Asian-Pacific Journal of Second and Foreign Language Education

RESEARCH Open Access

Exploring the impact of AI on language instructors' critical thinking abilities: Insights from higher education

Jaber Kamali[1*], Ramin Ali pour[2], Mina Naderi[3], Mahtab Afrasiabi[4] and Faramarz Piri[5]

*Correspondence:
Jaber.kamali@ihu.edu.tr

[1] School of Languages, Ibn Haldun University, Istanbul, Turkey
[2] Department of English Language, Hakim Sabzevari University, Sabzevar, Iran
[3] English Department, Shahrekord University, Shahrekord, Iran
[4] Department of English Language Teaching, University of Yazd, Yazd, Iran
[5] English Language Department, Kermanshah ACECR Institute of Higher Education, Kermanshah, Iran

Abstract

This study examines the impact of artificial intelligence (AI) on the critical thinking (CT) abilities of language instructors in higher education. To this end, 10 university language instructors participated in semi-structured interviews, which were analyzed thematically. Having adapted the CT framework, the data were discussed in four key domains: clarification, advanced clarification, basis of inference, and inference. Participants highlighted the benefits of AI for concept clarification and a subtle understanding of information. Furthermore, they pinpointed the potential of AI to facilitate advanced clarification and a deeper analysis of underlying assumptions. However, regarding the basis of inference, the reliance on AI is reduced, suggesting the need for a practical integration of AI in educational practices. In conclusion, this research highlights the complex perspectives of instructors and underscores the pivotal role of AI in CT in higher education contexts, while emphasizing the need for further research on its implications to inform AI-integrated pedagogical approaches.

Keywords: Artificial intelligence, Critical thinking, Higher education, Teachers' perspectives

Introduction

Artificial intelligence (AI), also referred to as machine intelligence, is a specialized field within computer science. It focuses on comprehending the nature of intelligence and creating innovative, intelligent machines that mimic, enhance, and augment human intelligence. The advancement of AI has led to numerous innovations and improvements that have influenced various aspects of human existence. Education, being a crucial element of societal progress and personal growth, has greatly benefited from these AI developments (Kamalov et al., 2023). In this study, we treat AI as an umbrella term that includes a range of applications (e.g., intelligent tutoring systems, feedback tools, recommendation and retrieval systems), while giving particular analytical attention to generative AI systems such as large language models (LLMs). Generative AI refers to probabilistic machine-learning models that produce novel text, images, or other media by learning patterns from large corpora (e.g., ChatGPT) (Moorhouse & Wong, 2025). When we refer specifically to this subset, we use the term "generative AI" or "LLMs";

Springer

when we use "AI" more broadly, this reflects the way participants themselves often spoke about "AI" alongside other AI-mediated and data-driven technologies.

The study of AI necessitates integrating knowledge from various disciplines, including computer science, mathematics, sociology, psychology, and philosophy (Huang & Qiao, 2022). The use of AI in education signifies not just a technological improvement but also a fundamental shift towards more personalized and adaptive learning experiences. However, despite its potential advantages, the incorporation of AI in educational settings presents notable challenges (Kamali et al., 2024a). A primary concern is the cultivation of Critical Thinking (CT) skills among instructors and students, especially in environments where traditional educational practices are still dominant (Liu & Wang, 2024b). Engaging with AI can significantly enhance instructors' and students' CT skills as they explore AI applications and assess their real-world implications and ethical considerations. (Huang & Qiao, 2022). In our rapidly changing, technology-driven society, CT is essential for instructors and students to evaluate the extensive information available, make informed choices, and tackle the intricate challenges posed by swift technological progress, including advancements in AI (Li et al., 2024; Liu & Wang, 2024a). According to Popenici and Kerr (2017), as advanced as AI solutions may become, a future where algorithms can replace the complexity of the human mind and CT cannot yet be envisaged.

To examine this, this research aims to explore how AI- and generative-AI–supported instructional environments influence the development of CT skills among higher education language instructors. It seeks to identify specific features of AI tools that enhance or hinder these skills. By investigating the impact of AI on CT skills, this research can provide valuable insights for instructors and higher education institutions on the integration of AI technologies into curricula. Furthermore, this study aims to evaluate how instructors' perceptions and experiences with AI technologies shape their integration into the classroom and their effectiveness in fostering CT. In effect, it seeks to understand the barriers and facilitators that teachers face regarding their use of CT when adopting AI tools.

Literature Review

AI has become a crucial component of higher education, influenced by advancements in information technology and intelligent machine capabilities (Kamali et al., 2024b; McGrath et al., 2024). This integration offers both opportunities and challenges for teaching and learning (Kuleto et al., 2021; Popenici & Kerr, 2017). Recent work frames a central tension: AI as a cognitive aid that scaffolds analysis versus an epistemic threat that can outsource judgment and induce automation bias (Chan & Tsi, 2024; Lu et al., 2024), especially in the language classroom (Seo, 2025; Tay & Xie, 2025). In particular, scholars warn of risks to academic integrity and the validity of assessment with generative AI, alongside difficulties detecting and deterring misconduct (Cotton et al., 2024), and meta-reviews catalog persistent concerns around bias, opacity, data privacy, uneven access, and staff capacity/infrastructure (Bond et al., 2024). In line with our definitional stance, this literature spans both earlier AI applications (e.g., learning analytics, intelligent tutoring systems, robotics) and more recent generative AI tools; we signal "generative AI" when referring specifically to text/image generators such as LLMs. Instructors

are increasingly utilizing AI tools to broaden students' perspectives, such as employing intelligent robotic devices for English learning (Zheng et al., 2018) and developing teachable agent-based games for preschool mathematics (Gulz et al., 2020). Additionally, chatbots and expert systems are seen as potential enhancements for student engagement (Guo et al., 2024). In higher-education writing, for instance, generative AI-generated feedback has been shown to complement rather than replace human feedback, with differential strengths across criteria and revision planning (Banihashem et al., 2024; Javahery et al., 2025), foregrounding orchestration rather than substitution.

Technological evolution, particularly in the field of AI, has transformed educational settings into intelligent learning environments that utilize neural networks and natural language processing (Lin et al., 2023). Research indicates that AI can support the development of computational thinking skills. For example, Lin et al. (2021) demonstrated that artificial intelligence of things (AIoT) learning, combined with Augmented Reality, can improve students' problem-solving abilities and comprehension. Understanding AI requires interdisciplinary knowledge from fields like computer science, mathematics, sociology, psychology, and philosophy, which can enhance students' CT through the examination of AI applications and their ethical implications (Huang & Qiao, 2022). However, these benefits are conditional on design choices and governance: AI infrastructures can enhance feedback only when aligned with transparent criteria, student agency, and teacher mediation (Banihashem et al., 2022). Teacher-education research similarly shows that generative AI can structure lesson planning and scaffold critical inquiry, while requiring ethical guardrails and literacies that keep human judgment central (van den Berg & du Plessis, 2023).

In the twenty-first century, CT has become vital, with AI-driven instruction showing promise in helping instructors and students develop these skills by fostering trust and self-confidence in learning (Muthmainnah et al., 2022). The traditional "one size fits all" approach to education is being challenged by the potential of AI for personalized learning (Hwang et al., 2020). Intelligent (AI-supported) instructors can provide customized feedback and instruction tailored to individual student needs (Guo et al., 2024). However, personalization can slide into epistemic outsourcing if AI outputs are treated as authoritative rather than fallible prompts. Studies therefore advocate for the deliberate pairing of AI feedback with instructor input and student self-evaluation (Banihashem et al., 2024; Lu et al., 2024). Design proposals for "hybrid intelligent feedback" specify roles for humans and GenAI to support higher-order inquiry and evidence appraisal, emphasizing process transparency and calibrational mentoring (Banihashem et al., 2025).

For the purpose of this study, building on Ennis's early and updated formulations, we adopt a four-domain view of CT: clarification, advanced clarification, basis of inference, and inference (Ennis, 1978, 2018). In this context, clarification involves identifying the issue, stating and interpreting claims, and mapping reasons; i.e., making the problem space and argument structure explicit. Advanced clarification extends this by defining key terms, surfacing presuppositions, handling ambiguity and vagueness, and testing the consistency of positions. The basis of inference involves the evaluation of reasons and evidence, including assessing the credibility of sources, the quality of data, warrants, and the fit between evidence and the claim. Finally, inference addresses drawing and

evaluating conclusions (judging deductive validity, estimating inductive strength, considering alternatives and counter-evidence, and calibrating confidence). Taken together, these domains provide a practical blueprint for aligning tasks, feedback, and assessment, including in AI-supported learning environments where evidence appraisal and conclusion formation can be scaffolded in domain-specific ways (Ennis, 1978, 2018). Empirical and conceptual work on AI-enabled feedback aligns with these domains by strengthening clarification (e.g., descriptive summaries) and the basis of inference (e.g., targeted identification/justification), while highlighting complementarity between human and AI feedback (Banihashem et al., 2024, 2025).

The promotion of CT is a significant goal in higher education (Manning, 2024; McPeck, 2016). Power dynamics play a crucial role in CT through the pedagogic device, which influences CT via distributive, recontextualizing, and evaluative rules that shape teaching and learning experiences (Bernstein, 2000). Distributive rules manage power relationships among social groups by allocating various forms of CT, thereby creating different professional identities (Singh, 2002). For instance, AI-driven personalized learning systems may assign distinct CT tasks based on an individual's learning style. Recontextualizing rules govern the formation of pedagogic discourses by relocating and refocusing them, reflecting the distribution of power and control over what constitutes acceptable CT (Bernstein, 2004). In datafied classrooms, AI-driven feedback pipelines can quietly reconfigure evaluative rules, what counts as evidence, and who judges it, shifting authority from teachers to dashboards unless design makes criteria visible and contestable (Banihashem et al., 2022). This helps explain why debates about AI and CT are simultaneously about automation versus teacher agency (Chan & Tsi, 2024) and about institutional governance of assessment (van den Berg & du Plessis, 2023).

AI-powered platforms can manipulate CT tasks to align with specific power distributions (Walter, 2024). Evaluative rules establish pedagogic practices that validate CT within instructional and regulative discourses, such as AI-driven assessment tools that evaluate students' CT skills. Furthermore, power and control are expressed through classification and framing, which are conceptually distinct but influence how CT is organized and transmitted in educational contexts (Bernstein, 2000). AI can affect the classification of knowledge areas and the framing of pedagogical relationships, impacting students' autonomy in their learning processes.

Therefore, power dynamics significantly shape the teaching and learning of CT in educational contexts. Fuchs (2017) emphasizes the need to examine power dynamics in AI development and implementation critically. Coeckelbergh (2020) and Kamali et al. (2024b) highlight the importance of AI ethics in shaping policy discussions, while also emphasizing the need for concrete policies to ensure that the benefits of AI are accessible to everyone. Taken together, the literature suggests a partnering model: leverage AI for ideation, organization, and formative cues while safeguarding human-led evaluation, ethics, and contextual judgment (Banihashem et al., 2025; Chan & Tsi, 2024; Lu et al., 2024).

The integration of AI in education also necessitates a change in the role of instructors. Moving away from being the "sage on the stage," instructors are now expected to be the "guide on the side" (King, 1993, p. 30), facilitating learners in seeking, finding, and integrating information, thereby fostering independent thinking. Learner-instructor

interaction plays a crucial role in language learning (Hagenauer et al., 2024); however, further research is needed to understand how AI systems impact this interaction (Felix, 2020). Concerns have also been raised about the changes in the role of teachers in the classroom and the transparency of decisions made by AI systems (Kim & Kim, 2022). Instructors have historically expressed concerns that AI could replace their jobs, rather than enhancing learning and instruction, influenced by media and science fiction (Luckin & Holmes, 2016). However, recent studies indicate a shift in their expectations regarding significant changes in education through the implementation of AI (Kim, 2024). Attitudinal modeling in post-pandemic blended contexts shows instructors and students negotiating evolving roles and emotions within AI ecosystems, highlighting PD needs in critical data literacy and feedback design (Banihashem et al., 2023).

Teachers' perceptions of AI vary based on factors such as pedagogical beliefs, teaching experience, prior exposure to educational technology, and views on the effectiveness of specific technologies (Alwaqdani, 2024). For example, many STEM teachers report positive experiences with AI, recognizing its potential for providing personalized support, immediate feedback, and customized learning materials (Kim et al., 2021). Despite these benefits, some teachers remain hesitant to adopt AI due to negative perceptions and anxiety about new technologies, preferring familiar methods (Hébert et al., 2021; Hopcan et al., 2024; Istenic et al., 2021). Therefore, addressing teachers' perceptions and offering professional development is essential to enhance their understanding and acceptance of AI in education (Kim & Kim, 2022).

The significance of the present study lies in addressing a rapidly evolving but still under-explored intersection between generative AI and CT in higher education, with a specific focus on language instructors rather than students. Recent systematic reviews indicate that AI can both scaffold and erode CT, depending on its integration into learning and assessment practices (Melisa et al., 2025; Salido et al., 2025). At the same time, work on teachers' and academic staff's views highlights strong interest in AI alongside concerns about over-reliance, ethics, and a lack of pedagogical guidance (Mah et al., 2025). By examining instructors' perspectives through Ennis's CT domains (clarification, advanced clarification, basis of inference, and inference), this study aims to focus on the following research questions:

> *How, from higher education instructors' points of view, do AI-driven instructional contexts impact higher education instructors' CT skills?*

Method

A qualitative multiple-case study design was employed to explore the perspectives of Iranian university instructors regarding AI's impact on CT in higher education. The study aimed to understand the relationship between AI technologies and instructors' attitudes and skills. Convenient sampling was used to recruit 10 Iranian language instructors from various universities. Data collection occurred during the fall semester to accommodate participants' teaching schedules. Semi-structured interviews were conducted, and a Google Form questionnaire was utilized for demographic data collection. The interviews were audio-recorded, transcribed, and translated from Persian to English for analysis.

Participants

Ten Iranian university instructors specializing in TEFL (Teaching English as a Foreign Language) from different universities participated in this study. Participants were recruited through purposive, criterion-based sampling, and recruitment proceeded iteratively alongside preliminary analysis. To ensure that participants could meaningfully engage with the research focus, instructors with demonstrable exposure to AI tools were selected. In this study, AI literacy was treated as an inclusion criterion rather than an outcome variable, because participants needed at least a basic awareness of how AI operates in educational settings in order to reflect on its affordances, limitations, and implications for CT in higher education.

Accordingly, eligible participants were those who had (a) prior experience with AI-driven platforms or applications, (b) the ability to critically comment on their affordances and limitations, and (c) varying degrees of AI literacy, which was expected to generate a richer and more nuanced dataset. As presented in Table 1, participants also demonstrated clear and transparent profiles, with diversity in age (28–60), gender (female/male), highest degree (MA/PhD), teaching experience (3–30 years), and professional teaching qualifications (e.g., TTC, CELTA, TESOL, IELTS).

Importantly, the final sample size was determined by data saturation, operationalized as the point at which additional interviews (the 10th teacher) no longer yielded substantively new codes or insights relevant to the research question. Interviews and initial coding were conducted concurrently, and recruitment continued until saturation was reached; in this dataset, no meaningfully new themes emerged in the final interviews, and saturation was achieved by the tenth interview. This approach aligns with qualitative methodological guidance that emphasizes saturation as the key criterion for adequacy in sample size determination (e.g., Creswell) and is consistent with work suggesting that thematic saturation often occurs within approximately 6–12 interviews in relatively homogeneous samples (Guest et al., 2006).

Materials

A two-phase instrumentation approach was employed to ensure that the research instruments were both purpose-built for the study's open-ended research question and grounded in established qualitative design principles. Specifically, the instruments were

Table 1 Participants' demographic information

Teachers	Age	Gender	Education	Teaching Experience	Teaching Qualifications	AI Tool Used
T1	50	female	MA TEFL	14 years	TTC, CELTA	Chat-GPT
T2	30	female	PhD TEFL	3 years	None	ZeroAI, Chat-GPT
T3	49	female	MA ELL	26 years	TTC	Chat-GPT
T4	30	female	PhD TEFL	14 years	TTC	Google Bard, Chat-GPT, Bing
T5	28	female	PhD TEFL	8 years	None	Chat-GPT
T6	35	male	PhD TEFL	17 years	TTC, TESOL	Bing, Chat-GPT
T7	31	male	MA TEFL	13 years	TTC, IELTS	Bing, Chat-GPT
T8	48	male	PhD TEFL	29 years	None	ZeroAI, Chat-GPT
T9	34	male	PhD TEFL	17 years	TESOL	Chat-GPT, Perplexity
T10	60	male	PhD Linguistics	30 years	TTC, TESOL	Chat-GPT

developed to elicit (a) instructors' conceptualizations of CT and AI, (b) their lived experiences of using AI in higher education, and (c) their perceived influences of AI on CT, including opportunities, constraints, and contextual conditions (cf. Kallio et al., 2016; Merriam & Tisdell, 2016; Patton, 2015).

Phase 1: Screening and profiling questionnaire (Google Form)

Prior to developing the interview protocol, a brief Google Form questionnaire was designed for two purposes that directly supported the study's qualitative aims: (1) participant profiling and eligibility screening, and (2) initial elicitation of participants' views to inform the interview design. The questionnaire included demographic and professional background items (e.g., age, gender, highest degree, and years of teaching experience), which facilitated transparent reporting of participant profiles and enabled contextual interpretation of responses. It also included one concise, open-ended prompt: *How do you think AI affects CT?* – with a short response requirement (approximately 50 words). This prompt functioned as an entry-point narrative that helped capture participants' preliminary sense-making and guided the refinement of the interview questions to better align with the study's open-ended research focus (Creswell & Poth, 2018; Patton, 2015).

The face validity and clarity of the Google Form were examined by an expert in the relevant field, who reviewed item wording, relevance to the research focus, and potential ambiguities. Minor revisions were made based on this feedback to improve comprehensibility and alignment with the study aims (Kallio et al., 2016).

Phase 2: Semi-structured interview protocol

Following the questionnaire stage, a semi-structured interview guide was developed by the research team based on (a) the study's research question, (b) related empirical and conceptual literature on AI use in higher education and teacher cognition, and (c) established conceptualizations of CT (e.g., Ennis, 2018). The interview format was selected because it allows for depth, flexibility, and probing, which are necessary for addressing open-ended qualitative questions about perceptions and experiences (Brinkmann & Kvale, 2015; Merriam & Tisdell, 2016).

The final protocol consisted of 12 core questions (Appendix A). To make the alignment explicit, the guide was organized to elicit three complementary types of data:

> Conceptual understanding (participants' definitions and interpretations of AI and CT),
> Experience-based accounts (how participants actually use AI tools in their teaching contexts), and
> Perceived influence and evaluation (how, why, and under what conditions AI may support or constrain CT, including affordances, limitations, and ethical concerns).

For example, the item *"What specific AI technologies or tools do you believe have the greatest potential to enhance your critical thinking skills?"* was designed to elicit concrete tool-mediated experiences and to connect participants' perceptions to dimensions of CT associated with Ennis's framework (e.g., making judgments, evaluating evidence,

and reflecting on decisions). The interview questions were reviewed and refined by the same expert who evaluated the questionnaire, confirming the relevance and appropriateness of the protocol for the study's open-ended research aims (Kallio et al., 2016). The interview schedule was adjusted to accommodate participants' availability and preferred mode (in-person at the university or virtually via Skype).

Ethical Procedures

Ethical considerations were addressed to ensure adherence to confidentiality and voluntary participation principles. Participation was fully voluntary, participants could withdraw at any time without consequence, and all shared information was used solely for research purposes. Interviews were anonymized at the transcription and translation stage, and pseudonyms (T1, T2, etc.) were used in place of real names. No identifiable personal details were reported.

Data Analysis

Deductive and inductive thematic analysis (Braun & Clarke, 2015) were employed to analyze the interview data, identifying major themes and subthemes. The analysis began with repeated reading of the Persian transcripts to achieve familiarization with the data, followed by initial coding of meaning units across the entire dataset. We used a both–and thematic approach because it best matched our aims and design. A deductive lens (Braun & Clarke, 2015) lets us organize what participants said in relation to our a priori framework: Ennis's four CT domains (clarification, advanced clarification, basis of inference, inference), so findings spoke directly to the research questions and existing theory. In parallel, an inductive lens ensured that analysis remained open to what the data added beyond the framework (e.g., unanticipated practices, tensions, and contextual constraints around AI use). This combination enabled us to move beyond descriptive summaries toward analytic claims that highlighted patterns, contradictions, and contextualized meanings in instructors' accounts. Practically, major themes and subthemes were identified through the following steps: transcripts were first coded into broad, theory-led families (deductive), then open-coded within each family to surface finer-grained, participant-led patterns (inductive); candidate themes were iteratively reviewed for internal coherence and distinctiveness across cases, then defined and named, with analytic memos documenting interpretive decisions at each stage. To ensure trustworthiness (Lincoln & Guba, 1985), three online sessions were conducted among researchers, resulting in modifications to the themes and subthemes. In reporting the findings, we sought to demonstrate data density by presenting each theme with illustrative and, where relevant, contrasting quotations, as well as noting nuances and tensions rather than only summarizing dominant views. Following Rowlands's (2021) interviewee transcript review (ITR) process, the verbatim transcripts were sent to the participants for member checking.

The interviews were conducted in Persian. Initial translations were produced using OpenAI's ChatGPT (GPT-3.5), after which all outputs were independently reviewed and corrected by the bilingual authors to ensure semantic fidelity. To further verify accuracy, a back-translation into Persian was conducted by a bilingual fellow researcher, with discrepancies reconciled by the research team. The translated interview texts

were subsequently uploaded into the MAXQDA software. Five interviews were coded by one researcher, and a second researcher analyzed the remaining five interviews following Ennis's (1987, 2018) protocol. To ensure consistency and reliability, three review sessions were held via Skype, during which all codes from the 10 interviews were compared and discussed among the researchers. Discrepancies were examined, negotiated, and resolved through consensus, resulting in an agreed coding frame and a consolidated set of themes and subthemes. Rather than computing a formal inter-coder reliability coefficient, we adopted this reflexive, team-based approach to coder agreement as more consistent with our qualitative–interpretivist orientation. The consolidated themes and subthemes (Table 2) were then reviewed thoroughly by the first author. The interpretations were meticulously examined, and the conceptual meanings were discussed with the other researchers. Taken together, member checking, iterative team-based coding, negotiated consensus, and careful translation procedures functioned as our strategy for validating the consistency and credibility of the analysis. This process resulted in detailed reports summarizing the findings, including categorized themes and subthemes, supported by quotes and visual data representations.

Findings

The qualitative analysis of semi-structured interviews with 10 participants revealed that the participants' perceptions of utilizing AI for enhancing educators' CT skills were structured around four key areas adapted from Ennis (1987): clarification, advanced clarification, basis of inference, and inference. Subthemes within these areas were identified,

Table 2 Themes and subthemes

	Frequency	Percentage	Percentage (valid)
Clarification\Analyzing Arguments	13	15.5	15.5
Basis ofinference\Judging credibility	12	14.3	14.3
Clarification\Providing Information	11	13.1	13.1
Clarification\Replying Questions	9	10.7	10.7
Clarification\Focusing	7	8.3	8.3
Advanced Clarification\Contextualization	4	4.8	4.8
Unknown	4	4.8	4.8
Advanced Clarification\Teaching Model	4	4.8	4.8
Basis ofinference\Feedback	4	4.8	4.8
Advanced Clarification\Defining Terms	3	3.6	3.6
Inference\Deducing &Inducing	3	3.6	3.6
Clarification\Modeling	3	3.6	3.6
Clarification\Exploration	2	2.4	2.4
Basis ofinference\Observing	1	1.2	1.2
Inference\Value Judgements	1	1.2	1.2
Basis ofinference\Reflection	1	1.2	1.2
MAGENTA	1	1.2	1.2
Advanced Clarification\IdentifyingAssumptions	1	1.2	1.2
Inference	0	0.0	0.0
Advanced Clarification	0	0.0	0.0
Basis ofinference	0	0.0	0.0
Clarification	0	0.0	0.0
TOTAL (valid)	84	100.0	100.0
Missing	0	0.0	
TOTAL	84	100.0	

aligning with Ennis's taxonomy (1987, 2018) and emerging from the study (Fig. 1). Across domains, participants consistently positioned AI as a partner for upstream meaning-making rather than a substitute for evaluative judgment, a stance that aligns with recent higher-education work on human-AI complementarity. Rather than treating AI as an autonomous decision-maker, participants described it as a generative, exploratory resource that supports, but does not replace, their own critical scrutiny, especially at later stages of reasoning.

Clarification

The skill of clarification involves the precise comprehension of content and identification of primary points, which manifests through diverse subthemes. Across the dataset, participants most frequently invoked AI in relation to clarification, describing it as a "first-pass" tool for organizing, condensing, and rendering complex content more tractable. Notably, participants utilized AI for tasks such as analyzing arguments and providing information. Additionally, although less frequently used, other important aspects included enhancing focus and creating models within this skill domain. For instance, six participants mentioned using AI for "argument analysis" and content organization through summaries and tables; in these cases, AI was not merely a convenience tool but a scaffold that helped them surface main claims, counterclaims, and supporting reasons more quickly than through manual processing alone. Notably, participants highlighted its role in drafting academic works and supporting instructors in lesson planning, curriculum design, and material development. For example, T1 mentioned "I use it for material development … planning … curriculum development … and lesson planning" (T1, Interview). In this view, AI serves as a powerful tool in different instructional contexts. However, even where AI seemed deeply integrated into planning practices, instructors framed it as a structuring aid that still required human oversight to ensure alignment with course aims, institutional constraints, and students' needs. Instructors utilized AI most in curricular changes, lesson planning, and material development, suggesting that it facilitates the creation of educational materials in a more structured approach.

In terms of "information provision", six participants referenced AI's ability to furnish comprehensive data. In today's digital era, instructors harness the power of AI for

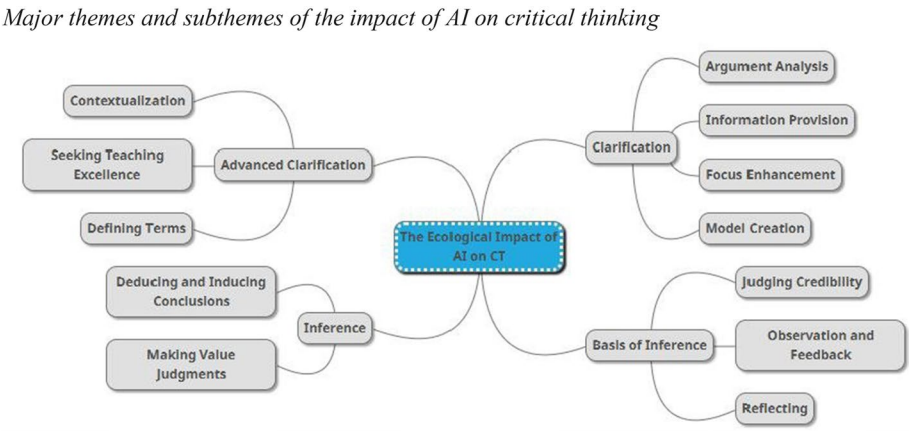

Fig. 1 Major themes and subthemes of the impact of AI on critical thinking

information retrieval and analysis. With AI's capability to comprehend vast amounts of data and provide detailed responses, instructors can streamline their research efforts and delve deeper into specific topics. Moreover, AI's ability to offer extensive information on diverse subjects proves invaluable in navigating the complexities of the modern educational landscape. As T8 argued, it can result in "a focused search on a topic to know the details" (T8, Interview). Alternatively, in the words of T6, "It has helped me a lot, providing a lot of information. It provides teachers and students with a wealth of information and lots of data that was previously inaccessible" (T6, Interview). For these teachers, AI broadened the evidential base on which they reasoned, allowing them to juxtapose multiple perspectives and sources that would otherwise remain hidden or time-consuming to locate. Integrating AI into educational practices noticeably improves the ability of instructors and students to access and analyze the data. AI tools help instructors streamline their research processes and engage deeply with specific topics by providing comprehensive data retrieval and detailed analysis. Moreover, a wide range of information accessible through AI can address the complex challenges in an educational environment. Participants nevertheless treated such outputs as starting points to refine prompts, cross-check sources, and select relevant evidence, maintaining teacher agency over what ultimately counts as credible. For instance, T6 stated that.

> *If you know how to ask ChatGPT or other similar AI applications, they can help you a lot. But it doesn't mean that they're writing a book for you or a paper for you, but they help you a lot. You know, they may give you ideas that may not come to you. But using them, you can have a much more open mind about what you're doing (T6, Interview).*

Additionally, three participants referred to the subtheme of "focus enhancement". Specifically, they emphasized the role of AI in providing precise and comprehensive information essential for enhancing focus, which involves identifying key elements such as main topics, claims, and conclusions. For example, T1 argued that AI "provides you with the exact and accurate kind of information that you need" (T1, interview). Another significant sub-theme is "model creation". In this context, instructors leverage AI to provide illustrative examples and exemplary models, enhancing students' comprehension and learning outcomes. This is evident in the next extract:

> *So now imagine that you can have the best kind of model for your students, and they can learn from the best. I also believe that teachers should use AI to bring some writing examples and use them in classes. (T1, Interview)*

Participants noted the transformative nature of AI in education, particularly in "focus enhancement" and information processing, which enables them to identify key materials that foster a deeper understanding. Additionally, considering "model creation", AI generates practical examples to improve academic performance and learning. Yet, a minority of participants voiced concerns that over-reliance on AI-generated models might narrow students' exposure to idiosyncratic or imperfect texts they must also learn to critique, underscoring the need for deliberate curation of AI outputs. Overall, sub-themes reflect a shift toward an effective AI-centered pedagogy.

Advanced Clarification

In the domain of advanced clarification, three participants (i.e., T1, T2, and T10) frequently engage in "contextualization", emphasizing the "seeking teaching excellence", and the "precise definition of terms". The sub-theme of "contextualization" within the overarching theme of advanced clarification primarily involves the use of AI to contextualize vocabulary and grammar concepts. Through this approach, AI generates examples and provides contextualized instances to aid in understanding, thereby facilitating learning within tangible contexts. For instance, T10 expressed that "AI delivers this vocabulary to me in the form of text" (T10, Interview). T1 noted that "you can devise different kinds of tasks, you can actually teach one simple vocabulary activity, which is a gap filler in different ways" (T1, Interview). Here, AI was described as a flexible generator of micro-contexts (short texts, dialogues, or tasks) that allowed instructors to experiment with multiple representations of the same concept and to select those that best matched learners' proficiency and interests.

Another notable sub-theme identified is "seeking teaching excellence". This encompasses the various methods, models, and approaches that instructors seek to employ in order to enhance their teaching efficacy within the classroom setting. For instance, one teacher expressed a desire for an optimal model that would facilitate optimal learning outcomes for students.

> *I don't know what method and combination of methods to use to teach, for example, how this content will be learnt better. We can get help from AI to get the task. we can get help from this. Now, how to implement the task. (T2, Interview)*

AI plays a significant role in teaching and learning by contextualizing words and grammatical structures. Instructors strive for teaching excellence by exploring different methods and instructional models that can enhance students' engagement and learning outcomes. Integrating AI into educational settings enables instructors to provide contextualized examples and receive guidance on teaching methods tailored to specific objectives. At the same time, some participants were cautious about generic AI pedagogical advice, noting that recommendations still needed to be adapted to local curricular demands, assessment regimes, and institutional cultures. By adopting a collaborative approach, instructors can refine their teaching practices.

"Defining terms" is a prominent application of AI and a key aspect of advanced clarification. This method facilitated participants' understanding of unfamiliar concepts by providing comprehensive explanations and defining the meaning of ambiguous terms by inferring from the context, as T3 exemplified it as finding "the meaning of the word in the context by giving the paragraph to the software (T3, Interview). In this extract, AI was not only a dictionary-like tool but also a semantic explainer that unpacked technical or emerging AI-related terminology, thereby enabling instructors to position themselves more confidently in disciplinary debates around AI and CT.

Basis of Inference

Seven participants predominantly utilized AI to judge credibility and four participants used AI to provide feedback. These functions were the primary ways AI was applied.

In contrast, the subthemes of observing and reflecting were identified as less frequently used, indicating that participants did not rely on AI as much for these purposes. "Judging credibility" involved assessing the trustworthiness of information, while "providing feedback" entailed offering constructive critiques. Judging credibility, crucial for evaluating the validity of information, was widely employed. This process involves systematically comparing and analyzing the content across multiple sources to ensure reliability and accuracy.

> *For example, I myself used artificial intelligence … to know whether it is good for teaching language or not. I used critical thinking here. Because we wonder whether artificial intelligence can be a substitute for books and traditional education. (T9, Interview)*

Four participants primarily engaged with AI for its "judging credibility" and three participants for "feedback" to enhance their critical evaluation in education and information assessment. They also pinpointed the AI's effectiveness as an alternative to obsolete educational sources. Furthermore, they emphasize a careful evaluation of AI's role in language learning, suggesting cautious optimism concerning AI's application in education. In practice, AI-assisted credibility checks (e.g., generating criteria, compiling candidate references), but final warranting remained human-led.

One participant described AI as simultaneously supporting "observation": the appraisal of evidence, coherence, and the validity of conclusions, and four participants declared "feedback": targeted guidance that helps correct errors and deepen understanding within the basis of inference domain. As one teacher noted, "sometimes AI can help you with your critique and critical thinking" (T1, Interview), indicating its role as a first-pass observer that flags inconsistencies and prompts further scrutiny. In parallel, AI's feedback functions were seen as especially useful when immediate human input is unavailable: "for writing skills, when a student gets stuck in writing a second or third language, and there is no teacher, give him feedback" (T3, Interview). In such cases, real-time suggestions can identify errors and propose revisions that facilitate learning in the moment. At the same time, four participants cautioned that uncritical reliance on AI-generated feedback may displace instructors' own evaluative reasoning; they emphasized using AI outputs as prompts for reflective judgment rather than substitutes for it, thereby preserving and potentially sharpening instructors' CT. This tension between efficiency and dependence recurred across interviews and marked the basis-of-inference domain as a key site where instructors negotiated boundaries around acceptable AI involvement.

"Reflecting" is a crucial aspect of the basis of inference, involving the thorough examination of content to ensure consistency of evidence and validity of conclusions. AI enhances this reflective process by providing tools for deeper analysis and highlighting patterns or inconsistencies. "If you actually try to be a kind of reflective teacher, to learn from your past and have a prospective view toward your future, you need to be in constant contact with technology" (T1, Interview). This integration of AI supports and strengthens CT, leading to more accurate inferences. In this regard, AI can assist instructors in assessing content effectively and deepen instructors' understanding of the information's inconsistencies, leading them to valid conclusions. However, only a subset

of instructors reported using AI explicitly for post-hoc reflection (e.g., asking AI to summarize their own lesson plans or teaching narratives), suggesting that reflective use of AI remains an emerging rather than a fully consolidated practice.

Inference

Finally, within the inference domain, four participants extensively utilized AI for "deducing and inducing conclusions", demonstrating the significant role of AI in facilitating the process of drawing logical inferences and making informed judgments based on available evidence and "making value judgments". For instance, T10 mentioned, "If we want to give a text to the student that focuses on the method and wants to draw conclusions from the text, it will make the strong imagination work more and become more creative" (T10, Interview). Participants utilized AI to deduce and induce conclusions, deriving accurate hypotheses and inferences. This involved analyzing complex data to identify patterns and relationships and developing decision-making processes and overall outcomes. Some participants also used AI dialogically, simulating "what-if" scenarios and asking AI to generate alternative explanations or solutions, which they then compared with their own reasoning. AI's role in induction and deduction enables participants to draw meaningful inferences and hypotheses. The more they are engaged with AI, the better their CT skills will get. Making an impartial decision involves considering a multitude of factors, including background elements, and achieving balance. This process involves meticulously evaluating various factors to ensure fairness and equity in decision-making. In this regard, T6 stated that "If we can systematically analyze the arguments and information, we can arrive at a well-informed decision" (T6, Interview). To ensure fairness and impartiality, decision-makers need to consider various factors and process the information thoroughly. This minimizes bias and contributes to the credibility of the decision-making process. Yet, participants were clear that the final responsibility for judgment rested with human teachers, particularly when ethical, political, or high-stakes pedagogical decisions were involved; AI was framed as an input to, not an arbiter of, value judgments.

In sum, AI catalyzed brainstorming and scenario exploration, while teachers moderated confidence, weighed counter-evidence, and aligned conclusions with explicit criteria. Overall, the dataset portrays selective, pragmatic uptake: AI is embraced for Clarification and Advanced Clarification (summarizing, structuring, disambiguating, assumption-surfacing), while the Basis of Inference and Inference are safeguarded through teacher-led verification, provenance checks, and dialogic justification. This pattern suggests that, for these Iranian university instructors, AI currently functions most comfortably as a cognitive amplifier at earlier stages of thinking, whereas later evaluative and ethical moves remain firmly anchored in human judgment, thereby preserving a human-in-the-loop model of CT with AI.

Discussion

This study offers a situated account of how 10 Iranian higher-education language instructors perceive the impact of using AI on CT. We interpret the findings as a context-bound elaboration of how instructors align AI uses with the domains of clarification, advanced clarification, basis of inference, and inference (Ennis, 1978, 2018). In a field where most

empirical work on AI and CT has focused on students in STEM or Anglophone settings, we add a small but underrepresented perspective: TEFL instructors working in a resource-constrained, exam-oriented higher-education context in Iran. Our contribution is therefore descriptive-interpretive and practice-oriented, not theory-revising.

Across clarification and advanced clarification, participants reported using AI to rephrase dense texts, map arguments, surface ambiguities, and propose definitions. These patterns align with evidence that generative AI can assist students in interpreting instructor feedback and planning revisions when combined with human judgment (Lu et al., 2024), and that AI feedback can complement (rather than replace) peer/human feedback in academic writing (Banihashem et al., 2024). In teacher education settings, generative AI has also been explored for lesson planning and fostering openness and critical inquiry (van den Berg & du Plessis, 2023; Javahery et al., 2025). In summary, these alignments suggest that AI can accelerate meaning-making processes upstream (summarizing, structuring, and disambiguating) when its outputs are framed as tentative prompts rather than authoritative answers. This emerges from the data showing that some instructors simultaneously worry that heavy reliance on AI-generated "models" may narrow students' exposure to imperfect, authentic texts they must also learn to critique, signaling a tension that is not foregrounded in much of the existing, more celebratory work on GenAI as a clarification aid.

By contrast, instructors were markedly cautious about relying on AI as a basis of inference; that is, as a warrant for claims or as a substitute for source appraisal. This hesitation mirrors broader perceptions of AI as an assistive partner rather than a teacher substitute (Chan & Tsi, 2024) and reflects design concerns raised in learning analytics reviews: feedback pipelines can reconfigure what counts as evidence and who judges it unless criteria are made visible and contestable (Banihashem et al., 2022). Participants' stance therefore converges with a "human-in-the-loop" position: use AI to scaffold exploration, but anchor evaluative judgments in teacher-mediated verification and explicit standards. What our mapping adds here is a more granular, domain-level gradient: AI is welcomed for clarification and advanced clarification, treated ambivalently for the basis of inference, and tightly constrained for final inferential moves, suggesting a practical boundary line that instructors draw between acceptable cognitive outsourcing and non-negotiable human judgment.

With respect to inference (generating alternatives, considering counter-evidence, and calibrating conclusions), instructors described AI as a catalyst for idea generation and scenario exploration, yet they emphasized the need for human calibration to avoid hallucinated details, spurious warrants, and over-confident conclusions. This division of cognitive labor is consonant with a design-oriented hybrid intelligent feedback framework in which human and GenAI roles are specified across the feedback cycle (Banihashem et al., 2025). In our data, the "hybrid" sweet spot appears upstream (clarification/advanced clarification), whereas downstream (basis of inference/inference) requires tighter human control over credibility checks and justificatory fit. We also note that reflective uses of AI (e.g., asking AI to summarize one's own teaching narratives or lesson plans for self-review) were reported only by a minority of participants, indicating an emergent, not yet mainstream, practice that future work could intentionally cultivate and study.

These dynamics also touch teacher identity and agency (Satvati et al., 2025). Far from being displaced, participants framed their role as orchestrators of evidence and mentors of judgment, a stance consistent with reports that teachers and students favor coexistence models that keep human creativity, ethics, and socio-emotional capacities central (Chan & Tsi, 2024). Attitudinal modeling in post-pandemic blended contexts similarly reveals that instructors and students negotiate evolving roles and emotions within datafied, AI-adjacent ecosystems, highlighting the need for professional development in critical data literacy and feedback design (Banihashem et al., 2023). Our participants' accounts concretize these identity negotiations in a specific TEFL context: they describe themselves less as "guards against AI misuse" and more as designers of tasks and criteria that discipline AI's role, which may be a useful reframing for PD in comparable settings.

Incorporating recent literature helps situate our findings within two foundational tensions: (1) AI as cognitive aid versus epistemic threat (Chan & Tsi, 2024; Lu et al., 2024), and (2) automation logics versus teacher agency (van den Banihashem et al., 2022; Berg & du Plessis, 2023). Our participants largely resolve these tensions by adopting a partnering model: leveraging AI for ideation, organization, and formative cues, while safeguarding human-led evaluation, ethics, and contextual judgment. What this study adds is not a new tension, but an illustration of how that partnering model is locally instantiated across Ennis's four domains of CT in one under-represented higher-education context.

A further implication concerns task architecture. Instructors can operationalize a CT-aligned workflow across a sequence of activities: (a) Clarification: use AI to draft argument maps, list reasons, and expose definitional ambiguities; (b) Advanced clarification: prompt AI to surface assumptions and competing interpretations, then require students to label which are plausible and why; (c) Basis of inference: shift to human-led verification (source provenance, triangulation, and warrant–claim fit) using rubrics aligned with course outcomes and visible to students; (d) Inference: stage a dialogic justification (student + teacher) in which conclusions are graded on explicit criteria for validity and attention to counter-evidence. Evidence from higher education suggests this orchestration keeps AI's strengths while mitigating epistemic outsourcing (Banihashem et al., 2022, 2025; Lu et al., 2024). Our proposed workflow is deliberately modest in scope: it is reverse-engineered from the patterns and tensions reported by participants rather than imposed a priori, and is intended as a context-sensitive starting point that other programs can adapt and empirically test rather than as a universal model.

Context also matters. Resource variability, institutional policy, and assessment regimes likely shape how far teachers push AI beyond preliminary analysis. For example, where assessment rubrics foreground source credibility and methodological fit, instructors may more readily contain AI's role to pre-writing and exploratory synthesis; where production pressures dominate, AI may drift into unvetted evidence provision unless governance makes criteria visible and contestable (Banihashem et al., 2022; van den Berg & du Plessis, 2023). This echoes our participants' pragmatic selectivity and helps explain cross-study variability. In the Iranian universities represented here, participants described navigating additional constraints (e.g., infrastructural

instability, uneven institutional guidance), which may partially explain their strong preference for using AI in low-stakes, upstream tasks rather than in high-stakes inferential work; this context-specific pattern may be informative for systems with similar constraints.

For professional development, we suggest micro-credentialed modules that pair generative-AI literacies (prompt design, critique, and revision planning) with critical data literacy (understanding dashboards, bias, and evidence provenance). Department-level communities of practice can curate vetted prompt–rubric bundles and exemplars of "good" justification chains, thereby accelerating responsible reuse (Banihashem et al., 2023, 2025). These recommendations directly reflect areas of uncertainty voiced by participants (e.g., how to constrain AI's role in assessment, how to explain credibility checks to students), and should therefore be read as grounded in their reported needs rather than extrapolated solely from the international literature.

We deliberately temper our contribution claims. We do not assert that our data modernizes Ennis's taxonomy or yields "novel insights" into AI-integrated pedagogy at a theoretical level. Rather, we offer three incremental, practice-facing contributions: (a) an empirically grounded mapping of instructors' selective AI uptake onto Ennis's four domains in a specific TEFL context; (b) the identification of domain-specific tensions (e.g., enthusiasm for AI-generated models co-existing with concern about narrowed textual exposure; emergent but limited reflective uses of AI) that flesh out how "human-in-the-loop" logics are lived in practice; and (c) a plausible, literature-aligned workflow for hybrid use that prioritizes teacher agency. These insights are necessarily exploratory and context-bound, given our small, homogeneous sample and constrained analytic approach. We therefore frame them as hypotheses and design prompts to be tested, challenged, and refined in larger, multi-site studies, rather than as definitive claims.

Conclusion

This study reveals that Iranian language instructors recognize the significant potential of AI to enhance their CT skills in higher education. Participants identified AI's ability to facilitate concept clarification and advanced analysis, enabling them to deconstruct complex ideas and effectively evaluate underlying assumptions. However, the reliance on AI for foundational inference skills was perceived as limited, indicating a need for careful integration of AI technologies in teaching practices. Participants emphasized that while AI can support tasks such as summarizing content and drafting academic documents, instructors should be vigilant to critically assess information and make sound judgments. Overall, the findings highlight a nuanced understanding of AI's role in supporting various aspects of CT, from basic clarification to advanced inference.

The implications of this research for higher education are profound, suggesting a need for curricula that effectively integrate AI into pedagogical practices while fostering CT skills. As AI tools become increasingly prevalent, instructors must adapt their teaching strategies to leverage these technologies while ensuring students and themselves engage in meaningful critical analysis. Training programs for teachers should emphasize the use of AI to enhance instructional methods, promote deeper cognitive development, and encourage independent inquiry among students. Furthermore, educational institutions should consider developing frameworks for evaluating the credibility of AI-generated

content, fostering an environment where students can discern reliable information in an era dominated by digital resources. Ultimately, this research advocates for a balanced approach that recognizes AI as a powerful ally in education, one that complements rather than replaces the essential role of instructors in nurturing CT.

Despite its contributions, this study has several limitations. First, the small, purposive sample (n = 10) of Iranian university language instructors limits transferability to other disciplines, institutions, and contexts; moreover, the inclusion criterion of AI literacy (prior demonstrable use of AI tools) likely attracted instructors who are more familiar with or positively disposed toward AI, introducing potential self-selection bias and under-representing low-use or more sceptical perspectives. Second, the length and depth of interviews varied due to teaching schedules, which may have limited probing and narrowed the range of examples. The data are self-reported and cross-sectional, with no classroom observations or direct assessments of CT performance to triangulate claims; therefore, the findings pertain to perceived rather than measured CT. Third, transcripts were translated from Persian to English using ChatGPT (GPT-3.5) and then reviewed by the bilingual authors with back-translation to check fidelity; nonetheless, subtle nuances of meaning, register, and culture may have been lost or reshaped by the translation pipeline. Fourth, while the adapted Ennis framework provided analytic scaffolding, it may also have constrained coding, and participants likely held heterogeneous understandings of both "AI" (used broadly for a range of tools) and CT, introducing interpretive variability. Fifth, although we employed team-based coding and negotiated consensus, we did not compute a formal inter-coder reliability coefficient (e.g., Krippendorff's alpha), which limits comparability with more quantitatively oriented qualitative studies. Finally, although thematic saturation was judged pragmatically, it cannot be assumed with a relatively homogenous, modest sample. These constraints necessitate larger, multi-site, mixed-methods, and longitudinal designs with richer triangulation.

Appendix A

Semi-structured interview core questions

(1) What are the upsides and downsides of using AI when teaching?

(2) How often do you use AI in the classroom, institution, and society?

(3) To what extent has AI helped you to become critical in the classroom, institution, and society? give examples?

(4) How do you perceive the impact of AI on students' critical thinking abilities in your own classrooms? Why do you think it's effective or not?

(5) What specific AI technologies or tools do you believe have the greatest potential to enhance critical thinking skills in your students?

(6) How do you feel about your own ability to effectively integrate AI technologies into your teaching practices to promote critical thinking?

(7) How do school policies and resources shape your perception and use of AI technologies to promote critical thinking skills?

(8) In what ways did professional development programs influence your understanding and utilization of AI technologies to enhance critical thinking?

(9) How do educational stakeholders (e.g., policymakers, administrators) affect your perception of the potential benefits and risks of integrating AI into classrooms for critical thinking development?

(10) What systemic factors (e.g., funding, government regulations) influence the inclusion of AI in educational settings to support critical thinking?

(11) What are the potential long-term impacts of widespread AI integration on critical thinking skills at a societal level?

(12) How do global trends and advancements in AI impact your perceptions and readiness to incorporate AI for critical thinking development?

Author Contribution

Jaber Kamali: conceptualization (lead); data curation (supporting); formal analysis (supporting); methodology (equal); supervision (lead); writing – original draft (equal); writing – review and editing (equal). Ramin Ali Pour: conceptualization (supporting); data curation (lead); formal analysis (lead); methodology (equal); writing – original draft (equal); writing – review and editing (equal). Mina Naderi: conceptualization (supporting); data curation (supporting); formal analysis (equal); methodology (supporting); writing – original draft (supporting); writing – review and editing (equal). Mahtab Afrasiabi: conceptualization (supporting); data curation (supporting); formal analysis (equal); methodology (supporting); writing – original draft (supporting); writing – review and editing (equal). Faramarz Piri: conceptualization (supporting); data curation (supporting); formal analysis (supporting); methodology (supporting); writing – original draft (supporting); writing – review and editing (equal).

Funding

This research received no external funding.

Data Availability

Data from this study (interview audios and their transcriptions) are available and will be shared upon request.

Declarations

Ethics approval and consent to participate

This study received ethical approval from the authorities of the higher education institutions where the participants were working and was conducted in accordance with the ethical guidelines and principles outlined in the Declaration of Helsinki. All research procedures adhered to institutional and international ethical standards for research involving human participants. All participants were fully informed about the purpose, procedures, potential risks, and benefits of the study. All participants in this study provided informed consent to take part. Participation was entirely voluntary, and individuals had the right to withdraw at any time without consequence. Confidentiality and anonymity of the participants were maintained throughout the research process.

Consent for publication

All authors consent to the submission of this manuscript to the ***Asian-Pacific Journal of Second and Foreign Language Education*** for consideration. We confirm that the work is original, has not been published elsewhere, and is not under review in any other journal. If accepted, we agree to its publication in accordance with the journal's policies.

Conflict of interests

The authors declare no competing interests.

Received: 1 February 2025 Accepted: 3 March 2026
Published online: 08 May 2026

References

Alwaqdani, M. (2024). Investigating teachers' perceptions of artificial intelligence tools in education: Potential and difficulties. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-12903-9

Banihashem, S. K., Noroozi, O., van Ginkel, S., Macfadyen, L. P., & Biemans, H. J. A. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review, 37*, Article 100489. https://doi.org/10.1016/j.edurev.2022.100489

Banihashem, S. K., Noroozi, O., den Brok, P., Biemans, H. J. A., & Taghizadeh Kerman, N. (2023). Modeling teachers' and students' attitudes, emotions, and perceptions in blended education: Towards post-pandemic education. *The International Journal of Management Education, 21*(2), Article 100803. https://doi.org/10.1016/j.ijme.2023.100803

Banihashem, S. K., Taghizadeh Kerman, N., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: Peer-generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education, 21*(1), Article 23. https://doi.org/10.1186/s41239-024-00455-4

Banihashem, S. K., Noroozi, O., Khosravi, H., Schunn, C. D., & Drachsler, H. 2025. Pedagogical framework for hybrid intelligent feedback. *Innovations in Education and Teaching International.* Advance online publication. https://doi.org/10.1080/14703297.2025.2499174
Bernstein, B. (2000). *Pedagogy, symbolic control, and identity : Theory, research, critique*. Rowman & Littlefield.
Bernstein, B. (2004). The structuring of pedagogic discourse. *Routledge*. https://doi.org/10.4324/9780203011263
Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & 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*(1), Article Article 4. https://doi.org/10.1186/s41239-023-00436-z
Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
Brinkmann, S., & Kvale, S. (2015). *InterViews: Learning the craft of qualitative research interviewing* (3rd ed.). SAGE.
Chan, C. K. Y., & Tsi, L. H. Y. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation, 83*, Article 101395. https://doi.org/10.1016/j.stueduc.2024.101395
Coeckelbergh, M. (2020). AI ethics. *The MIT Press*. https://doi.org/10.7551/mitpress/12549.001.0001
Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*. https://doi.org/10.1080/14703297.2023.2190148
Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.
Ennis, R. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9–26). W. H. Freeman and Company.
Ennis, R. (2018). Critical thinking across the curriculum: A vision. *Topoi, 37*(1), 165–184. https://doi.org/10.1007/s11245-016-9401-4
Felix, C. V. (2020). The Role of the Teacher and AI in Education. In E. Sengupta, P. Blessinger, & M. Makhanya (Eds.). *International Perspectives on the Role of Technology in Humanizing Higher Education* (pp. 33–48). Emerald. https://doi.org/10.1108/s2055-364120200000033003
Fuchs, C. (2017). *Social Media: A Critical Introduction (2nd edition)*. Sage.
Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59–82. https://doi.org/10.1177/1525822X05279903
Gulz, A., Londos, L., & Haake, M. (2020). Preschoolers' understanding of a teachable agent-based game in early mathematics as reflected in their gaze behaviors – An experimental study. *International Journal of Artificial Intelligence in Education, 30*(1), 38–73. https://doi.org/10.1007/s40593-020-00193-4
Guo, Y., & Wang, Y. (2024). Exploring the effects of artificial intelligence application on EFL students' academic engagement and emotional experiences: A mixed-methods study. *European Journal of Education*. https://doi.org/10.1111/ejed.12812
Hagenauer, G., Muehlbacher, F., & Ivanova, M. (2023). "It's where learning and teaching begins-is this relationship"—Insights on the teacher-student relationship at university from the teachers' perspective. *Higher Education, 85*(4), 819–835. https://doi.org/10.1007/s10734-022-00867-z
Hébert, C., Jenson, J., & Terzopoulos, T. (2021). Access to technology is the major challenge: Teacher perspectives on barriers to DGBL in K-12 classrooms. *E-Learning and Digital Media, 18*(3), 307–324. https://doi.org/10.1177/2042753021995315
Hopcan, S., Türkmen, G., & Polat, E. (2024). Exploring the artificial intelligence anxiety and machine learning attitudes of teacher candidates. *Education and Information Technologies, 29*(6), 7281–7301. https://doi.org/10.1007/s10639-023-12086-9
Huang, X., & Qiao, C. (2024). Enhancing computational thinking skills through artificial intelligence education at a STEAM high school. *Science & Education, 33*, 383–403. https://doi.org/10.1007/s11191-022-00392-6
Hwang, G., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers & Education: Artificial Intelligence, 1*, Article 100001. https://doi.org/10.1016/j.caeai.2020.100001
Istenic, A., Bratko, I., & Rosanda, V. (2021). Pre-service teachers' concerns about social robots in the classroom: A model for development. *Education & Self Development, 16*, 60–87. https://doi.org/10.26907/esd.16.2.05
Javahery, P., Alpat, M. F., & Kamali, J. (2025). Exploring Iranian novice EFL trainees' perceptions of ChatGPT use for lesson planning through a critical digital literacy lens. *Discover Computing, 28*(1), Article 194. https://doi.org/10.1007/s10791-025-09720-0
Kallio, H., Pietilä, A.-M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing, 72*(12), 2954–2965. https://doi.org/10.1111/jan.13031
Kamali, J., Freeman, D., Larsen-Freeman, D., Norton, B., & Farrell, T. S. C. (2024a). Editorial: Language teacher education research – key trends, challenges, and questions. *Language Teacher Education Research, 1*, 1–16. https://doi.org/10.32038/lter.2024.01.01
Kamali, J., Alpat, M. F., & Bozkurt, A. (2024b). AI ethics as a complex and multifaceted challenge: Decoding educators' AI ethics alignment through the lens of activity theory. *International Journal of Educational Technology in Higher Education, 21*(1), 1–20. https://doi.org/10.1186/s41239-024-00496-9
Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability, 15*(16), Article 12451. https://doi.org/10.3390/su151612451
Kim, J. (2024). Leading teachers' perspective on teacher-AI collaboration in education. *Education and Information Technologies, 29*(7), 8693–8724. https://doi.org/10.1007/s10639-023-12109-5
Kim, N., & Kim, M. K. (2022). Teacher's perceptions of using an artificial intelligence-based educational tool for scientific writing. *Frontiers in Education*. https://doi.org/10.3389/feduc.2022.755914
Kim, N., Vicentini, C., & Belland, B. R. (2021). Influence of scaffolding on information literacy and argumentation skills in virtual field trips and problem-based learning for scientific problem solving. *International Journal of Science and Mathematics Education, 20*(2), 215–236. https://doi.org/10.1007/s10763-020-10145-y

King, A. (1993). From sage on the stage to guide on the side. *College Teaching, 41*(1), 30–35. https://doi.org/10.1080/87567555.1993.9926781
Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M., Păun, D., & Mihoreanu, L. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. *Sustainability, 13*(18), Article 10424. https://doi.org/10.3390/su131810424
Li, G., Zarei, M. A., Alibakhshi, G., & Labbafi, A. (2024). Teachers and educators' experiences and perceptions of artificial-powered interventions for Autism groups. *BMC Psychology, 12*(1), Article 199. https://doi.org/10.1186/s40359-024-01664-2
Lin, Y.-S., Chen, S.-Y., Tsai, C.-W., & Lai, Y.-H. (2021). Exploring computational thinking skills training through augmented reality and AIoT learning. *Frontiers in Psychology, 12*, 1–9. https://doi.org/10.3389/fpsyg.2021.640115
Lin, C. C., Huang, A. Y., & Lu, O. H. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: A systematic review. *Smart Learning Environments, 10*(1), Article 41. https://doi.org/10.1186/s40561-023-00260-y
Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Sage.
Liu, G. X., & Wang, Y. (2024a). Modelling EFL teachers' intention to integrate informal digital learning of English (idle) into the classroom using the theory of planned behavior. *System, 120*, Article 103193. https://doi.org/10.1016/j.system.2023.103193
Liu, W., & Wang, Y. (2024b). The effects of using AI tools on critical thinking in English literature classes among EFL learners: An intervention study. *European Journal of Education*. https://doi.org/10.1111/ejed.12804
Lu, Q., Yao, Y., Xiao, L., Yuan, M., Wang, J., & Zhu, X. (2024). Can ChatGPT effectively complement teacher assessment of undergraduate students' academic writing? *Assessment & Evaluation in Higher Education, 49*(5), 616–633. https://doi.org/10.1080/02602938.2024.2301722
Luckin, R., & Holmes, W. H. (2016). *Intelligence Unleashed: An argument for AI in Education*. Pearson Education.
Mah, D.-K., Knoth, N., & Egloffstein, M. (2025). Perspectives of academic staff on artificial intelligence in higher education: Exploring areas of relevance. *Frontiers in Education, 10*, Article 1484904. https://doi.org/10.3389/feduc.2025.1484904
Manning, M. N. J. (2024). University lecturers' lived experiences of teaching critical thinking in Australian university: A hermeneutic phenomenological research. *Higher Education*. https://doi.org/10.1007/s10734-024-01200-6
McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2024). Generative AI chatbots in higher education: A review of an emerging research area. *Higher Education*. https://doi.org/10.1007/s10734-024-01288-w
McPeck, J. E. (2016). Critical Thinking and Education. *Routledge*. https://doi.org/10.4324/9781315463698
Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., Marlini, C., Zefrin, Z., & Fuad, Z. A. (2025). Critical thinking in the age of AI: A systematic review of AI's effects on higher education. *Educational Process: International Journal, 14*, Article e2025031. https://doi.org/10.22521/edupij.2025.14.31
Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.
Moorhouse, B. L., & Wong, K. M. (2025). *Generative Artificial Intelligence and Language Teaching*. Cambridge University Press.
Muthmainnah, Seraj, & P. M. I., & Oteir, I. (2022). Playing with AI to investigate human-computer interaction technology and improving critical thinking skills to pursue 21st century age. *Education Research International*. https://doi.org/10.1155/2022/6468995
Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE.
Popenici, S., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning, 12*, Article 22. https://doi.org/10.1186/s41039-017-0062-8
Rowlands, J. (2021). Interviewee transcript review as a tool to improve data quality and participant confidence in sensitive research. *International Journal of Qualitative Methods*. https://doi.org/10.1177/16094069211066170
Salido, A., Syarif, I., Sitepu, M. S., Suparjan, Wana, & P. R., Taufika, R., & Melisa, R. (2025). Integrating critical thinking and artificial intelligence in higher education: A bibliometric and systematic review of skills and strategies. *Social Sciences & Humanities Open*. https://doi.org/10.1016/j.ssaho.2025.101924
Satvati, N., Kamali, J., Safian Boldaji, F., Khodadadi, M., & Akhondi, S. (2025). AI integration into language education and teacher identity: An ecological perspective. *Language Teaching Research Quarterly, 47*, 1–19. https://doi.org/10.32038/ltrq.2025.47.01
Seo, J. Y. (2025). Cornell note-taking strategy instruction for Gen Z: Enhancing EFL students' reading comprehension. *Asian-Pacific Journal of Second and Foreign Language Education, 10*(1), 40. https://doi.org/10.1186/s40862-025-00347-8
Singh, P. (2002). Pedagogising knowledge: Bernstein's theory of the pedagogic device. *British Journal of Sociology of Education, 23*(4), 571–582. https://doi.org/10.1080/0142569022000038422
Tay, D., & Xie, D. (2025). Fingerprints of EFL writing: An AI deep learning approach. *Asian-Pacific Journal of Second and Foreign Language Education, 10*(1), Article 43. https://doi.org/10.1186/s40862-025-00351-y
van den Berg, G., & du Plessis, E. (2023). ChatGPT and generative AI: Possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. *Education Sciences, 13*(10), Article 998. https://doi.org/10.3390/educsci13100998
Walter, Y. (2024). Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education, 21*(1), Article 15. https://doi.org/10.1186/s41239-024-00448-3
Zheng, J., Zhang, Q., Xu, S., Peng, H., & Wu, Q. (2018). Cognition-based context-aware cloud computing for intelligent robotic systems in mobile education. *IEEE Access, 6*, 49103–49111. https://doi.org/10.1109/access.2018.2867880

Publisher's Note