International Journal of TESOL Studies (2026)
Vol. 8 (4) 164-184 https://doi.org/10.58304/ijts.250522

Article

Critical Thinking in the Age of Generative AI: Effects of a Short-Term Experiential Learning Intervention on EFL Learners

Ngo Cong-Lem*
Thang Tat Nguyen
Khanh Nhat Hoang Nguyen
Dalat University, Vietnam

Received: 20 March 2025/Received in revised form: 10 May 2025/Accepted: 16 May 2025/ Available online: 28 May 2025

Abstract
The integration of generative AI tools such as ChatGPT has transformed English as a Foreign Language (EFL) education, offering new opportunities for supporting writing, research, and critical inquiry. However, unguided use of AI may foster cognitive passivity and over-reliance, highlighting the need for targeted pedagogical interventions. Grounded in experiential learning theory, this quasi-experimental study, employing a pretest–posttest control group design, evaluated the effectiveness of a 90-minute workshop, the Critical AI Engagement Cycle, in enhancing EFL learners' critical thinking skills when using ChatGPT. Despite the short duration, the workshop included multiple scaffolded activities designed to stimulate immediate critical reflection. Seventy-two undergraduate and graduate students at a Vietnamese public university participated, with 38 assigned to the experimental group and 34 to the control group. Participants were selected using convenience sampling based on course enrollment and availability. Pre- and post-test results demonstrated statistically significant improvements in overall critical thinking and each of the four subdomains—analytical skills, logical reasoning, evidence evaluation, and open-mindedness—among participants in the experimental group. Notably, the consistent and large effect sizes across all critical thinking subdomains (Cohen's d = 0.94 to 1.23) underscore the robust impact of the intervention. The experimental group significantly outperformed the control group in post-intervention critical thinking scores, even after controlling for pretest scores, gender, prior AI knowledge, and AI skill level, as confirmed by ANCOVA analyses. The results suggest that even brief, theoretically grounded interventions can significantly enhance critical thinking skills in AI-mediated EFL environments. These findings underscore the importance of evidence-informed practices and highlight the need for explicit critical thinking training to ensure sustainable and responsible educational practices in the age of generative AI.

Keywords
ChatGPT, critical thinking, EFL, experiential learning, intervention, Vietnam

*Corresponding author. Email: lemnc@dlu.edu.vn

1 Introduction

The integration of generative artificial intelligence (AI) tools such as ChatGPT has rapidly reshaped English as a Foreign Language (EFL) education, providing unprecedented opportunities for enhancing research skills, writing proficiency, and critical thinking (Abdelhalim, 2024; Alshammari, 2024; Zou et al., 2024). ChatGPT has been found to support learner autonomy, facilitate personalized feedback, and foster curiosity-driven inquiry (Alshammari, 2024; De La Puente et al., 2024). In particular, studies have noted its potential to stimulate critical engagement with content when pedagogically scaffolded (Chen, 2024; Furze et al., 2024).

However, this transformation is accompanied by significant pedagogical challenges. While AI tools can enhance metacognitive reflection and research competency (Abdelhalim, 2024; Chen, 2024), concerns persist regarding students' over-reliance on AI outputs, limited verification of content accuracy, and diminished higher-order thinking when AI use is unguided (Darwin et al., 2024; Liang & Wu, 2024; Teng, 2023). Furthermore, while learners appreciate the convenience and immediate support provided by AI, they often lack the critical skills needed to discern bias, inaccuracies, and ethical risks embedded in AI-generated content (Avsheniuk et al., 2024; Hu, 2025). Therefore, pedagogical models that promote strategic and ethical engagement with AI are urgently needed.

Among these challenges, ethical reasoning represents a particularly underdeveloped dimension in current pedagogical responses to AI. Learners need to be equipped not only to question content accuracy but also to recognize and respond to ethical issues such as unverified claims, biased outputs, or fabricated citations (Hu, 2025; Avsheniuk et al., 2024). These risks demand a deliberate instructional focus on transparency, academic integrity, and evaluative judgment in AI-mediated tasks. Accordingly, the present study integrates discussions on ethical awareness and responsible tool use as part of its intervention design.

To address both the cognitive and ethical dimensions of critical AI engagement, this study draws on experiential learning theory (Kolb, 1984) as a guiding framework. This theory offers a useful foundation for designing AI-integrated activities that encourage learners to move beyond passive consumption. By engaging learners in cycles of experience, reflection, conceptualization, and active experimentation, experiential approaches can cultivate deeper metacognitive awareness and critical thinking skills necessary for responsible AI use (Fullana et al., 2016; Daradoumis & Arguedas, 2020).

Despite a burgeoning interest in AI-assisted education, empirical evidence quantifying the effects of structured AI interventions on EFL learners' critical thinking remains scarce. Much of the existing research has relied on qualitative designs (Darwin et al., 2024; Liang & Wu, 2024) or self-reported perceptions (Abdelhalim, 2024; Almazrou et al., 2024), limiting objective assessment of cognitive outcomes. Recent experimental studies have demonstrated that AI integration can foster learner engagement (Huang & Teng, 2025), and higher-order thinking (Deng et al., 2025; Liu & Wang, 2024), yet there is a notable paucity of short-term, targeted interventions grounded in established pedagogical frameworks. Moreover, few studies have employed validated quantitative measures to assess changes in critical thinking following AI-focused interventions in EFL contexts (Yusuf et al., 2024; De La Puente et al., 2024).

Additionally, while some research highlights the role of metacognitive strategies and structured prompting in maximizing AI's educational benefits (Chen, 2024; Teng, 2025), practical models that operationalize these strategies within brief, scalable interventions are lacking. Particularly in resource-constrained environments like Vietnam, where digital literacy training remains emergent, there is an urgent need for accessible, effective pedagogical models that foster critical AI engagement (Furze et al., 2024).

This study aims to address these gaps by evaluating the effectiveness of a 90-minute workshop, the Critical AI Engagement Cycle, designed to enhance EFL learners' critical thinking skills in interacting

with ChatGPT. Grounded in experiential learning theory (Kolb, 1984) and critical thinking frameworks (Facione, 2011), the intervention provides scaffolded opportunities for learners to engage, reflect, analyze, and strategize AI use.

The study is guided by the following research question:

- To what extent does a structured workshop grounded in experiential learning enhance EFL learners' critical thinking skills in using ChatGPT, as measured by a validated critical thinking scale?

This study contributes to the evolving field of AI-assisted language education by offering empirical evidence on the impact of a structured, short-term intervention designed to foster critical thinking in EFL learners. By focusing on a theoretically grounded and practically scalable workshop, the study addresses key gaps in the literature related to the cognitive and ethical dimensions of AI use in language education. The findings have important implications for curriculum design, particularly in resource-constrained contexts where time-efficient and pedagogically sound approaches are urgently needed to equip learners with the critical competencies required to engage with AI tools responsibly and reflectively.

2 Literature Review

2.1 Generative AI and Its Role in EFL Education

Recent research on generative AI tools such as ChatGPT has moved beyond general debates over their benefits and risks, focusing instead on how pedagogical design shapes their effectiveness in EFL education. While early discourse emphasized AI's ability to personalize learning and offer real-time support, more recent studies highlight the importance of guided interaction, critical verification, and ethical reflection to avoid surface-level engagement and misinformation (Abdelhalim, 2024; Xu & Liu, 2025). This section reviews empirical evidence on how structured use of AI tools impacts EFL learners' cognitive, metacognitive, and affective outcomes.

Emerging research suggests that structured pedagogical interventions are crucial for realizing AI's educational potential. Abdelhalim (2024) found that EFL undergraduates' use of ChatGPT was significantly shaped by their metacognitive awareness, underscoring the need for explicit training. Similarly, Chen (2024) and Teng (2025) showed that AI tools can enhance reflective performance and writing self-efficacy when used with scaffolding and clear objectives. Strobl et al. (2024) demonstrated that using ChatGPT as a writing model in advanced L2 German classrooms promoted higher-order thinking during revision, as students critically evaluated AI-generated texts and improved their own writing. These findings highlight the need for structured pedagogical guidance to balance perceived usefulness with critical engagement.

Studies in the speaking and writing domains further demonstrate the value of AI tools when used reflectively. Hapsari and Wu (2022) found that AI chatbots reduced speaking anxiety and stimulated critical thinking during oral practice. Shen and Tao (2025) revealed that AI-based writing feedback reduced anxiety and improved metacognitive strategy use. Dizon et al. (2025) found that although university-level Japanese EFL students perceived ChatGPT as a useful tool for translation and summarization, they expressed concerns about over-reliance and the potential hindrance to authentic language development. However, these studies primarily examine specific skill domains or affective outcomes, often relying on self-reported data and lacking rigorous measurement of cognitive gains.

Despite promising results, several gaps remain. First, most prior research adopts qualitative designs or perception-based surveys (e.g., Darwin et al., 2024; Alshammari, 2024), with few studies employing validated instruments to objectively assess gains in critical thinking (see Deng et al., 2025; Liu & Wang, 2024). Second, ethical reasoning—an essential component of critical AI engagement—receives limited attention, despite emerging risks such as misinformation, bias, and over-reliance (Avsheniuk et al., 2024; Hu, 2025). Third, few interventions are brief, replicable, and theoretically grounded, making them

difficult to scale in resource-constrained EFL contexts (Yusuf et al., 2024; Furze et al., 2024). Fourth, comparative analyses across subdimensions of critical thinking remain rare, leaving open the question of which cognitive skills benefit most and why (Imjai et al., 2025). As Todd (2025) argued, generative AI may act as a disruptive force in language education, requiring reconceptualized teaching models that prioritize human oversight, critical thinking, and innovation.

In response, the present study proposes a short-term, structured intervention grounded in experiential learning theory (Kolb, 1984) and aligned with a multi-dimensional model of critical thinking (Facione, 2011; Imjai et al., 2025). Rather than focusing solely on academic performance or technical skills, the intervention aims to cultivate reflective judgment, ethical reasoning, and evaluative thinking through scaffolded interactions with ChatGPT. By doing so, this study contributes to current literature by offering a theoretically informed and practically feasible model of AI-integrated instruction for EFL learners.

2.2 Critical Thinking Development in EFL Contexts

Critical thinking is a core academic skill and a central outcome in modern EFL education, particularly in AI-mediated learning environments. Defined by Facione (2011, p. 27) as "purposeful, self-regulatory judgment" encompassing analysis, evaluation, inference, and open-mindedness, critical thinking underpins essential abilities such as academic literacy, autonomous learning, and the responsible use of information.

Recent scholarship has increasingly emphasized the importance of embedding critical thinking within practical English language skills, especially in contexts where generative AI tools are integrated into learning. For instance, Darwin et al. (2024) reported that EFL learners conceptualize critical thinking as a process of questioning, contextual analysis, and evidence-based reasoning—skills essential for academic reading and writing. In AI-supported writing contexts, Teng (2025) found that students with higher metacognitive awareness made more effective use of ChatGPT for feedback, showing improved self-efficacy and evaluative capacity. Similarly, Hapsari and Wu (2022) demonstrated that AI chatbot use reduced speaking anxiety while enhancing critical thinking in oral interaction. Almazrou et al. (2024) confirmed these trends in a broader context, finding that students perceived ChatGPT as beneficial in generating diverse perspectives and encouraging reflection. Yang et al. (2024) further observed that both students and teachers must navigate ChatGPT as a pervasive "ghostwriter," calling for stronger focus on ethical judgment and critical interpretation in academic writing classrooms. Collectively, these findings affirm that critical thinking can be cultivated through targeted English language tasks such as argumentative writing, text analysis, speaking practice, and source evaluation—particularly when mediated through structured AI use.

Systematic reviews further reinforce these findings. Wei and Li (2024) noted that interactive and constructive uses of AI—rather than passive consumption—are most effective in promoting learners' analytical and inferential reasoning. Shen and Teng (2024) identified a reciprocal relationship between AI-assisted writing, self-directed learning, and critical thinking, noting that these skills mutually reinforce one another. Teng's (2024) review emphasized that while ChatGPT supports writing development, its effective integration requires deliberate scaffolding to prevent cognitive dependency. Such findings suggest that AI tools, when embedded within purposeful and scaffolded pedagogical designs, can enhance rather than diminish students' cognitive engagement.

To guide both instructional design and assessment in this study, four key dimensions of critical thinking were adopted from Imjai et al. (2025): Analytical Skills, Logical Reasoning, Evidence Evaluation, and Open-Mindedness. These dimensions were selected due to their empirical grounding and alignment with the specific cognitive tasks posed by AI interaction in English learning. Analytical Skills refer to the capacity to break down complex input (e.g., AI-generated texts) and identify relevant connections. Logical Reasoning denotes the ability to make consistent, well-supported decisions based

on structured argumentation. Evidence Evaluation involves the judgment of information credibility, accuracy, and source validity—critical in contexts where AI may hallucinate or fabricate data. Open-Mindedness captures the willingness to consider alternative viewpoints and revise assumptions in light of new information, particularly relevant in learner-AI dialogue. These dimensions reflect both the cognitive and dispositional facets of critical thinking and are particularly salient in AI-mediated academic tasks.

Accordingly, the current study builds on this growing body of evidence by evaluating how a brief, structured and theoretically grounded intervention—guided by Kolb's experiential learning theory—can support the development of these four interrelated aspects of critical thinking. The intervention aimed not only to build cognitive skill but also to foster reflective awareness and ethical discernment in AI-supported English academic communication. As AI becomes increasingly embedded in language learning ecosystems, developing EFL students' critical capacities remains essential—not only for academic success but also for navigating the complex, evolving digital landscape with autonomy and responsibility.

2.3 Theoretical Framework: Experiential Learning as a Theoretical Foundation

Experiential learning theory (Kolb, 1984; Kolb & Kolb, 2009) provides a foundational lens for understanding how learners actively construct knowledge through iterative cycles of doing, reflecting, conceptualizing, and applying. Grounded in the work of Dewey, Piaget, and Lewin, experiential learning conceptualizes learning not as the passive absorption of information, but as a continuous process whereby knowledge emerges from the transformation of lived experience. This model offers a potentially valuable pedagogical framework in the context of emerging technologies, where learners are increasingly expected to navigate unfamiliar tools, assess ambiguous information, and develop the cognitive and ethical judgment required for autonomous learning.

Kolb's model is composed of four interconnected stages. The first, Concrete Experience, refers to direct engagement with a task or phenomenon that initiates the learning process. Learners encounter new content, perform an action, or interact with an environment that disrupts existing mental models and invites inquiry. The second stage, Reflective Observation, involves critical reflection on that experience—what occurred, what was observed, and what outcomes were unexpected. Learners begin to notice patterns, contradictions, or gaps in understanding. The third stage, Abstract Conceptualization, requires the learner to integrate those reflections into broader theoretical insights. This may involve revising assumptions, generating hypotheses, or articulating generalizable principles. Finally, in Active Experimentation, learners test their new understandings in novel contexts, applying strategies, adapting behavior, and producing new experiences that continue the cycle.

This recursive process supports the development of metacognitive awareness, self-regulated learning, and higher-order thinking skills. Kolb and Kolb (2009) argue that experiential learning fosters the ability to shift flexibly between modes of action and reflection—an essential capacity in today's complex, AI-mediated educational environments. Similarly, Fullana et al. (2016) found that reflective learning contributes not only to students' academic development, but also to their motivation and sense of self as learners. More recently, studies have applied this framework to digital contexts. For instance, Lin et al. (2025) demonstrated that experiential learning cycles enhanced reflective thinking in AI-supported STEM activities, while Hu (2025) and Ward et al. (2025) highlight how guided interaction, reflection, and ethical questioning can cultivate critical AI literacy in communication and multicultural education contexts.

Importantly, the experiential learning framework underscores the human-centered nature of critical engagement—particularly when learners interact with technologies that produce fluent yet fallible outputs. As Lewis and Sarkadi (2024) caution, while generative AI systems like ChatGPT can generate plausible responses, they do not possess reflective or ethical capacity. Thus, it is important that

pedagogical designs seek to scaffold learners' capacity to interrogate, verify, and ethically interpret AI-generated content. Experiential learning theory provides a structured yet flexible model for fostering these dispositions, enabling learners to build understanding not only of what AI produces, but of how and why those outputs should be critically evaluated.

In this study, experiential learning theory functions as both a conceptual foundation and a pedagogical rationale. It guided the design of the Critical AI Engagement Cycle, a brief, structured workshop aimed at enhancing EFL learners' critical thinking when interacting with ChatGPT. Rather than teaching technical AI use in isolation, the intervention sought to embed learners in cycles of inquiry, reflection, abstraction, and reapplication, thereby operationalizing experiential learning in the service of AI literacy. This theoretical foundation also informs the study's research questions and analysis, as we investigate whether experiential engagement can lead to measurable gains in critical thinking across cognitive and ethical domains. In doing so, the study contributes to broader efforts to align instructional design with both the affordances and limitations of generative AI in English language education.

2.4 Positioning the Present Study

Building on the limitations identified in recent literature, this study offers a practical model for enhancing EFL learners' critical thinking in AI-supported contexts. Unlike many existing interventions, which are long-term or exploratory in nature (e.g., Teng, 2025; Liu & Wang, 2024), the present research tests a brief, structured workshop—the Critical AI Engagement Cycle—designed to promote both cognitive and ethical engagement with AI.

Grounded in experiential learning theory, the workshop provides scaffolded opportunities for reflection, analysis, and application of AI tools in academic tasks. The study employs a validated scale (Imjai et al., 2025) to assess changes in four critical thinking dimensions and targets a resource-constrained Vietnamese EFL setting. In doing so, it contributes empirical evidence to ongoing conversations about scalability, measurement, and pedagogy in AI-mediated English education.

3 Methods

3.1 Participants

The study was conducted at a public university in Vietnam's Central Highlands, where English as a Foreign Language (EFL) programs have begun integrating generative AI tools such as ChatGPT. A total of 72 students (38 in the experimental group; 34 in the control group) participated. Participants were enrolled in undergraduate and graduate programs in English Language Studies and Applied Linguistics. Their English proficiency ranged from upper-intermediate to advanced, based on institutional placement tests and successful completion of English-medium academic coursework. This background helps ensure that participants had sufficient linguistic competence to engage critically with AI-generated English content.

Convenience sampling was employed due to logistical constraints, and participants were selected based on course enrollment and willingness to participate. The experimental and control groups were drawn from two intact class sections within the same program but taught separately, with no overlapping instruction, group work, or scheduled interaction. Both groups were instructed by the same teacher using the same syllabus, ensuring consistency in instructional delivery across groups.

The intervention was implemented as a brief, 90-minute workshop delivered only to the experimental group, with the posttest administered immediately afterward. This design reduced the likelihood of contamination, as students had limited opportunity to discuss the intervention content across groups before data collection concluded. In-class discussions were confined to each section, and participants were not informed of the study conditions assigned to the other group. Baseline characteristics such

as age, gender, and AI familiarity were similar across groups, and descriptive statistics confirmed their general comparability, supporting the internal validity of between-group comparisons.

3.2 Intervention

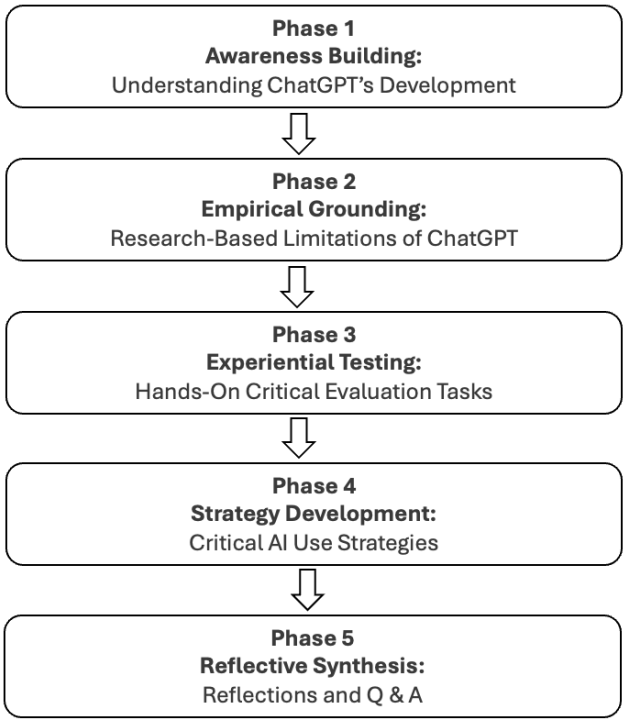
A 90-minute workshop titled *Developing Critical Thinking Usage of Generative AI* was delivered to the experimental group. The aim of the workshop was to raise participants' awareness of both the affordances and limitations of ChatGPT and to cultivate their ability to use the tool critically, reflectively, and ethically in academic contexts. The workshop design was informed by constructivist, experiential, and reflective learning theories (Kolb, 1984; Schön, 2017), emphasizing learning through active experimentation, dialogic interaction, critical evaluation, and metacognitive self-awareness. These principles were embedded in both the overall structure of the workshop and the nature of the tasks, which systematically encouraged students to test assumptions, scrutinize outputs, and reflect on responsible AI use.

The workshop was developed based on a structured instructional model that served as a practical lesson plan. Its content and sequencing were informed by experiential learning theory and empirical studies on AI literacy in EFL education. Prior to implementation, the intervention was reviewed by two experts in TESOL and instructional design to ensure theoretical coherence and pedagogical clarity. Although no formal pilot test was conducted, feedback from these reviewers was used to refine the timing, complexity, and flow of tasks.

The intervention followed a five-phase instructional sequence, conceptualized as the Critical AI Engagement Cycle, which scaffolded the learning experience around progressively deeper engagement with ChatGPT. Students engaged primarily in individual tasks supplemented by guided peer discussions to foster collaborative reflection. A visual overview of the Critical AI Engagement Cycle is provided to illustrate the pedagogical structure (see Figure 1).

Figure 1

The Critical AI Engagement Cycle: A Five-Phase Structure for Guiding Critical and Reflective Use of Generative AI



In the first phase, Awareness-Building, students were introduced to ChatGPT's development process and inherent limitations, drawing on documentation from OpenAI that described reinforcement learning processes, fine-tuning procedures, and common issues such as factual hallucinations and response variability. In the second phase, Empirical Grounding, participants engaged with synthesized findings from recent empirical research (Cong-Lem et al., 2024), which highlighted ChatGPT's fabrication of references, factual inaccuracies, inconsistent domain-specific performance, and ethical risks in academic settings.

The third phase, Experiential Testing, involved a series of guided, hands-on tasks requiring participants to directly evaluate ChatGPT's outputs. Activities included verifying AI-generated references, evaluating ChatGPT's accuracy in summarizing scholarly research, assessing its performance on logical reasoning prompts, and evaluating its mathematical accuracy. For example, in one task, students prompted ChatGPT to generate a paragraph about a common language learning theory and then asked it to list references with DOI links. They then attempted to locate these references via Google Scholar, discovering that several were fabricated. In another activity, students crafted prompts based on real studies (e.g., "Tell me about the findings of the 2025 study by Yu on vocabulary acquisition") and evaluated whether ChatGPT could identify the correct content. Additional exercises involved asking ChatGPT to compose a 25-word sentence and checking for precision in word count, or prompting it with a logic problem involving multiple possible causes of failure to assess the AI's reasoning consistency. Students were also instructed to challenge ChatGPT's responses—e.g., by refuting its initial suggestion on the best method for vocabulary learning—and observe whether it maintained or revised its position. Through these exercises, students developed firsthand awareness of the model's cognitive strengths and limitations. In the fourth phase, Strategy Development, students received explicit instruction and practice in refining prompts, validating information sources, and treating ChatGPT's responses as preliminary starting points for further human inquiry rather than as definitive answers.

Finally, in the fifth phase, Reflective Synthesis, students participated in a concluding instructor-facilitated discussion and completed a written reflection activity. They were encouraged to articulate their evolving perspectives on the responsible use of AI tools in academic work, focusing on critical judgment, recognition of AI's limitations, and ethical considerations. The instructor served as a facilitator and critical thinking coach throughout, guiding students' reflections without prescribing answers.

Approximately 15 to 20 minutes were allocated to each phase to ensure structured progression within the 90-minute session. The workshop was conducted in a computer-equipped classroom to enable immediate interaction with ChatGPT during activities. This design ensured that students moved through Kolb's experiential learning cycle—concrete experience, reflective observation, abstract conceptualization, and active experimentation—while building competencies across analytical reasoning, evidence evaluation, logical inference, and open-mindedness.

3.3 Instruments

Participants completed a two-part survey administered pre- and post-intervention. Part 1 collected demographic information (age, gender, AI use frequency, prior AI knowledge, and self-reported AI skill level). Part 2 consisted of a 12-item critical thinking scale adapted from Imjai et al. (2025), assessing four subdomains: Analytical Skills (AS), Logical Reasoning (LR), Evidence Evaluation (EE), and Open-Mindedness (OM). Minor contextual adjustments were made to reflect the study's focus on AI-assisted English learning, such as replacing general terms (e.g., "data" or "information") with references to "AI-generated content" or "tools like ChatGPT."

These surface-level wording changes preserved the conceptual integrity and original structure of each subscale. For example, the item "You consistently analyse complex data…" was revised to "I consistently analyze complex AI-generated content…" to situate the statement within an educational

AI context. To improve clarity and balance across dimensions, one additional item was added to each subscale, resulting in three items per dimension.

Each item was rated on a 5-point Likert scale (1 = Strongly Disagree; 5 = Strongly Agree). Internal consistency reliability was high, with Cronbach's alpha coefficients of .90 (pretest) and .85 (posttest) in this study. The slight decrease in posttest alpha likely reflects natural variation in participants' response patterns following the intervention, and both values remain within the range indicating strong reliability.

3.4 Data Collection and Analysis

Pre- and post-tests were administered in controlled classroom settings immediately before and after the intervention. No missing data were reported; all participants completed both assessments. Descriptive statistics (means, standard deviations) were calculated for all critical thinking measures. Paired-samples t-tests or Wilcoxon signed-rank tests were employed to analyze within-group changes, depending on data normality. Cohen's d was computed to evaluate effect sizes. Between-group comparisons at post-test were analyzed using Analysis of Covariance (ANCOVA), controlling for pretest scores and covariates (gender, AI use frequency, AI skill level, and prior AI knowledge). Homogeneity of regression slopes and normality of residuals were tested to validate ANCOVA assumptions. All analyses were conducted using R version 4.4.2, with a significance threshold set at $p < .05$.

3.5 Ethical Considerations

The study received approval from the president of the university where the participants were enrolled. Informed consent was obtained from all participants. Participation was voluntary, with the option to withdraw at any time without penalty. Data were anonymized using unique participant codes. Beyond procedural ethics, the workshop explicitly emphasized responsible AI use, encouraging students to develop critical thinking not only for cognitive advancement but also for ethical engagement with AI-generated content, addressing concerns about misinformation, bias, and academic integrity.

4 Results

This section presents the findings related to the impact of the Critical AI Engagement Cycle workshop on EFL learners' critical thinking skills. Descriptive statistics, inferential analyses (within- and between-group comparisons), pretest equivalence testing, and ANCOVA results are reported systematically to address the research question: To what extent does a structured workshop grounded in experiential learning enhance EFL learners' critical thinking skills in using ChatGPT, as measured by a validated critical thinking scale?

4.1 Participant Background Characteristics

Table 1 presents descriptive statistics for participant background variables. The sample consisted of 72 participants (experimental group: n = 38; control group: n = 34), with a mean age of 22 years (SD = 4.28). The majority of participants were female (83.3%). Regarding AI familiarity, 56.9% of participants reported moderate to high AI use frequency. Mean scores for AI skill level and prior knowledge were comparable across groups, indicating similar levels of technological readiness at baseline. To account for potential confounding effects, all background variables—including gender, AI use frequency, prior AI knowledge, and self-rated AI skill level—were included as covariates in subsequent ANCOVA analyses.

Table 1

Participant Background Characteristics

Variable	Experimental (n = 38)	Control (n = 34)	Total (N = 72)
Age (Mean ± SD)	22.1 ± 4.2	21.9 ± 4.4	22.0 ± 4.3
Gender (Female %)	84.20%	82.40%	83.30%
AI Use Frequency (Mod-High %)	58%	55%	56.90%
AI Skill Level (Mean ± SD)	3.45 ± 0.62	3.41 ± 0.59	3.43 ± 0.61
Prior Knowledge (Mean ± SD)	3.39 ± 0.67	3.37 ± 0.65	3.38 ± 0.66

4.2 Pretest Equivalence Testing

To examine baseline equivalence between groups, Mann-Whitney U tests were conducted comparing pretest critical thinking scores between the experimental and control groups. No significant differences were found for Analytical Skills ($U = 680$, $p = .702$), Logical Reasoning ($U = 637$, $p = .923$), Evidence Evaluation ($U = 714$, $p = .440$), Open-Mindedness ($U = 660.5$, $p = .872$), or Overall Critical Thinking ($U = 664.5$, $p = .839$). These results suggest that the two groups were statistically comparable at the outset of the study.

4.3 Critical Thinking Scores: Descriptive Statistics

As shown in Table 2, the experimental group exhibited notable improvements across all measures of critical thinking. Overall Critical Thinking (Overall_CT) scores increased from a pretest mean of 3.33 (SD = 0.69) to a posttest mean of 4.02 (SD = 0.37). Similarly, substantial gains were observed across the four critical thinking subdomains, particularly in Evidence Evaluation (EE) and Logical Reasoning (LR). In contrast, the control group showed smaller gains across all measures, with Overall_CT increasing from 3.29 (SD = 0.80) to 3.61 (SD = 0.63).

Table 2

Descriptive Statistics for Critical Thinking Scores

Measure	Group	Pretest Mean (SD)	Posttest Mean (SD)
Overall Critical Thinking	Experimental	3.33 (0.69)	4.02 (0.37)
	Control	3.29 (0.80)	3.61 (0.63)
Analytical Skills	Experimental	3.27 (0.73)	3.87 (0.49)
	Control	3.25 (0.75)	3.50 (0.69)
Logical Reasoning	Experimental	3.33 (0.78)	4.07 (0.45)
	Control	3.29 (0.98)	3.64 (0.70)
Evidence Evaluation	Experimental	3.30 (0.79)	4.07 (0.49)
	Control	3.23 (0.86)	3.56 (0.86)
Open-Mindedness	Experimental	3.40 (0.81)	4.07 (0.47)
	Control	3.38 (0.82)	3.75 (0.73)

4.4 Within-Group Pretest and Posttest Comparisons

To examine the workshop's impact within the experimental group, paired-samples tests were conducted. Overall Critical Thinking significantly improved, as indicated by the Wilcoxon signed-rank test ($W = 19$, $p < .001$, Cohen's $d = 1.23$), reflecting a large effect size. Analytical Skills (AS) also showed a significant increase ($W = 37$, $p < .001$, $d = 0.94$). Logical Reasoning (LR) improved significantly, with a paired t-test yielding $t(37) = -6.07$, $p < .001$, $d = 1.11$. Evidence Evaluation (EE) exhibited significant gains as well ($t(37) = -4.78$, $p < .001$, $d = 1.19$). Lastly, Open-Mindedness (OM) significantly improved ($W = 45$, $p < .001$, $d = 1.00$). In comparison, the control group showed modest improvements with smaller effect sizes, suggesting limited development without the intervention.

4.5 Between-Group Comparisons at Posttest

Further analyses using Mann-Whitney U tests were conducted to compare posttest critical thinking scores between the experimental and control groups. Table 3 summarizes the results.

Table 3
Between-Group Posttest Comparisons (Mann-Whitney U Tests)

Measure	U Statistic	p-value	Cohen's d
Overall Critical Thinking	384	.003	0.80
Analytical Skills	443	.018	0.62
Logical Reasoning	379.5	.002	0.74
Evidence Evaluation	407.5	.005	0.75
Open-Mindedness	483	.057	0.52

The experimental group significantly outperformed the control group on Overall Critical Thinking, Analytical Skills, Logical Reasoning, and Evidence Evaluation. Differences in Open-Mindedness approached significance.

4.6 ANCOVA Results

To further validate the effects of the workshop, Analysis of Covariance (ANCOVA) was conducted. ANCOVA is a statistical technique that compares posttest scores between groups while controlling for potential confounding variables known as covariates. Tests of homogeneity of regression slopes and normality of residuals confirmed that ANCOVA assumptions were not violated. In this study, covariates included pretest critical thinking scores, gender, AI use frequency, AI skill level, and prior knowledge related to AI. Controlling for these variables ensures that differences observed at posttest are more confidently attributed to the intervention rather than pre-existing group differences.

Table 4
ANCOVA Results Controlling for Covariates

Measure	F (1, 62)	p-value	Partial η^2
Overall Critical Thinking	17.60	<.001	.22
Analytical Skills	9.07	.004	.13
Logical Reasoning	12.50	.001	.17
Evidence Evaluation	11.74	.001	.16
Open-Mindedness	3.68	.060	.06

The results confirm significant differences favoring the experimental group across Overall Critical Thinking and three critical thinking subdomains after controlling for baseline scores and covariates.

In summary, the findings provide robust evidence that the 90-minute Critical AI Engagement Cycle workshop significantly enhanced EFL learners' critical thinking skills when using ChatGPT. Improvements were consistent across Overall Critical Thinking and its subdomains, particularly in Evidence Evaluation and Logical Reasoning, supporting the efficacy of short-term, structured interventions grounded in experiential learning theory for fostering critical AI literacy.

5 Discussion

5.1 Summary and Interpretations of Key Findings

This study investigated the effectiveness of a 90-minute Critical AI Engagement Cycle workshop in enhancing EFL learners' critical thinking skills when interacting with ChatGPT. The results demonstrated that the experimental group achieved significant improvements across Overall Critical Thinking and four subdomains (Analytical Skills, Logical Reasoning, Evidence Evaluation, and Open-Mindedness), whereas the control group showed only modest gains. ANCOVA analyses, controlling for pretest scores, gender, AI use frequency, AI skill level, and prior knowledge, supported the robustness of these findings.

Several features of the intervention may help explain its effectiveness. Grounded in experiential learning theory, the design guided learners through a sequence of active engagement, reflection, conceptual analysis, and experimentation. This structure appears to have supported deeper metacognitive processing than passive exposure to AI tools. Participants critically examined AI-generated outputs, identified inconsistencies, and practiced verification strategies—behaviors aligned with higher-order thinking development.

Modest improvements were also observed in the control group across several critical thinking subdomains. These gains, while smaller than those of the experimental group, could be attributed to general cognitive development during the academic semester, increased test familiarity, or indirect exposure to critical thinking tasks through regular coursework. However, as confirmed by ANCOVA results, the improvements in the experimental group remained significantly greater even after controlling for pretest scores and background variables.

The differential improvements observed across the four critical thinking subdomains can be understood in light of existing empirical evidence and the design features of the intervention. Specifically, Evidence Evaluation and Logical Reasoning showed the greatest gains, which may be attributed to the hands-on, task-based activities that explicitly required students to verify AI-generated outputs and challenge flawed logic in ChatGPT's responses. This corresponds with findings from Abdelhalim (2024) and Teng (2025), who emphasize the importance of metacognitive awareness and evaluative reasoning in AI-mediated learning. Similarly, Almazrou et al. (2024) reported that students perceived ChatGPT as especially beneficial for generating diverse perspectives and prompting analytical thinking, especially when guided by structured tasks. In contrast, more dispositional subdomains like Open-Mindedness may require repeated practice and deeper reflective engagement to see substantial change (Darwin et al., 2024; Fullana et al., 2016). Thus, the results suggest that experiential, feedback-rich interactions with AI can differentially foster the development of specific critical thinking facets, depending on how they are scaffolded in instructional design.

5.2 Comparison with Previous Literature

The findings are broadly consistent with previous studies highlighting the positive influence of structured AI use on critical thinking development in EFL contexts (Abdelhalim, 2024; Deng et al., 2025; Liu &

Wang, 2024). The substantial gains, particularly in Evidence Evaluation and Logical Reasoning, reinforce the importance of embedding experiential cycles of reflection, testing, and strategy development into AI-integrated learning environments. This study extends prior research by demonstrating that even a single-session intervention, grounded in experiential learning theory (Kolb, 1984), can produce meaningful cognitive outcomes when pedagogically scaffolded.

However, these results contrast with findings by Liang and Wu (2024), who reported limited critical thinking gains despite ChatGPT use. One plausible explanation lies in the design of the intervention. Unlike unguided AI exposure, the Critical AI Engagement Cycle deliberately operationalized experiential learning stages. Students engaged in concrete interaction (generating outputs), reflective observation (identifying limitations), abstract conceptualization (developing theories about AI behavior), and active experimentation (strategizing AI use). This structured experiential design likely enabled deeper and more sustained critical engagement.

Moreover, the present findings resonate with broader research emphasizing strategic and ethical AI literacy development in EFL education (Chen, 2024; Darwin et al., 2024). Students learned not only to identify logical fallacies or evaluate evidence but also to question AI outputs' credibility and reliability—skills essential for navigating English academic communication contexts.

Compared to more traditional expository or lecture-based instruction, which often emphasizes procedural use of AI tools or passive knowledge transfer, the experiential learning approach used in this study actively engaged learners in questioning, evaluating, and experimenting with AI-generated outputs. This pedagogical contrast may explain the observed gains: rather than being shown how ChatGPT functions, participants had to uncover its limitations through guided discovery and reflection. This active involvement, supported by structured prompts and peer discussion, aligns with previous research highlighting the superiority of experiential models in fostering deep learning and metacognitive skills (Kolb & Kolb, 2009; Lin et al., 2025).

This study offers several contributions to the field of EFL education. First, it demonstrates the potential for critical thinking gains to be achieved through a brief, targeted, and resource-efficient intervention, making such initiatives highly promising and scalable for resource-constrained settings. Second, by integrating covariates such as AI skill level and prior knowledge into ANCOVA analyses, the study addresses methodological gaps noted in earlier AI-in-EFL studies (e.g., Hapsari & Wu, 2022; Teng, 2025). Third, by proposing a conceptual model that views AI literacy as encompassing cognitive, ethical, and strategic dimensions, the study advances understanding of how to critically engage EFL learners with emerging technologies.

The outcomes of this study resonate strongly with the objectives of the United Nations Sustainable Development Goals, notably SDG 4: Quality Education and SDG 9: Industry, Innovation, and Infrastructure (United Nations, n.d.). SDG 4 emphasizes the provision of inclusive and equitable quality education and the promotion of lifelong learning opportunities for all. By equipping EFL learners with critical thinking skills and AI literacy, the intervention contributes to the transformation of education systems to address the emerging challenges and opportunities of the generative AI era. This aligns with UNESCO's advocacy for integrating AI competencies into education to foster human-centered and ethical use of technology.

Simultaneously, the study contributes to SDG 9 by fostering innovation in educational practices through the integration of AI tools like ChatGPT. By developing a structured framework—the Critical AI Engagement Cycle—the research promotes sustainable industrialization and innovation within the educational sector. This approach not only enhances the technological capabilities of learners but also encourages the development of infrastructure that supports innovative teaching and learning methodologies. Such initiatives are crucial for building resilient educational systems that can adapt to technological advancements and prepare students for the demands of the modern workforce.

5.3 Implications for English Language Teaching, Learning, and Policy

The findings of this study offer significant implications for English language teaching, learning, and educational policy. Beyond fostering general critical thinking, the Critical AI Engagement Cycle suggests potential contributions to EFL education in several ways. First, it may support the development of academic literacy skills fundamental to EFL success, including source evaluation, evidence-based argumentation, and critical academic writing (Abdelhalim, 2024). Encouraging students to verify information, identify bias, and critically assess AI-generated texts can strengthen their reading comprehension and analytical writing abilities in English.

Second, the intervention may promote critical language awareness (Darwin et al., 2024), sensitizing learners to discourse patterns, rhetorical inconsistencies, and persuasive techniques embedded in AI-generated language. This metalinguistic sensitivity is particularly important for advanced EFL learners navigating increasingly complex English academic texts. Third, by emphasizing strategic prompting and critical verification, the workshop may foster greater learner autonomy (Xu & Liu, 2025), encouraging students to engage with AI tools in a self-directed and critical manner.

Furthermore, by raising awareness of ethical considerations, the intervention encourages more responsible language practices among students, addressing growing concerns about plagiarism and academic integrity in AI-assisted writing (Shen & Tao, 2025). Developing ethical AI literacy alongside critical thinking skills is essential for preparing EFL learners to participate responsibly in AI-mediated academic environments.

In addition, this study contributes to the global call for evidence-informed and sustainable education in the generative AI era (Baskara, 2023; Corbeil & Corbeil, 2025). Rather than relying on speculative claims or anecdotal enthusiasm surrounding generative AI, the intervention was explicitly grounded in theoretical frameworks and supported by empirical validation. This alignment between pedagogical design, cognitive theory, and assessment underscores the importance of basing instructional decisions on rigorous, context-sensitive research. As AI tools continue to be integrated into English language education, adopting an evidence-informed approach can help educators avoid premature implementation, ensure ethical standards, and tailor interventions to specific learner needs. The present study serves as an example of how short-term, resource-efficient programs can be both theoretically grounded and empirically tested, supporting scalable innovation in AI-assisted EFL contexts.

For educators, the Critical AI Engagement Cycle provides a practical framework for integrating critical thinking and AI literacy into English language instruction without requiring major curricular overhauls. It can be flexibly implemented through targeted classroom tasks that prompt students to critically engage with AI outputs. For instance, learners might evaluate the consistency of AI-generated explanations across different prompts, detect logical fallacies or unsupported claims in model responses, or challenge ChatGPT's suggestions using counterexamples. These activities cultivate habits of questioning, verification, and evidence-based reasoning that are transferable across academic contexts.

Modular workshops, embedded within writing, reading, or research skills courses, can meaningfully enhance students' critical and communicative competencies while adapting flexibly to different institutional contexts. Such integrations require minimal technological infrastructure and can be facilitated using guided prompts, peer critique, and reflective tasks—tools already familiar to language educators.

For policymakers and curriculum designers, these findings highlight the urgency of incorporating AI literacy components into EFL programs. Equipping students with the cognitive, ethical, and strategic skills needed to engage with AI-generated information is vital for fostering informed, ethical, and autonomous English language users. Moreover, the success of a short, focused intervention suggests that even minimal curriculum adaptations can yield substantial benefits. Resource-limited institutions may particularly benefit from implementing scalable, experiential models like the Critical AI Engagement

Cycle to build 21st-century skills among EFL learners without extensive financial or technological investments.

5.4 Limitations and Future Research Directions

Despite its contributions, the study has several limitations. First, the quasi-experimental design, with participants drawn from intact classes rather than randomized groups, may introduce selection biases that limit causal inference. However, baseline characteristics such as age, gender, AI familiarity, and pretest scores were comparable across groups, and no significant pre-intervention differences were found. To further minimize potential confounding effects, these variables were included as covariates in the ANCOVA analysis. Although random assignment would provide stronger internal validity, these design and statistical safeguards help mitigate the impact of selection bias. Second, the sample was restricted to one Vietnamese university, limiting generalizability to broader EFL populations. Future research should replicate this intervention across diverse educational contexts, including secondary and international EFL settings.

Third, the study measured critical thinking immediately post-intervention; longitudinal studies are needed to assess the durability of gains over time. Future research could also explore how critical thinking skills transfer to academic writing, reading comprehension, and oral communication tasks in English. Comparative studies examining single-session versus multi-session interventions, or experiential versus expository AI literacy approaches, would deepen understanding of effective pedagogical models.

Finally, although the study used a validated critical thinking scale adapted for AI-assisted EFL contexts, relying solely on self-reported measures may introduce biases such as social desirability or overestimation of ability. Future research should consider supplementing survey data with additional sources such as reflective journals, classroom observations, or performance-based assessments. Triangulating these data sources would offer richer insights into how learners apply critical thinking skills in real-world AI-mediated academic tasks and enhance the validity of findings.

6 Conclusion

This study provides compelling evidence that a short, structured experiential learning intervention can significantly enhance EFL learners' critical thinking skills when interacting with generative AI tools such as ChatGPT. By operationalizing Kolb's experiential learning cycle into the Critical AI Engagement Cycle workshop, learners were guided through a systematic process of exploration, reflection, conceptualization, and strategic experimentation. The intervention not only improved students' analytical reasoning, evidence evaluation, logical thinking, and open-mindedness but also emphasized ethical and strategic engagement with AI outputs.

Importantly, the results underscore that critical thinking development in AI-enhanced environments does not necessarily require long, resource-intensive programs. Even a 90-minute focused session, when thoughtfully designed, can yield substantial cognitive benefits. These findings have significant implications for AI literacy education, particularly in resource-constrained EFL contexts where time and institutional capacity for curricular innovation may be limited.

Nevertheless, further research is necessary to explore the long-term sustainability of these gains, their transferability to real-world academic tasks, and their adaptability across diverse cultural and linguistic settings. Future work could also refine the experiential model, incorporating students' reflections and emotional responses to better scaffold critical engagement with emerging technologies.

Overall, the Critical AI Engagement Cycle offers a practical, scalable, and theoretically grounded pathway for cultivating discerning, reflective learners capable of navigating the complexities of the AI era responsibly and ethically.

Acknowledgements

The authors would like to express their sincere gratitude to all the students who participated in this study for their time and engagement. Special thanks are extended to Ms. Nguyen Hoang Nhat Quyen for her valuable assistance during the early stages of the project. We also appreciate the constructive feedback provided by the editor and the anonymous reviewers, which greatly contributed to improving the clarity and quality of this manuscript.

Funding Statement

This research is funded by the Foundation for Science and Technology Development of Dalat University. Grant No: 1153/QĐ-ĐHDL.

Ethics Approval Statement

All procedures performed in the study were in accordance with the ethical standards of the Institutional and/or National Research Committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Declaration of AI Use

The authors used generative AI technology (ChatGPT by OpenAI) to support the proofreading and language refinement of this manuscript. The use of AI was conducted under human oversight, and the final manuscript was thoroughly reviewed and edited to ensure accuracy and integrity. The author remains fully responsible and accountable for the content and conclusions presented in this work.

Appendix A. Critical Thinking Instrument

The following items were developed based on the four critical thinking dimensions adapted from Imjai et al. (2025): Analytical Skills, Logical Reasoning, Evidence Evaluation, and Open-Mindedness. Participants rated each statement on a 5-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). Vietnamese translations are provided in parentheses for participant accessibility.

Analytical Skills

AS1. I can identify the main points and supporting ideas in AI-generated content.
(Tôi có thể xác định các luận điểm chính và các ý hỗ trợ trong nội dung do AI tạo ra.)
AS2. I am able to break down AI-generated information into smaller parts to better understand it.
(Tôi có thể phân tích nội dung do AI tạo ra thành các phần nhỏ hơn để hiểu rõ hơn.)
AS3. I can distinguish between factual content and opinions in AI-generated responses.
(Tôi có thể phân biệt giữa thông tin thực tế và ý kiến trong phản hồi do AI tạo ra.)

Logical Reasoning

LR1. I use logical reasoning when deciding how much to trust AI-generated suggestions.
(Tôi sử dụng lập luận logic khi quyết định mức độ tin tưởng vào các đề xuất do AI tạo ra.)

LR2. I make sure that my conclusions are not solely influenced by AI outputs.

(Tôi đảm bảo rằng quyết định của tôi không chỉ bị ảnh hưởng bởi các kết quả của AI.)

LR3. I can detect inconsistencies or contradictions in AI-generated content and adjust my conclusions accordingly.

(Tôi có thể phát hiện sự không nhất quán hoặc mâu thuẫn trong nội dung do AI tạo ra và điều chỉnh kết luận của mình cho phù hợp.)

Evidence Evaluation

EE1. I assess the credibility of AI-generated information by cross-referencing it with reliable sources. (Tôi đánh giá độ tin cậy của thông tin do AI tạo ra bằng cách đối chiếu với các nguồn đáng tin cậy.)

EE2. I verify the accuracy and source of AI-generated content before applying it in my teaching or learning tasks.

(Tôi kiểm tra độ chính xác và nguồn gốc của nội dung do AI tạo ra trước khi áp dụng vào công việc giảng dạy hoặc học tập của mình.)

EE3. I critically evaluate whether AI-generated content is supported by sufficient evidence or data. (Tôi đánh giá một cách khoa học liệu nội dung do AI tạo ra có được hỗ trợ bởi đủ bằng chứng hoặc dữ liệu hay không.)

Open-Mindedness

OM1. I consider and respect perspectives that differ from AI-generated suggestions, integrating both AI and non-AI insights into my decision-making.

(Tôi xem xét và tôn trọng các quan điểm khác với các đề xuất do AI đưa ra, kết hợp cả những hiểu biết từ AI và không phải AI vào quá trình ra quyết định của mình.)

OM2. I am willing to experiment with new ways to critically assess and use AI-generated content in my academic work.

(Tôi sẵn sàng thử nghiệm các cách thức mới để đánh giá một cách phản biện và sử dụng nội dung do AI tạo ra trong công việc học tập của mình.)

OM3. I remain open to revising my ideas when AI-generated information presents new insights or challenges my assumptions.

(Tôi luôn cởi mở trong việc điều chỉnh ý tưởng của mình khi thông tin do AI tạo ra đưa ra những hiểu biết mới hoặc thách thức các giả định của tôi.)

References

Abdelhalim, S. (2024). Using ChatGPT to promote research competency: English as a Foreign Language undergraduates' perceptions and practices across varied metacognitive awareness levels. *Journal of Computer Assisted Learning*, *40*(3), 1261–1275. https://doi.org/10.1111/jcal.12948

Almazrou, S., Alanezi, F., Almutairi, S. A., AboAlsamh, H. M., Alsedrah, I. T., Arif, W. M., Alsadhan, A. A., AlSanad, D. S., Alqahtani, N. S., AlShammary, M. H., Bakhshwain, A. M., Almuhanna, A. F., Almulhem, M., Alnaim, N., Albelali, S., & Attar, R. W. (2024). Enhancing medical students critical thinking skills through ChatGPT: An empirical study with medical students. *Nutrition and Health*, 02601060241273627. https://doi.org/10.1177/02601060241273627

Alshammari, J. (2024). Revolutionizing EFL learning through ChatGPT: A qualitative study. *Revista Amazonia Investiga*, *13*(82), 208–221. https://doi.org/10.34069/AI/2024.82.10.17

Avsheniuk, N., Lutsenko, O., Svyrydiuk, T., & Seminikhyna, N. (2024). Empowering language learners' critical thinking: Evaluating ChatGPT's role in english course implementation. *Arab World English Journal*, 210–224. https://doi.org/10.24093/awej/ChatGPT.14

Baskara, F. R. (2023). Integrating ChatGPT into EFL writing instruction: Benefits and challenges. *International Journal of Education and Learning, 5*(1), 44–55. https://doi.org/10.31763/ijele.v5i1.858

Chen, K., Tallant, A. C., & Selig, I. (2024). Exploring generative AI literacy in higher education: Student adoption, interaction, evaluation and ethical perceptions. *Information and Learning Sciences*. https://doi.org/10.1108/ILS-10-2023-0160

Cong-Lem, N., Soyoof, A., & Tsering, D. (2024). A systematic review of the limitations and associated opportunities of ChatGPT. *International Journal of Human–Computer Interaction*, 1–16. https://doi.org/10.1080/10447318.2024.2344142

Corbeil, J. R., & Corbeil, M. E. (2025). *Teaching and learning in the age of generative AI: evidence-based approaches to pedagogy, ethics, and beyond*. Taylor & Francis.

Daradoumis, T., & Arguedas, M. (2020). Cultivating students' reflective learning in metacognitive activities through an affective pedagogical agent. *Educational Technology & Society*, *23*(2), 19–31. JSTOR.

Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E. D., & Marzuki. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, *11*(1), 2290342. https://doi.org/10.1080/2331186X.2023.2290342

De La Puente, M., Torres, J., Troncoso, A. L. B., Meza, Y. Y. H., & Carrascal, J. X. M. (2024). Investigating the use of ChatGPT as a tool for enhancing critical thinking and argumentation skills in international relations debates among undergraduate students. *Smart Learning Environments*, *11*(1), 55. https://doi.org/10.1186/s40561-024-00347-0

Deng, R., Jiang, M., Yu, X., Lu, Y., & 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

Dizon, G., Gold, J., & Barnes, R. (2025). ChatGPT for self-regulated language learning: University English as a foreign language students' practices and perceptions. *Digital Applied Linguistics*, *3*, 102510. https://doi.org/10.29140/dal.v3.102510

Facione, P. A. (2011). Critical thinking: What it is and why it counts. *Insight Assessment*, *1*(1), 1–23.

Fullana, J., Pallisera, M., Colomer, J., Fernández Peña, R., & Pérez-Burriel, M. (2016). Reflective learning in higher education: A qualitative study on students' perceptions. *Studies in Higher Education*, *41*(6), 1008–1022. https://doi.org/10.1080/03075079.2014.950563

Furze, L., Perkins, M., Roe, J., & MacVaugh, J. (2024). The AI assessment scale (AIAS) in action: A pilot implementation of GenAI-supported assessment. *Australasian Journal of Educational Technology*, *40*(4), 38–55. Scopus. https://doi.org/10.14742/ajet.9434

Hapsari, I. P., & Wu, T.-T. (2022). *AI chatbots learning model in english speaking skill: Alleviating speaking anxiety, boosting enjoyment, and fostering critical thinking*. *13449 LNCS*, 444–453. Scopus. https://doi.org/10.1007/978-3-031-15273-3_49

Hu, Y. (2025). Generative AI, communication, and stereotypes: Learning critical AI literacy through experience, analysis, and reflection. *Communication Teacher*, *39*(1), 6–12.

Huang, J., & Teng, M. F. (2025). Peer feedback and ChatGPT-generated feedback on Japanese EFL students' engagement in a foreign language writing context. *Digital Applied Linguistics, 2,* 102469. https://doi.org/10.29140/dal.v2.102469

Imjai, N., Yordudom, T., Yaacob, Z., Saad, N. H. M., & Aujirapongpan, S. (2025). Impact of AI literacy and adaptability on financial analyst skills among prospective Thai accountants: The role of critical thinking. *Technological Forecasting and Social Change*, *210*, 123889. https://doi.org/10.1016/j.techfore.2024.123889

Kolb, A. Y., & Kolb, D. A. (2009). The learning way: Meta-cognitive aspects of experiential learning. *Simulation & Gaming*, *40*(3), 297–327. https://doi.org/10.1177/1046878108325713

Kolb, D. (1984). Experiential learning: Experience as the source of learning and development. In *Journal of Business Ethics* (Vol. 1).

Lee, H.-Y., Chen, P.-H., Wang, W.-S., Huang, Y.-M., & Wu, T.-T. (2024). Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, *21*(1), 16. https://doi.org/10.1186/s41239-024-00447-4

Lewis, P. R., & Sarkadi, Ş. (2024). Reflective artificial intelligence. *Minds and Machines*, *34*(2), 14. https://doi.org/10.1007/s11023-024-09664-2

Liang, W., & Wu, Y. (2024). Exploring the use of ChatGPT to foster EFL learners' critical thinking skills from a post-humanist perspective. *Thinking Skills and Creativity*, *54*, 101645. https://doi.org/10.1016/j.tsc.2024.101645

Lin, C.-J., Lee, H.-Y., Wang, W.-S., Huang, Y.-M., & Wu, T.-T. (2025). Enhancing reflective thinking in STEM education through experiential learning: The role of generative AI as a learning aid. *Education and Information Technologies*, *30*(5), 6315–6337. https://doi.org/10.1007/s10639-024-13072-5

Liu, W., & Wang, Y. (2024). The effects of using AI tools on critical thinking in english literature classes among EFL learners: An intervention study. *European Journal of Education*, *59*(4), e12804. https://doi.org/10.1111/ejed.12804

Schenck, A. (2024). Examining relationships between technology and critical thinking: A study of south korean EFL learners. *Education Sciences*, *14*(6), Article 6. https://doi.org/10.3390/educsci14060652

Schön, D. A. (2017). *The reflective practitioner: How professionals think in action.* Routledge.

Shen, X., & Tao, Y. (2025). Metacognitive strategies, AI-based writing self-efficacy and writing anxiety in AI-assisted writing contexts: A structural equation modeling analysis. *International Journal of TESOL Studies*, *7*(1), 70–87. Scopus. https://doi.org/10.58304/ijts.20250105

Shen, X., & Teng, M. F. (2024). Three-wave cross-lagged model on the correlations between critical thinking skills, self-directed learning competency and AI-assisted writing. *Thinking Skills and Creativity, 52,* 101524. https://doi.org/10.1016/j.tsc.2024.101524

Strobl, C., Menke-Bazhutkina, I., Abel, N., & Michel, M. (2024). Adopting ChatGPT as a writing buddy in the advanced L2 writing class. *Technology in Language Teaching & Learning*, *6*(1), 1168. https://doi.org/10.29140/tltl.v6n1.1168

Teng, M. F. (2023). Scientific writing, reviewing, and editing for open-access TESOL journals: The role of ChatGPT. *International Journal of TESOL Studies, 5(*1), 87-91. https://doi.org/10.58304/ijts.20230107

Teng, M. F. (2024). A systematic review of ChatGPT for English as a foreign language writing: Opportunities, challenges, and recommendations. *International Journal of TESOL Studies*, *6*(3), 36–57. https://doi.org/10.58304/ijts.20240304

Teng, M. F. (2025). Metacognitive awareness and EFL learners' perceptions and experiences in utilising ChatGPT for writing feedback. *European Journal of Education*, *60*(1). https://doi.org/10.1111/ejed.12811

Todd, R. W. (2025). Generative AI as a disrupter of language education. *International Journal of TESOL Studies*, 1–9. https://doi.org/10.58304/ijts.250127

Ulla, M. B., & Teng, M. F. (2024). Generative artificial intelligence (AI) applications in TESOL: Opportunities, issues, and perspectives. *International Journal of TESOL Studies*, *6*(3), 1–3. https://doi.org/10.58304/ijts.20240301

United Nations. (n.d.a). *Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*. https://sdgs.un.org/goals/goal4?utm_source=chatgpt.com

United Nations. (n.d.b). *Goal 9: Build resilient infrastructure, promote sustainable industrialization and foster innovation*. https://www.un.org/sustainabledevelopment/infrastructure-industrialization/

Uştuk, Ö., & De Costa, P. I. (2021). Reflection as meta-action: Lesson study and EFL teacher professional development. *TESOL Journal*, *12*(1), e00531. https://doi.org/10.1002/tesj.531

Ward, F., Cho, J., Jung, J. K., Tognocchi, C., Beadles, T., & Cheon, J. (2025). ChatGPT for the intellectual soul: A Deweyan perspective on AI-based multicultural classroom praxis. *Multicultural Education Review*, *17*(1), 1–18. https://doi.org/10.1080/2005615X.2025.2467759

Wei, J., & Li, H. (2024). A systematic review of critical thinking development in information and communication technology-supported english as a foreign language teaching from 2015 to 2024. *Forum for Linguistic Studies*, *6*(6), 990–1006. Scopus. https://doi.org/10.30564/fls.v6i6.7478

Xu, J., & Liu, Q. (2025). Uncurtaining windows of motivation, enjoyment, critical thinking, and autonomy in AI-integrated education: Duolingo Vs. ChatGPT. *Learning and Motivation*, *89*. Scopus. https://doi.org/10.1016/j.lmot.2025.102100

Yang, S., Liu, Y., & Wu, T.-C. (2024a). ChatGPT, a new "Ghostwriter": A teacher-and-students poetic autoethnography from an EMI academic writing class. *Digital Applied Linguistics*, *1*, 2244. https://doi.org/10.29140/dal.v1.2244

Yang, S., Liu, Y., & Wu, T.-C. (2024b). ChatGPT, a new "Ghostwriter": A teacher-and-students poetic autoethnography from an EMI academic writing class. *Digital Applied Linguistics*, *1*, 2244. https://doi.org/10.29140/dal.v1.2244

Yusuf, A., Bello, S., Pervin, N., & Tukur, A. K. (2024). Implementing a proposed framework for enhancing critical thinking skills in synthesizing AI-generated texts. *Thinking Skills and Creativity*, *53*, 101619. https://doi.org/10.1016/j.tsc.2024.101619

Zou, X., Su, P., Li, L., & Fu, P. (2024). AI-generated content tools and students' critical thinking: Insights from a Chinese university. *IFLA Journal*, *50*(2), 228–241. https://doi.org/10.1177/03400352231214963

Dr. Ngo Cong-Lem is a lecturer in the Faculty of Foreign Languages at Dalat University, Vietnam. He holds a PhD from Monash University's Faculty of Education, Australia, ranked 8th globally in Education (2024–2025 U.S. News rankings). His research interests involve TESOL, the use of generative AI in education, and influential psychological factors in learning and teaching such as critical thinking, emotion, and agency. He has published in prestigious journals, such as Teaching and Teacher Education and System (Elsevier), Language Learning & Technology, and the International Journal of Human-Computer Interaction (Taylor & Francis). ORCID: https://orcid.org/0009-0005-3299-477X

Thang Tat Nguyen, Ph.D., is an Associate Professor at the Faculty of Foreign Languages, Dalat University, Vietnam. His research interests involve cognitive linguistics, socio-linguistics and teaching English as a second language. ORCID: https://orcid.org/0000-0001-8486-2385

Khanh Nhat Hoang Nguyen, MA, is a Lecturer in English at the Center for Foreign Languages and Human Resource Training at Dalat University. She teaches Academic Writing and TESOL methodology, and conducts research in applied linguistics, language education, and teacher development in Vietnam. Recently, she has investigated AI applications in English teaching, including adaptive learning platforms and automated feedback systems. She also contributes to curriculum design for diverse learner populations. Outside work, she reads professional literature and collaborates on innovative teaching projects. ORCID: https://orcid.org/0000-0001-6074-9935