information generated by AI tools.	3.01	0.98	Neutral
Q41	I struggle to identify errors or biases in AI-generated content.	2.81	0.99	Neutral
Social and Psychological Impact				
Q45	I consult my teachers less frequently because I can ask AI instead.	3.14	1.12	Neutral
Q43	I experience technical problems when using AI tools.	3.03	1.03	Neutral
Q47	I experience stress or anxiety about keeping up with rapid changes in AI technology.	2.65	1.19	Neutral
Q44	Using AI reduces my interaction with classmates on academic tasks.	2.48	1.03	Low
Q46	I feel more isolated in my studies since I started using AI tools regularly.	2.44	0.85	Low
Q48*	Using AI has not negatively affected my relationships with classmates or teachers. (reverse-coded)	2.32	1.11	Low
Access and Privacy				
Q50	The cost of paid versions of AI tools makes it harder for some students to benefit from AI.	3.76	1.02	High
Q49	I am concerned about the privacy and security of my personal data when using AI tools.	3.63	1.08	High

5.3 AI and English Language Skills Development

Table 3 presents the nine items on perceived AI impact on language skills ($\alpha = .746$). The overall scale mean was 3.17 (SD = 0.59), Neutral on Pimentel's framework. The data split into two distinct clusters, visible in Figure 1. Text-based skills scored in the High range: vocabulary (Q52; M = 3.71), reading (Q56; M = 3.59), grammar (Q51; M = 3.57), and writing (Q57; M = 3.48). Oral and aural skills scored lower: speaking (Q55; M = 2.74), pronunciation (Q53; M = 2.80), and listening (Q54; M = 2.57), with listening falling in the Low range. The 95% confidence intervals for these two clusters do not overlap (Figure 1), confirming that the distinction is not attributable to marginal score differences.

Two items assessed students' self-reflective awareness of their dependence on AI in language learning. Students expressed moderate concern that AI may be doing the language work for them rather than helping them learn (Q58; M = 3.01) and that their skills would improve with less AI reliance (Q59; M = 3.05). Both fell at the Neutral level – students neither dismissed nor strongly endorsed the risk.

Table 3: AI and English Language Skills Development (N = 153, $\alpha = .746$)

Item	Statement	M	SD	Level
Q52	Using AI helps me expand my English vocabulary.	3.71	0.97	High
Q56	Using AI helps me improve my English reading skills.	3.59	0.98	High

Item	Statement	M	SD	Level
Q51	Using AI helps me improve my English grammar.	3.57	0.99	High
Q57	Using AI helps me improve my English writing skills.	3.48	1.05	High
Q59	I feel that my English language skills would be better if I relied less on AI.	3.05	1.07	Neutral
Q58	I am concerned that AI is doing the language work for me rather than helping me learn.	3.01	0.99	Neutral
Q53	Using AI helps me improve my English pronunciation.	2.80	1.09	Neutral
Q55	Using AI helps me improve my English speaking skills.	2.74	1.13	Neutral
Q54	Using AI helps me improve my English listening skills.	2.57	1.02	Low

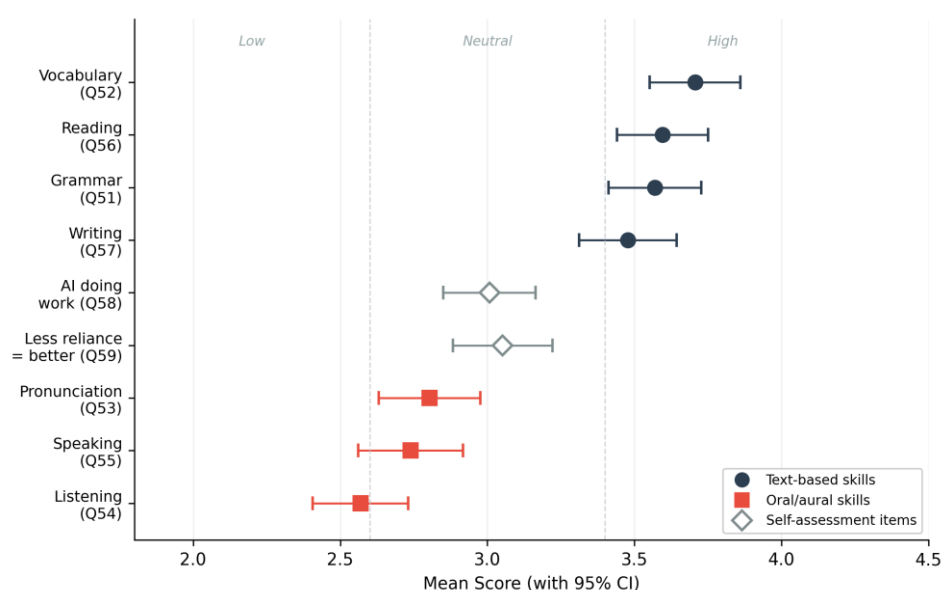


Figure 1: Mean scores with 95% confidence intervals for perceived AI impact on English language skills (N = 153). Dashed lines indicate Pimentel (2010) level boundaries. Text-based skills (circles) cluster in the High range, while oral/aural skills (squares) fall in the Low to Neutral range

5.4 Standalone Items

Table 4 presents the four items removed from their respective scales due to low item-total correlations, reported individually. The lowest score belongs to Q22 (M = 2.46, Low): students disagreed that their university had provided sufficient training on AI tools, a result that stands in tension with the high self-perceived readiness reported in Section 5.1. At the other end, Q42 (M = 3.97) indicates that most students report verifying AI-generated content before using it. The remaining two items fell in the Neutral range: Q30 (M = 3.24) on confidence in using AI without compromising integrity, and Q37 (M = 3.38) on whether AI helps students learn efficiently without eroding their own skills.

Table 4: Standalone Items (N = 153)

Item	Statement	M	SD	Level
Q42	I verify AI-generated content before using it in my work.	3.97	0.92	High
Q37	Using AI helps me learn more efficiently without reducing my own skills.	3.38	0.91	Neutral
Q30	I feel confident that I can use AI without compromising my academic integrity.	3.24	0.86	Neutral

Item	Statement	M	SD	Level
Q22	I have received sufficient training or guidance from my university on how to use AI tools.	2.46	1.00	Low

5.5 Qualitative Insights from Open-Ended Questions

Three open-ended questions were included to allow students to elaborate on their experiences. Of the 153 respondents, 44 (28.8%) responded to the first question, 73 (47.7%) to the second, and 58 (37.9%) to the third. After excluding non-substantive responses, the analysis was based on 21, 71, and 46 substantive responses, respectively. Responses were analysed using thematic content analysis (see Figure 2). Because individual responses frequently addressed multiple themes, frequencies reflect mentions rather than respondents.

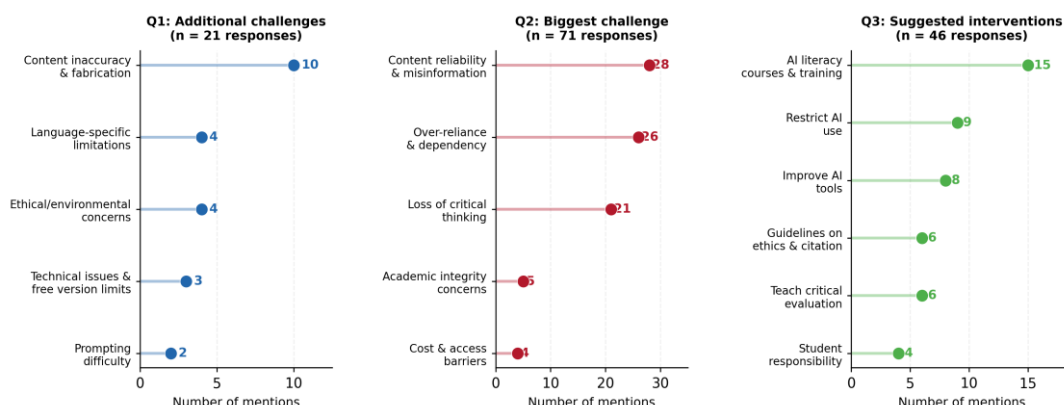


Figure 2: Thematic frequency of mentions across the three open-ended questions. Themes are ranked by frequency within each question. Because individual responses often addressed multiple themes, frequencies reflect the number of mentions, not the number of respondents

Question 1: “Are there any other challenges you have experienced when using AI tools for your studies that were not covered in the statements above?”

Among the 21 substantive responses, the most frequently mentioned additional challenge was content inaccuracy and fabrication ($n = 10$), with students describing fabricated sources and the need for repeated prompting. Language-specific limitations were raised by four respondents, who noted issues with translation quality, clichéd phrasing, and language mixing in AI output. Four respondents expressed ethical or environmental opposition to AI, and three mentioned technical issues and free version limitations. Two respondents described difficulty with prompting, noting that formulating effective queries requires considerable skill.

Question 2: “In your opinion, what is the biggest challenge that students face when using AI for learning and academic work?”

This question yielded the richest data (see Figure 2). Thematic analysis of the 71 substantive responses revealed three dominant, frequently overlapping themes. Content reliability and misinformation ($n = 28$) was the most frequently mentioned theme, followed closely by over-reliance and dependency ($n = 26$), with one respondent writing that students “stop relying on their knowledge and only consult everything with AI.” Loss of critical thinking ($n = 21$) was also prominent. Academic integrity ($n = 5$) and cost barriers ($n = 4$) were mentioned less frequently. The themes of over-reliance, critical thinking, and content reliability were deeply intertwined in many responses, suggesting that students perceive these not as separate issues but as interconnected dimensions of a broader problem.

Question 3: “Is there anything that your university, teachers, or AI tool developers could do to help you deal with these challenges?”

Of the 46 substantive responses, the most frequent suggestion was AI literacy courses and training ($n = 15$), covering prompt writing, AI limitations, and long-term learning strategies. A minority ($n = 9$) advocated restricting AI use rather than teaching it. Guidelines on ethical use and citation ($n = 6$), critical evaluation skills ($n = 6$), and improvements to AI tools ($n = 8$) were also mentioned. Four respondents placed responsibility on

the students themselves. The tension between those calling for integration and those advocating restriction reflects the broader ambivalence in the quantitative findings.

6. Discussion

The picture that emerges from these data is not a straightforward one. Students presented themselves as confident, critical AI users who verify outputs and understand risks, yet the same cohort described a peer culture of uncritical dependency and cognitive offloading. They rated content fabrication as a serious problem but reported only moderate concern about their own skill development. The quantitative data, taken alone, would paint a reassuringly balanced portrait; the qualitative data disrupt that portrait in ways that matter for how the findings should be interpreted and how institutional responses should be designed.

6.1 AI Readiness and the Institutional Gap (RQ1)

The high overall readiness score ($M = 3.85$) warrants closer examination rather than celebration. Students rated their understanding of AI limitations as the strongest dimension of their readiness (Q25; $M = 4.33$), and their prompting competence as adequate (Q20; $M = 3.65$), figures broadly in line with previous findings from Vietnam (Vo and Nguyen, 2024) and China (Xiao and Zhi, 2023). But there is reason to question how far self-report captures genuine competence. Du and Alm (2024) found that students often lacked the ability to effectively prompt AI or engage in post-editing of its output, even when they felt competent in using the tool, and the present data contain a similar tension: if students truly understand AI's limitations so well (Q25; $M = 4.33$), why do so many of their peers, as described in the open-ended responses, accept AI outputs uncritically? One plausible explanation is social desirability bias; another is the Dunning-Kruger adjacent possibility that students overestimate competencies they have never been formally taught to assess.

This last point connects to what is arguably the most consequential finding in the readiness data: the institutional gap. Students disagreed that their university had provided sufficient AI training (Q22; $M = 2.46$), placing this item firmly in the Low range. The implication is that whatever readiness these students possess, they acquired it on their own, through informal experimentation rather than structured guidance. This is not a neutral outcome. Informal learning is by definition uneven; students with greater digital capital, English proficiency, or curiosity will develop stronger AI literacy than those without, reproducing existing inequalities under the guise of individual initiative. Walczak and Cellary (2023) warned that institutions needed to respond with curricular reform; nearly three years later, the present data suggest that response has not materialised, at least not in this context. The qualitative data sharpen this interpretation. AI literacy courses were the single most frequently requested intervention ($n = 15$), covering prompt writing, AI limitations, and long-term learning strategies, and students separately called for guidelines on ethical use and citation ($n = 6$) and training in critical evaluation of AI output ($n = 6$). These are not vague complaints – they amount to a concrete curriculum that the institution has, to date, failed to provide.

6.2 Challenges and the Paradox of Critical Awareness (RQ2)

The overall challenge mean of 3.06 (Neutral) is, on its own, misleading. It averages together a High-level concern about fabricated references (Q40; $M = 4.18$) with Low-level disagreement about social isolation (Q46; $M = 2.44$), flattening meaningful variation into an uninformative midpoint. The item-level data tell a more interesting story. Content reliability was the clear standout: students recognised that AI fabricates sources (Q40; $M = 4.18$) and that its output often falls short of academic standards (Q39; $M = 3.75$). This is consistent with the phenomenon commonly termed hallucination (Goh and Aryadoust, 2025), but the present data add a practical dimension. Students are not just aware of hallucination in the abstract; when asked about challenges they had personally experienced, content inaccuracy and fabrication were the most frequently raised themes ($n = 10$), with students describing fabricated sources and the need for repeated prompting.

A genuine paradox lies at the heart of these data and deserves closer scrutiny. On the Likert scale, students reported diligent verification practices: they rejected uncritical acceptance of AI output (Q35; $M = 2.17$, the lowest item mean in the entire questionnaire) and rated their own verification behaviour highly (Q42; $M = 3.97$). Yet in an open-ended question about the biggest challenge their peers face, two dominant themes were over-reliance ($n = 26$) and loss of critical thinking ($n = 21$). This is not a minor discrepancy.

One reading is straightforward social desirability: students know they should verify, so they report doing so, while freely observing that others do not. But a less comfortable interpretation is also possible: that the gap between stated and actual verification behaviour may be smaller between the respondent and the described peers than the Likert data suggest. This cannot be resolved with the present data, but the pattern should give

pause to any study that relies on self-reported critical engagement as evidence that students are using AI responsibly.

The academic integrity items reveal genuine ethical uncertainty. Students worried about teachers discovering their AI use (Q31; $M = 3.44$) and felt uneasy about presenting AI-assisted work as their own (Q27; $M = 3.35$), but their concern was rooted not in a clear sense that AI use is wrong, but in uncertainty about where the line is. This mirrors Yusuf, Pervin and Román-González's (2024) cross-cultural finding that integrity perceptions vary with the availability of ethical guidelines – where guidelines are absent, anxiety fills the vacuum. The dependency data offer the most consequential example of why mixed-methods designs matter in this field. The quantitative aspects of dependency mostly received neutral ratings, making them easy to overlook. But the open-ended responses painted a different picture: over-reliance and critical thinking loss were two prominent and closely related themes, described not as separate problems but as a single reinforcing cycle. Lee et al. (2026) flagged dependency as under-researched; Derakhshan and Taghizadeh (2026) characterised AI as a double-edged sword for higher-order thinking. The qualitative data from the present study suggest that students already perceive the sword cutting, even if the Likert scale does not register the wound.

6.3 Perceived Impact on Language Skills (RQ3)

The text-based versus oral/aural split is the cleanest pattern in the dataset and, it is argued here, the most practically important. Vocabulary ($M = 3.71$), reading ($M = 3.59$), grammar ($M = 3.57$), and writing ($M = 3.48$) all fell in the High range; speaking ($M = 2.74$), listening ($M = 2.57$), and pronunciation ($M = 2.80$) did not exceed the Neutral threshold, with listening falling into the Low range. The non-overlapping 95% confidence intervals (Figure 1) confirm this is not a product of marginal score differences. The explanation is partly technological: current generative AI tools are text-processing systems at their core. They excel at manipulating written language because that is what they were built to do. Pronunciation, listening comprehension, and spontaneous speech production involve embodied, temporal, and auditory dimensions that a text interface cannot replicate. The pattern mirrors what Song and Song (2023), Shaikh et al. (2023), and Baskara (2023) reported for text-based skills, and the qualitative responses added a corrective: students noted that AI output tends toward clichéd phrasing and occasionally mixes languages – limitations invisible to the Likert items.

The speaking result deserves separate attention because it sits in tension with a body of experimental research reporting positive effects. Fathi, Rahimi and Derakhshan (2024) found improvements in speaking and willingness to communicate, Zhang, Meng and Ma (2024) reported reduced speaking anxiety, and Lyu, Lai and Guo's (2025) meta-analysis confirmed a medium effect across 31 studies ($g = 0.608$). But those studies employed structured, researcher-designed interventions, dedicated speaking practice sessions with AI chatbots under controlled conditions (Tai and Chen, 2024; Wang, C. et al., 2024). The students in the present study, by contrast, reported predominantly text-oriented uses of AI, including language polishing and text summarisation. This pattern may help explain why perceived benefits were stronger for text-based than for oral/aural skills. The discrepancy, then, is not between the present findings and the experimental literature; it is between the affordances of structured pedagogical AI use and the reality of how students actually use these tools when left to their own devices. Çelik, Yangın Ersanlı and Arslanbay (2024) showed that AI-simplified texts can improve reading comprehension, suggesting that even receptive skills may benefit when AI is used thoughtfully for input modification. The broader point is that generative AI's language learning potential is not an intrinsic property of the tool; it is a function of how the tool is used, and informal use skews heavily toward text.

6.4 The Methodological Value of Integrating Quantitative and Qualitative Data

A recurring argument throughout this discussion has been that the open-ended questions did genuine analytical work and were not merely decorative. It is worth making the case explicitly. The qualitative data operated on three levels. First, corroboration: the prominence of content reliability in the open-ended responses ($n = 28$ mentions) independently confirmed the highest-rated challenge item (Q40; $M = 4.18$), and the salience of over-reliance ($n = 26$) tracked the moderate quantitative concern about dependency. When two data sources converge, each strengthens the other's credibility. Second, contextualisation: the qualitative data showed what the numbers describe but cannot explain – what it is like to encounter a fabricated reference mid-assignment, how students develop ad hoc verification routines, and how skill atrophy registers in a learner's own perception of declining fluency. Third, and most important, extension: the open-ended responses revealed patterns that the Likert scale was structurally unable to capture. Students described over-reliance, critical thinking loss, and misinformation not as separate challenges but as a single reinforcing cycle. They split into opposing camps on whether institutions should integrate AI more deeply or restrict it, a division no closed-ended item had

anticipated. And the verification paradox discussed above would have been entirely invisible without the qualitative layer.

The methodological implication is direct. Survey-based research relying exclusively on Likert scales risks producing a sanitised picture that might not reflect how students experience these tools. Self-report scales are vulnerable to social desirability, acquiescence bias, and the averaging-out of complex attitudes into uninformative midpoints. The within-instrument triangulation employed here exposed contradictions that a purely quantitative instrument would have flattened, and researchers designing future surveys should consider qualitative components not as optional addenda but as core elements of instrument validity. Finally, this study is one of the first to report from Central Europe. The Slovak students exhibited patterns that share features with other contexts, including strong perceptions of text-based skills and concerns about content reliability, while diverging in others. The institutional training gap documented here may reflect a broader pattern in contexts where AI-related curricular reform is still at an early stage. It is worth distinguishing which findings are likely to travel and which are tied to this setting. The text-oral asymmetry should be the most transferable, since it follows from the architecture of current generative AI tools rather than from any local condition. The verification paradox is likewise a property of self-report methodology, and would be expected wherever socially desirable responding is possible. The size of the institutional training gap, by contrast, is a local finding, as is the prominence of cost as a barrier, which is sensitive to national income levels. Whether that structural factor accounts for the gap, or whether it is equally prevalent elsewhere but less frequently measured, remains a question for comparative research.

7. Conclusion

This study set out to examine how Slovak university students use generative AI for English language learning, and the answer turns out to be more complicated than any single data source could reveal. Students are confident, largely self-taught AI users who recognise the technology's limitations in the abstract but navigate a daily reality of fabricated references, unclear ethical boundaries, and an institutional environment that is still adapting to the tools they are already using. The tension between high self-reported readiness and limited institutional training carries practical implications, because informal learning produces uneven competencies across the student body.

Three findings carry direct practical weight. First, AI literacy deserves a more prominent place in university curricula; the students themselves identified this as the most needed intervention. Second, the sharp text-oral skills divide means that generative AI cannot substitute for the interactive, embodied language practice that develops speaking, listening, and pronunciation; institutions should therefore preserve dedicated practice in these skills and be explicit with students about why it matters. Third, the qualitative-quantitative divergence regarding dependency and verification behaviour should concern anyone designing assessments or policies based on what students say they do with AI; what they report about their peers suggests a more troubled picture than self-report alone can capture.

The study has limitations. The sample from a single Slovak institution, with a predominance of female participants, limits generalisability; self-report data carry inherent biases, and the cross-sectional design cannot capture how attitudes evolve over time. Future research should track students longitudinally, compare self-reported verification with actual practice through observational methods, and test whether structured AI integration into curricula can close the readiness gap. Cross-national comparisons within Central and Eastern Europe would clarify whether these patterns are regional or local.

The question facing higher education is no longer whether students will use generative AI for language learning. They already do so overwhelmingly, largely without institutional guidance. The question is whether universities will shape that use or merely react to it. Every semester of inaction risks widening the gap between what students need and what institutions provide, potentially affecting students with fewer digital, linguistic, or financial resources most strongly. The data presented here, from a context rarely represented in this literature, suggest that students are not passive consumers of AI but active, ambivalent agents navigating a technology that is simultaneously powerful and unreliable. Understanding how students engage with these tools is a necessary first step toward informed policy and curriculum design.

Acknowledgement

This article is from Project no. B4/2026, "Using Generative Artificial Intelligence Tools in the Foreign Language Teaching Process – Challenges and Perspectives", funded by the Faculty of Education, Trnava University, Slovakia

and from the project "Implementation of the Reform of Teacher Education at the University of Ostrava: Competence-based learning founded on practice, research, internationalisation and cooperation with the clinical school CZ.02.02.XX/00/23_019/0008272", which is co-financed by the European Union.

Ethics Statement: Ethics approval was not required for this study. All participants gave their informed consent for inclusion before they participated in the study. Participation was entirely voluntary, and respondents could withdraw at any point before submitting without consequences. No personally identifiable information was collected, and all data were stored securely and used solely for research purposes. The study was conducted in accordance with the Declaration of Helsinki and the ethical standards of the institution.

AI Declaration: Generative AI tools were used in the preparation of this manuscript. Specifically, Claude Opus 4 (Anthropic, 2025) was used to perform statistical computations (Cronbach's alpha coefficients and descriptive statistics) and to assist with manuscript formatting, language editing, and the preparation of tables and figures. GPT-5.6 Sol (OpenAI, 2026) was also used during the revision process to cross-check statistical calculations, verify the accuracy of literature-based claims and references, and assist with language and internal consistency checks. All AI-assisted statistical outputs and figures were independently verified by the author against the underlying data, and the statistical analyses were independently reproduced in JASP (an open-source statistics program). The author takes full responsibility for the content and integrity of the final manuscript.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding: This article was funded by Project no. B4/2026, "Using Generative Artificial Intelligence Tools in the Foreign Language Teaching Process – Challenges and Perspectives", funded by the Faculty of Education, Trnava University, Slovakia.

References

- Anthropic, 2025. Claude Opus 4 [Large language model]. Available at: <https://www.anthropic.com/claude> [Accessed 27 May 2026].
- Baskara, F.X.R., 2023. Integrating ChatGPT into EFL writing instruction: benefits and challenges. *International Journal of Education and Learning*, 5(1), pp 44–55. <https://doi.org/10.31763/ijelev.v5i1.858>.
- Çelik, F., Yangın Ersanlı, C. and Arslanbay, G., 2024. Does AI simplification of authentic blog texts improve reading comprehension, inferencing, and anxiety? A one-shot intervention in Turkish EFL context. *The International Review of Research in Open and Distributed Learning*, 25(3), pp 287–303. <https://doi.org/10.19173/irrodl.v25i3.7779>.
- Chan, C.K.Y. and Hu, W., 2023. Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20, 43. <https://doi.org/10.1186/s41239-023-00411-8>.
- Derakhshan, A. and Taghizadeh, M.S., 2026. Does artificial intelligence (AI) nurture or hinder language learners' higher-order thinking skills (HOTS)? A phenomenological study on L2 learners' perspectives and lived experiences. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.12824>.
- Du, J. and Alm, A., 2024. The impact of ChatGPT on English for Academic Purposes (EAP) students' language learning experience: a self-determination theory perspective. *Education Sciences*, 14(7), 726. <https://doi.org/10.3390/educsci14070726>.
- Fathi, J., Rahimi, M. and Derakhshan, A., 2024. Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, 121, 103254. <https://doi.org/10.1016/j.system.2024.103254>.
- George, D. and Mallery, P., 2003. *SPSS for Windows step by step: a simple guide and reference, 11.0 update*. 4th ed. Boston: Allyn & Bacon.
- Goh, C.C.M. and Aryadoust, V., 2025. Developing and assessing second language listening and speaking: does AI make it better? *Annual Review of Applied Linguistics*, 45, pp 179–199. <https://doi.org/10.1017/S0267190525100111>.
- Guo, K. and Wang, D., 2024. To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, 29, pp 8435–8463. <https://doi.org/10.1007/s10639-023-12146-0>.
- Huang, F. and Derakhshan, A., 2025. Learning motivation and digital literacy in AI adoption for self-regulated English learning. *European Journal of Education*, 60(4), e70254. <https://doi.org/10.1111/ejed.70254>.
- Huang, F. and Zou, B., 2024. English speaking with artificial intelligence (AI): the roles of enjoyment, willingness to communicate with AI, and innovativeness. *Computers in Human Behavior*, 159, 108355. <https://doi.org/10.1016/j.chb.2024.108355>.
- JASP Team, 2026. JASP (Version 0.98.0) [Computer software]. <https://jasp-stats.org> [Accessed 15 August 2026].
- Lee, S., Choe, H., Zou, D. and Jeon, J., 2026. Generative AI (GenAI) in the language classroom: a systematic review. *Interactive Learning Environments*, 34(1), pp 335–359. <https://doi.org/10.1080/10494820.2025.2498537>.
- Li, B., Tan, Y.L., Wang, C. and Lowell, V.L., 2025. Two years of innovation: a systematic review of empirical generative AI research in language learning and teaching. *Computers and Education: Artificial Intelligence*, 9, 100445. <https://doi.org/10.1016/j.caeai.2025.100445>.

- Lyu, B., Lai, C. and Guo, J., 2025. Effectiveness of chatbots in improving language learning: a meta-analysis of comparative studies. *International Journal of Applied Linguistics*, 35(2), pp 834–851. <https://doi.org/10.1111/ijal.12668>.
- Maxwell, D., Oyarzun, B., Kim, S. and Bong, J.Y., 2025. Generative AI in higher education: demographic differences in student perceived readiness, benefits, and challenges. *TechTrends*, 69, pp 1248–1259. <https://doi.org/10.1007/s11528-025-01109-6>.
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D.E., Thierry-Aguilera, R. and Gerardou, F.S., 2023. Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, 13(9), 856. <https://doi.org/10.3390/educsci13090856>.
- Nunnally, J.C., 1978. *Psychometric theory*. 2nd ed. New York: McGraw-Hill.
- OpenAI, 2026. GPT-5.6 Sol [Large language model]. Available at: <https://developers.openai.com/api/docs/models/gpt-5.6-sol> [Accessed 14 August 2026].
- Pikhart, M., Klímová, B. and Al-Obaydi, L.H., 2024. Exploring university students' preferences and satisfaction in utilizing digital tools for foreign language learning. *Frontiers in Education*, 9, 1412377. <https://doi.org/10.3389/feduc.2024.1412377>.
- Pimentel, J.L., 2010. A note on the usage of Likert scaling for research data analysis. *USM R&D Journal*, 18(2), pp 109–112.
- Rezai, A., Soyoo, A. and Reynolds, B.L., 2024. Disclosing the correlation between using ChatGPT and well-being in EFL learners: considering the mediating role of emotion regulation. *European Journal of Education*, 59(4), e12752. <https://doi.org/10.1111/ejed.12752>.
- Shaikh, S., Yayilgan, S.Y., Klímová, B. and Pikhart, M., 2023. Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigation in Health, Psychology and Education*, 13(9), pp 1937–1960. <https://doi.org/10.3390/ejihpe13090140>.
- Song, C. and Song, Y., 2023. Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>.
- Taber, K.S., 2018. The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48, pp 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>.
- Tai, T.-Y. and Chen, H.H.-J., 2024. Improving elementary EFL speaking skills with generative AI chatbots: exploring individual and paired interactions. *Computers & Education*, 220, 105112. <https://doi.org/10.1016/j.compedu.2024.105112>.
- Tour, E. and Zadorozhnyy, A., 2025. Conceptualizing and operationalizing prompt literacy for English language learners. *Journal of Adolescent & Adult Literacy*, 69(3), e70020. <https://doi.org/10.1002/jaal.70020>.
- Vo, T.K.A. and Nguyen, H., 2024. Generative artificial intelligence and ChatGPT in language learning: EFL students' perceptions of technology acceptance. *Journal of University Teaching and Learning Practice*, 21(6). <https://doi.org/10.53761/fr1rkj58>.
- Walczak, K. and Cellary, W., 2023. Challenges for higher education in the era of widespread access to generative AI. *Economics and Business Review*, 9(2), pp 71–100. <https://doi.org/10.18559/ebr.2023.2.743>.
- Wang, C., Li, X. and Zou, B., 2025. Revisiting integrated model of technology acceptance among the generative AI-powered foreign language speaking practice: through the lens of positive psychology and intrinsic motivation. *European Journal of Education*, 60(1), e70054. <https://doi.org/10.1111/ejed.70054>.
- Wang, C., Zou, B., Du, Y. and Wang, Z., 2024. The impact of different conversational generative AI chatbots on EFL learners: an analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System*, 127, 103533. <https://doi.org/10.1016/j.system.2024.103533>.
- Wang, Y., 2025. A study on the efficacy of ChatGPT-4 in enhancing students' English communication skills. *SAGE Open*, 15(1), pp 1–17. <https://doi.org/10.1177/21582440241310644>.
- Xiao, Y. and Zhi, Y., 2023. An exploratory study of EFL learners' use of ChatGPT for language learning tasks: experience and perceptions. *Languages*, 8(3), 212. <https://doi.org/10.3390/languages8030212>.
- Yusuf, A., Pervin, N. and Román-González, M., 2024. Generative AI and the future of higher education: a threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education*, 21, 21. <https://doi.org/10.1186/s41239-024-00453-6>.
- Zhang, C., Meng, Y. and Ma, X., 2024. Artificial intelligence in EFL speaking: impact on enjoyment, anxiety, and willingness to communicate. *System*, 121, 103259. <https://doi.org/10.1016/j.system.2024.103259>.