JOLLT Journal of Languages and Language Teaching
https://e-journal.undikma.ac.id/index.php/jollt
Email: jollt@undikma.ac.id
DOI: https://doi.org/10.33394/jollt.v13i4.15077

October 2025. Vol. 13, No. 4
p-ISSN: 2338-0810
e-ISSN: 2621-1378
pp. 1817-1831

EXAMINING THE ROLE OF AI-POWERED WRITING ASSISTANTS IN ENHANCING CRITICAL THINKING IN EFL ACADEMIC WRITING

[1]*Muhammad Tahir, [1]Andi Sahtiani Jahrir, [1]Andi Asrifan, [2]Ariatna
[1]English Language Education, Universitas Negeri Makassar, Jl. Bonto Langkasa, Banta-Bantaeng, Kota Makassar, South Selatan, Indonesia
[2]Faculty of Language and Arts, Universitas Negeri Medan, Jl. Willem Iskandar, Pasar V, Medan Estate, Kota Medan, North Sumatera, Indonesia

*Corresponding Author Email: muhammadtahir@unm.ac.id

Article Info

Article History
Received: March 2025
Revised: April 2025
Accepted: September 2025
Published: October 2025

Keywords
AI writing assistants;
EFL academic writing;
Critical thinking;
AI feedback;
Proficiency levels;

Abstract

AI-driven writing assistants in EFL classrooms have revolutionized academic writing pedagogy, offering students immediate feedback on grammar, coherence, and organization. Although AI tools have demonstrated efficacy in improving linguistic precision, their influence on developing critical thinking remains ambiguous, especially among varying competency levels. Current study predominantly emphasizes AI's impact on grammatical corrections, although there is a paucity of knowledge on its effect on higher-order cognitive involvement, including argumentation and reasoning abilities. This study examines the interaction between EFL students and AI feedback, assessing its impact on promoting or obstructing critical thinking. The study reveals that, through examining pre-test and post-test writing evaluations, student reflections, and AI feedback patterns, lower-proficiency students (B1) predominantly depend on AI for superficial adjustments. In contrast, advanced learners (C1) interact with AI-generated ideas more critically. Nonetheless, AI's constraints in assessing argument strength and logical reasoning demonstrate that it cannot entirely supplant human feedback. These findings indicate that AI should be deliberately integrated with teacher support to optimize its advantages while reducing over-dependence. Future studies should investigate AI-human hybrid feedback models to improve language proficiency and critical thinking skills in academic writing.

How to cite: Tahir, M., Jahrir, A.S., & Asrifan, A., & Ariatna, A. (2025). Examining the Role of AI-Powered Writing Assistants in Enhancing Critical Thinking In EFL Academic Writing, *JOLLT Journal of Languages and Language Teaching*, 13(4), 1817-1831. Doi: https://doi.org/10.33394/jollt.v13i4.15077

INTRODUCTION

AI in education has transformed language learning and instruction (Alam, 2021; Yadav, 2025). English as a Foreign Language (EFL) schools increasingly use AI-powered writing aides like Grammarly, Turnitin, and ChatGPT to provide real-time feedback on grammar, coherence, and writing style. These technologies promise to improve students' writing by providing quick feedback, reducing errors, and boosting clarity. Automating writing assessments reduces instructors' burden, allowing them to focus on instructional tactics rather than error repair.

Higher education, where students must study academic writing, has integrated technology-driven learning aids (Buragohain et al., 2023; Fontenelle-Tereshchuk, 2025; Sharma et al., 2024). AI assistants give users immediate feedback and alternative expressions to improve their writing, meeting the growing demand for digital literacy and autonomous learning. AI-powered writing tools are also accessible and adaptable, helping kids write academically.

Despite their benefits, AI-powered writing aids may impair critical thinking (Al-Sofi, 2024; Cardon et al., 2023; Song & Song, 2023a). Writing requires reasoning, argumentation, and creativity (Kim et al., 2025; Kormos & Suzuki, 2024; Werdiningsih et al., 2024). In EFL

courses, do AI-powered writing aides promote critical thinking? As AI's function in language acquisition evolves, educators, linguists, and researchers have debated this.

Academic writing in EFL requires critical thinking to evaluate sources, create arguments, and reflect. Research shows students need critical thinking abilities to write well-structured, compelling, and clear arguments (Pelenkahu et al., 2024). Brainstorming, drafting, editing, and peer and instructor feedback foster critical thinking in writing (Apridayani & Waluyo, 2025; Yin et al., 2023). These features encourage students to critically evaluate their writing, refine their arguments, and use metacognition to improve their reasoning.

AI-powered writing tools offer fast fixes and alternatives but do not teach critical thinking. Overreliance on AI tools may cause students to passively accept AI-generated suggestions without challenging the reason (Vieriu & Petrea, 2025; Zhai et al., 2024). This raises concerns that students may favor surface-level edits over substantive argumentation and idea growth, discouraging deep writing involvement.

If appropriately utilized, AI-powered writing helpers can aid critical thinking. Researchers say AI-generated feedback can help students reflect on their writing choices and increase clarity and coherence (Oates & Johnson, 2025; Simms, 2024; Yusuf et al., 2024). AI commentary on ambiguous or unsupported statements may encourage students to strengthen their thinking and provide more evidence. Thus, AI techniques may supplement critical thinking. Only a few studies have examined how AI-powered writing assistants affect critical thinking in EFL contexts. Most AI studies focus on improving grammatical accuracy and fluency (Fathi et al., 2025; Zhai et al., 2024; D. Zhao, 2024), but there is little evidence of their effects on students' higher-order thinking skills like analysis, synthesis, and evaluation.

The debate over whether AI tools encourage writing autonomy or dependency is continuing. Some educators worry pupils would uncritically accept AI-generated edits, resulting in shallow writing (Malik et al., 2024; Stuchlikova & Weis, 2024). In contrast, AI feedback with correct instructional tactics can improve metacognitive awareness and self-directed learning (Mizumoto, 2023; Rajaram, 2023; Sridevi et al., 2024). This opposing view emphasizes the need to study how AI-powered writing aids affect EFL students' critical thinking.

Accordingly, this study is guided by a central research problem: To what extent do AI-powered writing assistants support or inhibit the development of critical thinking in EFL academic writing, especially when considering students' varying proficiency levels? To address this problem, the study aims to: (1) evaluate the impact of AI-generated feedback on students' critical thinking performance in academic writing tasks; (2) explore how EFL learners at different CEFR proficiency levels (B1, B2, C1) engage with AI tools; and (3) identify interaction patterns that facilitate or impede the development of reasoning, coherence, and argumentation skills. By doing so, this research contributes to a more nuanced understanding of how AI can be integrated pedagogically to cultivate both linguistic accuracy and intellectual autonomy in EFL writing contexts.

RESEARCH METHOD

Research Design

This study employs a mixed-methods approach, combining quantitative and qualitative techniques to thoroughly investigate the impact of AI-powered writing assistants on the enhancement of critical thinking in EFL academic writing. The quantitative aspect entails a comparison of students' academic writing performance in pre-tests and post-tests, emphasizing indicators of critical thinking including argument clarity, coherence, and reasoning. The qualitative component encompasses students' written reflections, semi-structured interviews, and survey responses to obtain profound insights into their interaction with AI-generated feedback.

The justification for utilizing a mixed-methods design is its ability to furnish a comprehensive perspective—quantitative data supplies statistical proof of change, whereas qualitative data uncovers learners' perceptions and cognitive strategies. This design conforms to guidelines for rigorous educational research that prioritizes both quantifiable results and contextual comprehension (Alhassan, 2024; Vos & van Rijn, 2025).

Participants

The study's participants are EFL university students in academic writing courses at a medium-sized university. The sample comprises 100 undergraduate students, aged 18 to 25, from varied academic disciplines, predominantly majoring in English Language Education and Applied Linguistics. The participants exhibit diverse competence levels in English, classified according to standardized placement assessments (CEFR levels B1–C1). A purposeful sampling strategy is utilized to guarantee that the chosen students have past expertise in academic writing and are actively involved in writing projects necessitating critical thinking. This method guarantees that participants can offer pertinent insights regarding the efficacy of AI-driven writing assistance in cultivating advanced cognitive abilities.

The sample comprises students with varying degrees of competence with AI tools to ensure equitable representation. Participants certainly have prior experience with AI-driven writing helpers, whereas others are novice users. This variability facilitates a comparative investigation of the impact of AI tools on pupils across varying skill levels. All participants grant informed consent before participating in the study, and their identities are kept anonymous. This selection guarantees a representative sample of EFL learners, facilitating a comprehensive examination of the influence of AI on the enhancement or obstruction of critical thinking.

Table 1
Participant Data

Proficiency Level	Number of Participants	Percentage (%)
B1 (Intermediate)	40	40
B2 (Upper-Intermediate)	35	35
C1 (Advanced)	25	25

Table 1 illustrates the distribution of 100 EFL university students classified by competency level according to CEFR standards. Most participants (40%) are classified as B1 (Intermediate), followed by B2 (Upper-Intermediate) at 35%, and C1 (Advanced) at 25%. This selection guarantees a broad representation of students with differing levels of English ability, facilitating a comparative investigation of the effects of AI-powered writing assistance on various learner groups. The equitable distribution of participants yields significant information on whether AI technologies promote or obstruct the advancement of critical thinking across different levels of language competency.

The B1 level has the greatest number of participants, succeeded by B2 and C1. This distribution is essential for determining if AI-powered writing aids impact pupils variably according to their linguistic proficiency. The study seeks to determine if AI technologies assist less proficient students or predominantly enhance the capabilities of expert learners, given the correlation between critical thinking, writing complexity, and language competency. The graphical depiction facilitates the interpretation of the relative proportions of each proficiency category within the research sample.

Research Instruments

This research utilized three prevalent AI-driven writing tools—Grammarly, ChatGPT, and Turnitin—to enhance students' academic writing skills. Grammarly offered instantaneous

feedback on grammar, clarity, and sentence structure; ChatGPT facilitated idea generation, organization, and content expansion; and Turnitin guaranteed originality by identifying potential plagiarism and providing citation recommendations.

To assess students' critical thinking abilities, participants engaged in academic writing assignments, including argumentative essays and critical responses. The tasks were evaluated utilizing a validated critical thinking rubric, modified from Facione's (1990) Critical Thinking Scoring Rubric (CTS) and Paul and Elder's (2006) Elements of Thought (Adri and Abdullah 2022; Liu et al. 2023; Bates et al. 2025). The rubric included five essential dimensions. The rubric included five essential dimensions: Clarity of Argument (e.g., clearly stated thesis and purpose); use of Evidence (e.g., appropriate support, relevance, and credibility); logical Organization and Coherence (e.g., transitions and argument flow); evaluation of Counterarguments (e.g., refutation and balance); and reflective Judgment (e.g., synthesis, depth of reasoning, and originality). Each criterion was evaluated using a 5-point Likert scale (1 = Very Weak to 5 = Excellent). Two trained evaluators independently evaluated all student submissions, and inter-rater reliability was established (Cohen's Kappa = 0.82).

Data Collection Procedures

The data collection process spanned eight weeks and was organized into four distinct phases to guarantee systematic execution. During the initial phase, participants undertook a pre-test writing task devoid of AI assistance, thereby establishing baseline metrics for critical thinking and writing proficiency. During the second phase, participants engaged in two training sessions that presented the features, functionalities, and ethical applications of Grammarly, ChatGPT, and Turnitin. These sessions underscored the importance of critical engagement with AI recommendations, rather than passive acquiescence.

The third phase comprised three AI-integrated writing assignments, wherein participants were directed to employ AI tools for drafting and revising. Significantly, in this phase, instructor feedback was intentionally withheld to isolate the impact of AI-generated feedback. Instructors facilitated by overseeing students' interactions with AI, providing clarifications on tool functionalities, and conducting reflective mini-sessions post-task to encourage metacognitive awareness, such as questioning the relevance and logic of AI suggestions. This method guaranteed that learners interacted with AI assistance autonomously, while remaining within a structured and pedagogically sound framework.

In the concluding phase, students undertook a post-test writing task, again devoid of AI assistance, succeeded by semi-structured interviews and surveys. The final stage sought to assess the evolution of critical thinking, alterations in writing performance, and students' perceptions of AI feedback. The design facilitated a regulated evaluation of AI's influence and a contextual comprehension of learner experiences.

Data Analysis

The research used a mixed-methods strategy for data analysis, combining quantitative and qualitative techniques to evaluate the influence of AI-powered writing aids on critical thinking. Quantitative analysis examines pre-test and post-test writing scores through descriptive statistics, paired t-tests, and ANOVA to assess enhancement in critical thinking skills across varying proficiency levels. Statistical analyses ascertain whether AI-assisted feedback results in substantial advancements in argumentation, coherence, and reasoning.

Qualitative analysis examines students' interviews, surveys, and written comments through thematic coding to discern trends in their interaction with AI-generated feedback. The themes analyzed include reliance on AI, improved reasoning, and acceptance of passive correction. To guarantee reliability and validity, inter-rater dependability is established for rubric scoring, and triangulation is utilized by juxtaposing quantitative data with qualitative

insights. This thorough approach bolsters the study's trustworthiness and offers a nuanced comprehension of AI's influence on EFL writing development.

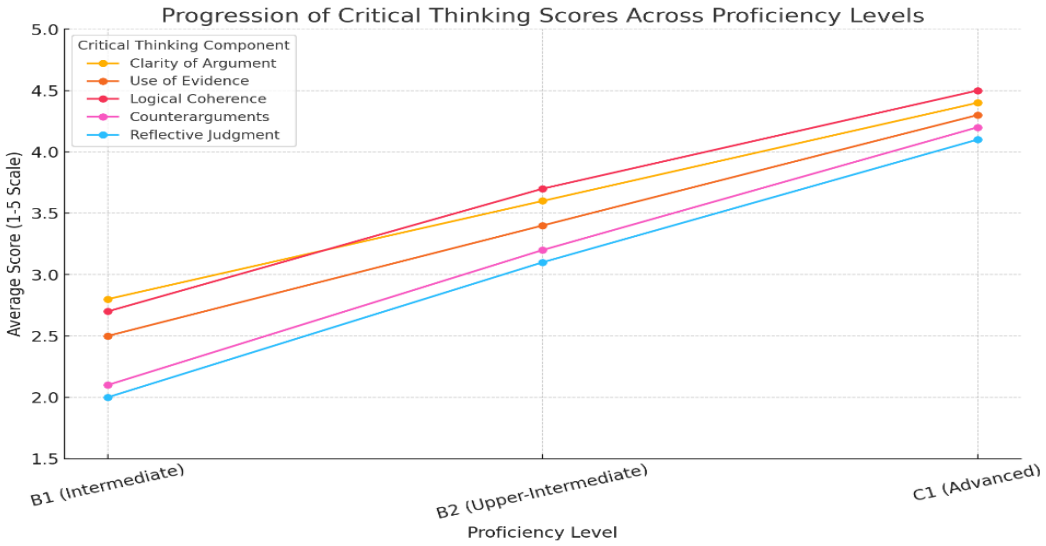

Figure 1. Progression of Critical Thinking Scores Across Proficiency Levels

Figure 1 depicts the advancement of mean scores across five essential components of critical thinking categorized by skill level. The data demonstrates a steady rising trajectory from B1 to C1, with C1 learners attaining the best ratings across all categories, especially in logical coherence and reflective judgment, signifying enhanced critical engagement with writing assignments.

RESEARCH FINDINGS AND DISCUSSION

Results

Quantitative Results

The comparison between pre-test and post-test indicates a notable enhancement in students' critical thinking abilities following the utilization of AI-powered writing assistance. Pre-test scores reveal that most students had difficulties in argument formation, coherence, and evidence-based reasoning, whereas post-test results significantly improved in these domains. The average pre-test score for all participants was 65.2 (SD = 8.4), whereas the post-test average rose to 78.5 (SD = 7.1). The 13.3-point rise indicates that AI tools promoted significant learning, while the degree of enhancement differed according to competency level. Although AI assisted students in enhancing sentence form and clarity, the development of critical thinking was more pronounced in those who actively interacted with AI feedback rather than those who only accepted corrections passively. These findings underscore AI incorporation's advantages and possible constraints in EFL writing pedagogy.

Table 2
Pre-test vs. Post-test Scores

Proficiency Level	Pre-Test Mean Score	Post-Test Mean Score	Score Improvement
B1 (Intermediate)	60.5	72.1	11.6
B2 (Upper-Intermediate)	66.8	79.4	12.6
C1 (Advanced)	71.2	85.2	14
Overall	65.2	78.5	13.3

Table 2 displays students' average scores at three competence levels (B1, B2, and C1) before and after the utilization of AI-powered writing aids. The pre-test mean score was 65.2,

which increased to 78.5 in the post-test, reflecting an enhancement of 13.3 points. The most significant enhancement was noted in C1 learners (14.0 points), succeeded by B2 learners (12.6 points) and B1 learners (11.6 points). The data demonstrate that all proficiency levels benefited from AI-assisted feedback, with higher-proficiency students seeing the greatest enhancement in reasoning and coherence, whilst lower-level students largely improved in grammar and clarity. The significant rise in scores among all groups underscores the beneficial effect of AI-powered writing aids on students' academic writing abilities. B1 students exhibited moderate improvement; however, B2 and C1 learners displayed more substantial advancements, suggesting that AI tools were especially advantageous for those with pre-existing critical thinking skills. The findings underscore that AI-assisted feedback has a facilitative function in enhancing students' writing. However, it may prove more beneficial for advanced EFL learners.

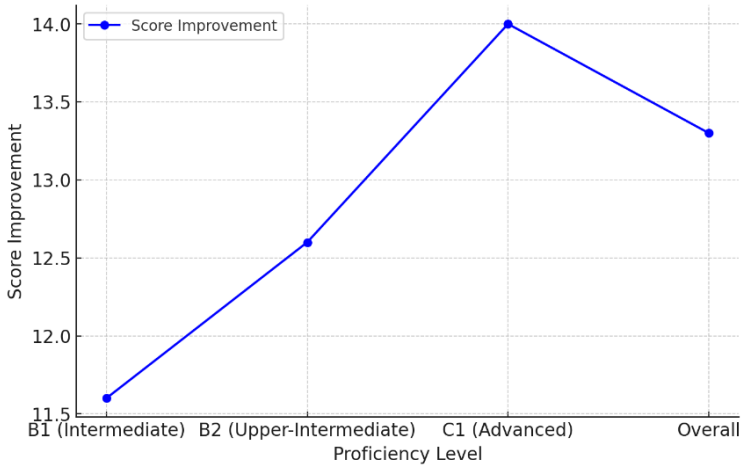

Figure 2. Score Improvement Across Proficiency Levels

Figure 2 depicts the trend of score enhancement throughout proficiency levels, indicating a steady rising trajectory from B1 to C1. C1 students demonstrated the most significant enhancement, indicating that advanced learners are more adept at leveraging AI feedback for critical analysis of their writing. The incremental progression from B1 to B2 suggests that although AI technologies assist lower-proficiency pupils, their improvements are confined to superficial fixes. The figure indicates that AI-driven writing assistants ought to be combined with instructional support, especially for B1 learners, to improve their capacity for critical engagement with AI-generated recommendations.

Proficiency Level Differences

The examination of AI-assisted writing assignments indicates significant disparities in the utilization and advantages experienced by B1, B2, and C1 pupils when employing AI-powered writing aids. B1 (Intermediate) pupils predominantly concentrated on grammatical fixes and enhancements in sentence structure, frequently acquiescing to AI-generated recommendations without substantial participation. Their writing improved clarity and coherence; nevertheless, advances in critical thinking were negligible since they predominantly depended on AI for superficial adjustments instead of developing arguments.

B2 (Upper-Intermediate) pupils exhibited a more equitable methodology, employing AI feedback to enhance linguistic precision and logical thinking. They demonstrated heightened awareness of the significance of AI input, selectively adopting it to improve cohesiveness, evidence integration, and argument clarity. Their critical thinking abilities enhanced, although AI technologies predominantly facilitated structural refinement instead of profound analytical involvement.

C1 (Advanced) students derived the most significant advantage from utilizing AI tools to critically evaluate and enhance their arguments instead of passively accepting corrections. They utilized AI input to enhance thesis statements, counterarguments, and logical consistency. Their enhancements were crucial, as they critically interrogated AI feedback, rewrote using logical reasoning, and included AI aid while maintaining originality. The findings indicate that AI-driven writing tools are most efficacious when pupils have established fundamental critical thinking abilities.

Table 3
Proficiency Level Differences in AI-Assisted Writing

Proficiency Level	Grammar & Clarity Improvement	Critical Thinking Improvement	Argumentation & Coherence Improvement
B1 (Intermediate)	80	50	55
B2 (Upper-Intermediate)	85	70	75
C1 (Advanced)	90	85	90

Table 3 delineates the enhancements in writing skills throughout three competency tiers (B1, B2, and C1) after utilizing AI-powered writing assistance. The enhancement of grammar and clarity was most pronounced among C1 learners (90%), succeeded by B2 (85%) and B1 (80%), demonstrating that all groups gained from AI-assisted feedback. Nevertheless, the enhancement of critical thinking was markedly inferior for B1 learners (50%) in comparison to B2 (70%) and C1 (85%), indicating that lower-proficiency pupils depended more on AI for superficial fixes. The most significant improvements were in argumentation and coherence, especially among C1 students (90%), who adeptly incorporated AI feedback into their critical thinking processes.

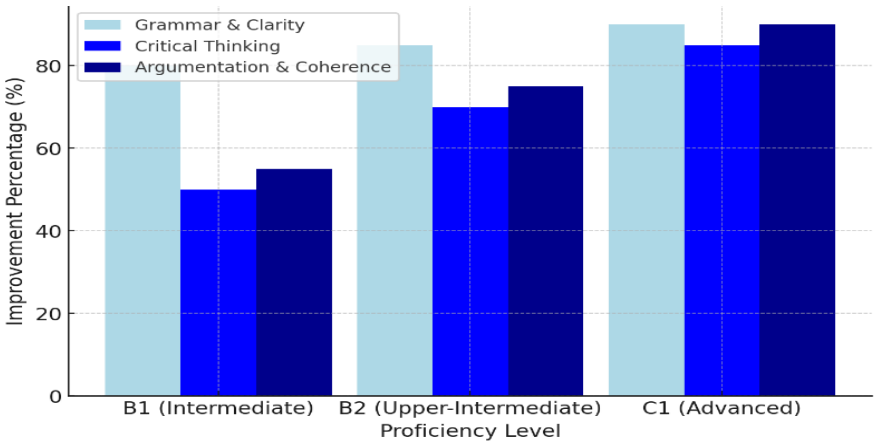

Figure 3. Proficiency Level Differences in AI-Assisted Writing Tasks

Figure 3 delineates the disparities in writing enhancement domains across varying competence levels. The findings indicate that B1 students predominantly gained from input on grammar and clarity, but C1 students exhibited the most significant advancements in argumentation and critical thinking. The disparity between critical thinking and grammatical enhancements at various competence levels indicates that novice students predominantly employ AI for linguistic corrections. In contrast, advanced learners engage with AI input more analytically to enhance their reasoning and argumentative framework. These data suggest that the efficacy of AI is contingent upon students' cognitive engagement with the input.

AI Tool Effectiveness Metrics

Examining students' changes informed by AI-generated feedback uncovers specific patterns in the engagement of varying competence levels with AI tools. The research classifies AI feedback into three categories: language and mechanics, coherence and organization, and argumentation and critical reasoning.

B1 students predominantly utilized grammar and mechanics input, such as spelling, punctuation, and sentence structure, implementing AI suggestions with minimal alteration. B2 students exhibited a more equitable methodology, employing AI to enhance grammatical precision and coherence, facilitate smoother transitions, and augment readability. B2 and C1 students more effectively employed coherence and organization input, such as paragraph structure and logical flow, revising their work according to AI recommendations to improve clarity and logical progression.

C1 students primarily utilized feedback on argumentation and critical reasoning, such as claim building, evidence support, and counterarguments, to critically engage with AI-generated ideas, thus enhancing their arguments while preserving original thought.

Research demonstrates that although AI tools improve language precision, their influence on critical thinking and argumentation is contingent upon students' capacity to evaluate and incorporate input instead of passively accepting changes actively. This highlights the necessity for instructional scaffolding to guarantee that AI technologies function as cognitive aids rather than crutches.

Table 4
AI Tool Effectiveness Metrics

Proficiency Level	Grammar & Mechanics Feedback Usage (%)	Coherence & Organization Feedback Usage (%)	Argumentation & Critical Reasoning Feedback Usage (%)
B1 (Intermediate)	85	50	30
B2 (Upper-Intermediate)	75	70	55
C1 (Advanced)	60	85	80

Table 4 displays the percentage of AI feedback utilization across three competency levels (B1, B2, and C1) in various writing dimensions: grammar and mechanics, coherence and organization, and argumentation and critical reasoning. B1 pupils predominantly depended on grammar input (85%), emphasizing superficial improvements. B2 students demonstrated a balanced methodology, employing coherence (70%) and grammatical (75%) criticism. C1 students had the highest level of engagement with argumentation feedback (80%), demonstrating a critical interaction with AI-generated recommendations for improving logical reasoning and argument form. The results indicate that elevated proficiency levels optimize AI for enhanced cognitive engagement.

Qualitative Results

The qualitative research, derived from student interviews and surveys, identifies three predominant themes about AI-powered writing assistants: autonomous learning, reliance on AI, and developing critical thinking.

Autonomous Learning: Numerous pupils indicated that AI feedback facilitated their development as independent authors by pinpointing frequent errors and proposing enhancements. C1 students demonstrated notable reflection, employing AI as a pedagogical instrument rather than a corrective device. They characterized AI feedback as a framework that directed them towards improved self-editing and revising techniques.

An opposing theme appeared, notably among B1 students, who tended to excessively depend on AI advice without critically analyzing the recommendations. Some acknowledged uncritically accepting AI corrections, although lacking a comprehensive understanding of the underlying rationale. This suggests that whereas AI enhances verbal precision, it may also diminish cognitive involvement if not employed judiciously.

Development of Critical Thinking: B2 and C1 students observed that AI comments on coherence and argumentation enhanced their thinking and logical progression. Nonetheless, they recognized the constraints of AI in evaluating profound analytical reasoning, necessitating human involvement. The findings indicate that AI-powered tools are advantageous but require the integration of essential engagement methods to enhance their effectiveness in writing development.

Table 5
Qualitative Results: AI Feedback Themes

Proficiency Level	Autonomous Learning (%)	AI Reliance (%)	Critical Thinking Development (%)
B1 (Intermediate)	40	85	30
B2 (Upper-Intermediate)	65	60	70
C1 (Advanced)	80	35	85

Table 5 displays the percentage of student reflections about AI feedback themes throughout B1, B2, and C1 competence levels. C1 students (80%) exhibited the greatest participation in autonomous learning, whereas B1 students (40%) demonstrated the least independence. The dependence on AI was greatest among B1 learners (85%), suggesting a propensity to accept AI-generated adjustments passively. Conversely, C1 students (35%) relied less on AI and used it judiciously to augment their case. The advancement of critical thinking was most pronounced among C1 students (85%), indicating that higher-proficiency learners critically engaged with AI input, while B1 students (30%) exhibited limited critical engagement.

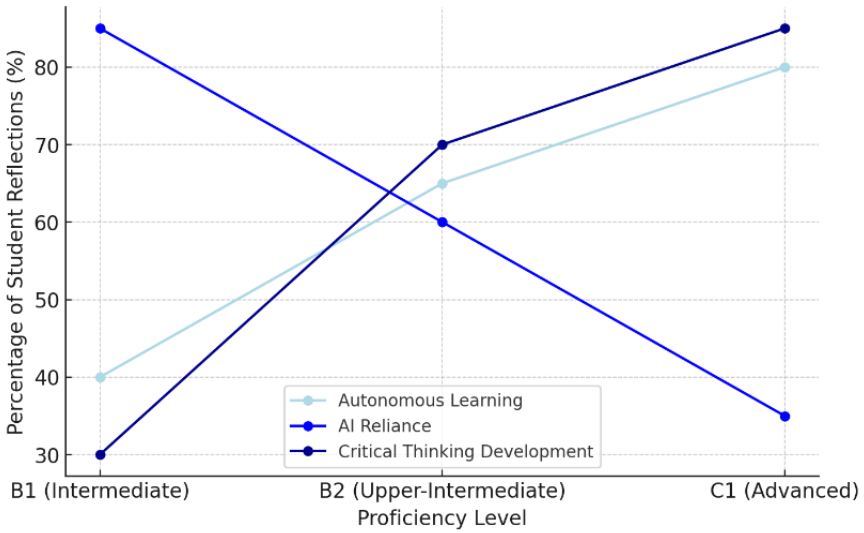

Figure 4. Trends in AI Feedback Utilization Themes

Figure 4 depicts trends in adopting AI feedback across B1, B2, and C1 learners. Enhancing autonomous learning and critical thinking correlates with proficiency, suggesting that advanced learners engage with AI technologies more adeptly. Conversely, dependence on AI diminishes, indicating that lower-level students are increasingly dependent on AI-

generated recommendations without engaging in thorough examination. The graphic underscores the necessity of incorporating AI with pedagogical practices to promote critical engagement among B1 and B2 learners rather than passive acceptance of AI feedback. This pattern indicates that AI technologies are most advantageous when students actively reflect and strategically implement feedback.

Challenges and Limitations of AI Usage

Notwithstanding the benefits of AI-driven writing assistance, students faced numerous obstacles in properly leveraging AI-generated feedback. A significant challenge was the interpretation of AI recommendations, especially among B1 learners, who frequently accepted modifications without comprehending the underlying reasoning. Numerous individuals indicated that AI feedback, particularly regarding coherence and argumentation, was deficient in clarity, hindering the ability to effectuate enhancements beyond superficial modifications. Students at the B2 and C1 levels showed greater discernment in utilizing AI suggestions, yet they encountered discrepancies since AI advice sometimes conflicted with established academic writing practices.

A notable constraint is AI's capacity to promote higher-order thinking. Although AI technologies effectively resolve grammar, clarity, and structural concerns, they falter in evaluating argument strength, critical reasoning, and originality. Numerous C1 students observed that AI feedback on reasoning frequently emphasized diction and phrase coherence above logical rigor or evidence assessment. B2 students articulated divergent perspectives—some regarded AI as beneficial for structuring ideas, but others perceived it as deficient in providing substantive insights for critical engagement.

These challenges suggest that although AI can improve the quality of technical writing, it should not be exclusively relied upon to cultivate critical thinking skills. Successful integration necessitates human supervision, wherein educators assist students in critically interpreting AI feedback and strategically using it in academic writing.

Table 6
Challenges and Limitations of AI Usage

Proficiency Level	Difficulty in Interpreting AI Suggestions (%)	Perceived AI Limitations in Higher-Order Thinking (%)	Over-Reliance on AI Feedback (%)
B1 (Intermediate)	80	50	85
B2 (Upper-Intermediate)	60	70	65
C1 (Advanced)	35	85	40

Table 6 delineates the problems and constraints associated with AI utilization at B1, B2, and C1 proficiency levels. B1 students (80%) indicated the greatest challenge in comprehending AI recommendations, frequently finding it difficult to implement comments effectively. Perceived limits of AI in higher-order thinking escalated with proficiency, as 85% of C1 students acknowledged AI's incapacity to assess the depth of arguments. B1 pupils exhibited the highest dependence on AI input at 85%, which rapidly diminished to 40% among C1 learners. These findings indicate that AI is effective for superficial adjustments; nonetheless, cultivating critical thinking necessitates human involvement to guarantee substantive interaction with AI-generated feedback.

Discussion

The findings of this study highlight a complex relationship between the use of AI-powered writing assistants and the development of critical thinking in EFL academic writing. The quantitative results demonstrate that AI-assisted feedback enhances overall writing performance, particularly in linguistic accuracy and coherence. This aligns with earlier studies

emphasizing the benefits of automated writing evaluation (AWE) systems for improving grammatical precision and fluency. For instance, Wei et al. (2023) reported that AI-based feedback significantly increased students' linguistic competence and reduced surface-level errors in writing tasks. Similarly, Fathi et al. (2025) found that intelligent personal assistants fostered fluency and comprehensibility in EFL learners. However, the current study extends these insights by revealing that while such gains are evident across proficiency levels, the development of higher-order thinking skills—such as argumentation and reflective judgment—occurs predominantly among advanced learners. This pattern corroborates findings by Kim et al. (2025) and Werdiningsih et al. (2024), who noted that proficient students tend to engage with AI-generated feedback more critically, transforming it into a tool for metacognitive reflection rather than passive correction.

The differences in how learners of varying proficiency levels interact with AI feedback reveal the cognitive scaffolding necessary for meaningful engagement. B1 students primarily relied on AI for mechanical and syntactic correction, indicating a surface approach to learning. This finding resonates with Zhai et al. (2024), who warned that excessive dependence on AI tools can inhibit students' cognitive effort and diminish deep learning. Conversely, B2 and C1 students showed evidence of analytical engagement, selectively incorporating AI suggestions to refine coherence, strengthen arguments, and enhance reasoning. These outcomes mirror the observations of Oates and Johnson (2025), who argued that AI feedback can stimulate metacognitive evaluation when accompanied by reflective strategies. The current findings therefore reinforce the importance of pedagogical mediation—teachers must guide learners to question and evaluate AI recommendations rather than accept them uncritically. This pedagogical implication echoes Stuchlikova and Weis (2024), who stressed the need to reimagine critical thinking instruction in the age of AI by embedding reflective dialogue around automated feedback.

From a theoretical standpoint, the findings can be interpreted through the lens of constructivist and metacognitive learning theories. Constructivism posits that learners actively construct knowledge through interaction and reflection; thus, AI feedback serves as a cognitive scaffold that prompts hypothesis testing, revision, and reasoning. For higher-proficiency learners, AI becomes an instrument of self-regulated learning, encouraging deeper analysis of argument structures and evidence use. This observation is consistent with Mizumoto's (2023) metacognitive resource use framework, which suggests that AI can promote reflective awareness when learners treat feedback as data to interpret rather than as instructions to follow. However, the study also uncovers a limitation of AI systems—they lack the epistemic capacity to evaluate logic, argument validity, or rhetorical balance. Such deficiencies reaffirm Guo's (2024) assertion that AI-generated corrective feedback, though useful, cannot replicate human judgment in fostering critical reasoning. Hence, while AI enhances procedural accuracy, it requires human mediation to cultivate interpretive depth and intellectual autonomy.

Practically, the results underscore the need for hybrid feedback models that integrate AI efficiency with human evaluative insight. The study found that C1 learners benefited most when they combined AI feedback with self-reflection, suggesting that instructional frameworks should blend automated suggestions with teacher-guided critical analysis. This approach supports the recommendation of Yusuf et al. (2024), who proposed synthesizing AI-generated texts with explicit training in argument evaluation to enhance critical thinking. Moreover, the data demonstrate that learners' attitudes toward AI evolve with proficiency—while B1 students perceived AI as an authoritative corrector, C1 students viewed it as a dialogic partner. Such transformation aligns with the progression from dependency to autonomy described by Bai and Wang (2023) in their exploration of self-regulated learning.

Consequently, fostering digital and AI literacy becomes integral to EFL pedagogy, ensuring students can discern, adapt, and interrogate automated feedback responsibly.

Despite its promising outcomes, the study reveals inherent challenges in using AI tools for developing higher-order cognition. Students frequently encountered ambiguities in AI feedback on coherence and reasoning, a limitation also noted by Song and Song (2023a), who found that AI-based systems excel in linguistic correction but perform poorly in assessing conceptual depth. This reinforces the argument of Kennedy and Romig (2024) that cognitive load may increase when students must interpret unclear feedback without adequate scaffolding. Therefore, instructional interventions must include explicit training on interpreting AI recommendations and integrating them into argument-based writing tasks. The findings also suggest a pedagogical imperative to cultivate critical digital awareness—students should not merely "use" AI tools but understand their epistemological boundaries. Such awareness aligns with Cardon et al. (2023), who advocated for developing "AI literacy" that combines technical proficiency with ethical and critical evaluation of algorithmic input.

Overall, this study contributes to the growing body of literature on AI-assisted learning by offering a nuanced perspective on its pedagogical and cognitive implications. It confirms that AI-powered writing assistants can effectively enhance writing accuracy and organization but only contribute meaningfully to critical thinking when learners possess sufficient linguistic and metacognitive maturity. The findings thus position AI not as a substitute for human instruction but as a complementary cognitive tool that can amplify reflective and analytical writing practices. For educators, this implies the need to design integrative learning environments that leverage AI's strengths—speed, precision, and accessibility—while embedding human guidance that nurtures reasoning, creativity, and independent thought. In sum, the synergy between AI and pedagogy, rather than the technology itself, determines the extent to which EFL learners evolve from passive writers into autonomous, critically engaged thinkers.

CONCLUSION

The findings of this study reveal that AI-powered writing assistants play a dual and nuanced role in the development of EFL students' critical thinking within academic writing contexts. On one hand, these tools substantially enhance linguistic precision, coherence, and organizational quality across proficiency levels, enabling learners to produce clearer and more structured texts. The significant improvement in post-test scores, particularly among C1 learners, demonstrates that AI-assisted feedback can facilitate metacognitive awareness and encourage reflective writing practices when learners engage critically with the technology. However, the study also exposes the limitations of AI in nurturing deeper cognitive abilities such as reasoning, argument evaluation, and synthesis. Lower-proficiency students tended to rely heavily on AI-generated suggestions, accepting corrections passively and focusing mainly on grammatical and stylistic improvements. This pattern suggests that without explicit pedagogical scaffolding, AI tools may inadvertently promote surface-level engagement rather than higher-order thinking. Therefore, while AI-driven feedback mechanisms hold transformative potential, their educational value depends largely on how learners interpret and interact with them. Teachers must play an essential mediating role—guiding students to critically evaluate AI input, integrate it meaningfully into their writing, and maintain intellectual autonomy throughout the learning process.

From a pedagogical and theoretical perspective, this study underscores the importance of integrating AI assistance within a human-centered framework that balances automation with critical pedagogy. Theoretically, the findings reinforce constructivist and metacognitive models of learning, demonstrating that active engagement and reflective reasoning are indispensable for translating technological feedback into cognitive growth. Practically, the

research advocates for hybrid feedback models that combine AI efficiency with human insight, ensuring that students not only refine linguistic accuracy but also develop argumentation, judgment, and evaluative reasoning. Educators are encouraged to incorporate AI literacy training into EFL curricula, enabling students to understand both the affordances and the limitations of these tools. Future research should explore adaptive AI-human feedback systems that can respond to learners' proficiency levels and cognitive needs dynamically. Ultimately, this study contributes to a broader understanding of AI's pedagogical role—not as a replacement for critical human instruction but as a catalyst for developing autonomous, reflective, and critically aware writers in the evolving landscape of digital education.

REFERENCES

Alam, A. (2021). Possibilities and Apprehensions in the Landscape of Artificial Intelligence in Education. *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)*, 1–8. https://doi.org/10.1109/ICCICA52458.2021.9697272

Alhassan, A. I. (2024). Analyzing the application of mixed method methodology in medical education: a qualitative study. *BMC Medical Education*, *24*(1), 225. https://doi.org/10.1186/s12909-024-05242-3

Al-Sofi, B. B. M. A. (2024). Artificial intelligence-powered tools and academic writing: to use or not to use ChatGPT. *Saudi Journal of Language Studies*, *4*(3), 145–161. https://doi.org/10.1108/SJLS-06-2024-0029

Apridayani, A., & Waluyo, B. (2025). Valuing students' mindsets in essay compositions: Active learning and feedback approaches. *Social Sciences & Humanities Open*, *11*, 101273. https://doi.org/10.1016/j.ssaho.2024.101273

Asrifan, A., Khristianto, K., Budiman, A., Astuti, P. I., & Rossydi, A. (2025). *Enhancing Critical Thinking* (pp. 413–448). https://doi.org/10.4018/979-8-3373-4576-5.ch012

Bai, B., & Wang, J. (2023). The role of growth mindset, self-efficacy and intrinsic value in self-regulated learning and English language learning achievements. *Language Teaching Research*, *27*(1), 207–228. https://doi.org/10.1177/1362168820933190

Buragohain, D., Punpeng, G., Jaratjarungkiat, S., & Chaudhary, S. (2023). Impact of E-Learning Activities on English as a Second Language Proficiency among Engineering Cohorts of Malaysian Higher Education: A 7-Month Longitudinal Study. *Informatics*, *10*(1), 31. https://doi.org/10.3390/informatics10010031

Cardon, P., Fleischmann, C., Aritz, J., Logemann, M., & Heidewald, J. (2023). The Challenges and Opportunities of AI-Assisted Writing: Developing AI Literacy for the AI Age. *Business and Professional Communication Quarterly*, *86*(3), 257–295. https://doi.org/10.1177/23294906231176517

Fathi, J., Rahimi, M., & Teo, T. (2025). Applying intelligent personal assistants to develop fluency and comprehensibility, and reduce accentedness in EFL learners: an empirical study of Google Assistant. *Language Teaching Research*. https://doi.org/10.1177/13621688251317786

Fontenelle-Tereshchuk, D. (2025). First-Year College Students: Perspectives on Technology and Wellness in Education. *Trends in Higher Education*, *4*(1), 13. https://doi.org/10.3390/higheredu4010013

Guo, X. (2024). Facilitator or thinking inhibitor: understanding the role of ChatGPT-generated written corrective feedback in language learning. *Interactive Learning Environments*, 1–19. https://doi.org/10.1080/10494820.2024.2445177

Kennedy, M. J., & Romig, J. E. (2024). Cognitive Load Theory: An Applied Reintroduction for Special and General Educators. *TEACHING Exceptional Children*, *56*(6), 440–451. https://doi.org/10.1177/00400599211048214

Kim, J., Yu, S., Detrick, R., & Li, N. (2025). Exploring students' perspectives on Generative AI-assisted academic writing. *Education and Information Technologies*, *30*(1), 1265–1300. https://doi.org/10.1007/s10639-024-12878-7

Kormos, J., & Suzuki, S. (2024). Creativity and the linguistic features of argumentative and narrative written task performance. *System*, *127*, 103531. https://doi.org/10.1016/j.system.2024.103531

Malik, M. A., Amjad, A. I., Aslam, S., & Fakhrou, A. (2024). Global insights: ChatGPT's influence on academic and research writing, creativity, and plagiarism policies. *Frontiers in Research Metrics and Analytics*, *9*. https://doi.org/10.3389/frma.2024.1486832

Mizumoto, A. (2023). Data-driven Learning Meets Generative AI: Introducing the Framework of Metacognitive Resource Use. *Applied Corpus Linguistics*, *3*(3), 100074. https://doi.org/10.1016/j.acorp.2023.100074

Oates, A., & Johnson, D. (2025). ChatGPT in the Classroom: Evaluating its Role in Fostering Critical Evaluation Skills. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-024-00452-8

Pelenkahu, N., Ali, M. I., Tatipang, D. P., Wuntu, C. N., & Rorintulus, O. A. (2024). Metacognitive strategies and critical thinking in elevating EFL argumentative writing proficiency: Practical insights. *Studies in English Language and Education*, *11*(2), 873–892. https://doi.org/10.24815/siele.v11i2.35832

Rajaram, K. (2023). Future of Learning: Teaching and Learning Strategies. In *Learning Intelligence: Innovative and Digital Transformative Learning Strategies* (pp. 3–53). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-9201-8_1

Shanmugavelu, G., Parasuraman, B., Ariffin, K., Kannan, B., & Vadivelu, M. (2020). Inquiry Method in the Teaching and Learning Process. *Shanlax International Journal of Education*, *8*(3), 6–9. https://doi.org/10.34293/education.v8i3.2396

Sharma, A., Bharti, & Pandey, A. (2024). Unleashing the Potential of Technology-Driven Learning Management Systems for Student-Centric Excellence to Empower Higher Education. *RAiSE-2023*, 195. https://doi.org/10.3390/engproc2023059195

Simms, R. C. (2024). Work With ChatGPT, Not Against. *Nurse Educator*, *49*(3), 158–161. https://doi.org/10.1097/NNE.0000000000001634

Song, C., & Song, Y. (2023a). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1260843

Song, C., & Song, Y. (2023b). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1260843

Sridevi, R., Ashokkumar, P., & Sathish, V. (2024). *Empowering Minds* (pp. 125–152). https://doi.org/10.4018/979-8-3693-3559-8.ch005

Stuchlikova, L., & Weis, M. (2024). From Information to Insight: Reimagining Critical Thinking Pedagogy in the Age of Artificial Intelligence. *2024 International Conference on Emerging ELearning Technologies and Applications (ICETA)*, 591–598. https://doi.org/10.1109/ICETA63795.2024.10850787

Vieriu, A. M., & Petrea, G. (2025). The Impact of Artificial Intelligence (AI) on Students' Academic Development. *Education Sciences*, *15*(3), 343. https://doi.org/10.3390/educsci15030343

Vos, J., & van Rijn, B. (2025). Using mixed methods in feasibility studies: The example of brief transactional analysis psychotherapy for depression. *Counselling and Psychotherapy Research*, *25*(1). https://doi.org/10.1002/capr.12871

Wei, P., Wang, X., & Dong, H. (2023). The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: a randomized controlled trial. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1249991

Werdiningsih, I., Marzuki, & Rusdin, D. (2024). Balancing AI and authenticity: EFL students' experiences with ChatGPT in academic writing. *Cogent Arts & Humanities*, *11*(1). https://doi.org/10.1080/23311983.2024.2392388

Yadav, S. (2025). Leveraging AI to Enhance Teaching and Learning in Education. In *Optimizing Research Techniques and Learning Strategies With Digital Technologies* (pp. 211–238). IGI Global. https://doi.org/10.4018/979-8-3693-7863-2.ch008

Yin, X., Saad, M. R. B. M., & Halim, H. B. A. (2023). A systematic review of critical thinking instructional pedagogies in EFL writing: What do we know from a decade of research. *Thinking Skills and Creativity*, *49*, 101363. https://doi.org/10.1016/j.tsc.2023.101363

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

Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*, *4*, 100134. https://doi.org/10.1016/j.caeai.2023.100134

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments*, *11*(1), 28. https://doi.org/10.1186/s40561-024-00316-7

Zhao, D. (2024). The impact of AI-enhanced natural language processing tools on writing proficiency: an analysis of language precision, content summarization, and creative writing facilitation. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-13145-5

Zhao, Y., Zhao, M., & Shi, F. (2023). Integrating Moral Education and Educational Information Technology: A Strategic Approach to Enhance Rural Teacher Training in Universities. *Journal of the Knowledge Economy*, *15*(3), 15053–15093. https://doi.org/10.1007/s13132-023-01693-z