

AI-Based Cognitive Scaffolding and Academic Writing Performance among EFL Students

 <https://doi.org/10.31004/jele.v11i5.2471>

Lam Quoc Huy, *Nguyen Tien Dung^{ab} 

¹² Hanoi University, Vietnam, Indonesia

Corresponding Author: dungnt@hanu.edu.vn

ABSTRACT

Academic writing is a core competence for learners of English as a foreign language (EFL), yet it remains one of the most demanding skills because it requires the simultaneous integration of content development, idea organization, and linguistic accuracy. Many EFL students consequently devote disproportionate cognitive resources to lower-level language concerns, which may hinder higher-order processes such as coherence building and argument development. Drawing on sociocultural perspectives on cognitive scaffolding and the growing capacity of artificial intelligence (AI) tools to deliver immediate, individualized support, this study investigated whether AI-based cognitive scaffolding could enhance EFL students' academic writing outcomes and whether such effects generalized across different dimensions of writing performance. A quasi-experimental pretest-posttest design was employed over 14 weeks with 53 first-year non-language-major students at a public university in Vietnam. Students' writing performance was compared between a group that received AI-based scaffolding and a group that received conventional instruction. The findings showed that AI-mediated scaffolding produced a significant improvement in the experimental group's overall academic writing performance (gain $\Delta = 0.81$) compared with the control group (gain $\Delta = 0.55$). However, the benefits were selective and domain-specific rather than uniform: the intervention markedly enhanced idea generation and language-related performance but had limited impact on organizational structure and content elaboration. These findings highlight both the strengths and the limitations of AI tools in EFL writing instruction and suggest that AI is most effective when embedded in an integrated pedagogy that combines technological scaffolds with explicit teaching and guided support to address higher-order writing demands.

Keywords: *Academic Writing, AI-Based Scaffolding, Cognitive Scaffolding, EFL Students*

Article History:

Received 03rd September 2026

Accepted 15th October 2026

Published 21st October 2026



INTRODUCTION

Academic writing is a core competence for learners of English as a foreign language (EFL), yet it remains one of the most challenging skills to master. According to Ferris & Hedgcock (2014), effective academic writing requires the integration of multiple components, including content development, organization, and language accuracy. Many EFL learners struggle to manage these tasks and tend to focus their cognitive resources on lower-level concerns such as grammar and vocabulary at the expense of higher-order processes such as argument development, critical thinking, and textual coherence (Zhang et al., 2025). These challenges draw attention to instructional approaches that can provide timely, personalized, and adaptive support during the writing process.

One such approach is cognitive scaffolding, which is rooted in Vygotsky's (1978) sociocultural theory and its emphasis on guided support that helps learners perform beyond their current level of competence. The metaphor of scaffolding was introduced by Wood et al. (1976), and in classroom practice such support is usually provided by teachers through feedback, modelling, or interactive guidance. However, this support is often constrained by time, class size, and variability in learners' needs. With the rapid development of artificial intelligence (AI), interest has grown in leveraging AI-based tools as scalable scaffolding mechanisms in language learning contexts. AI-based tools can function as cognitive supports

by offloading lower-level linguistic processing, assisting with planning and revision, and providing immediate, individualized feedback during writing tasks (Hong et al., 2025; Nguyen et al., 2024; Zulaikha et al., 2025).

Empirical evidence increasingly supports the effectiveness of AI-based scaffolding in enhancing EFL learners' academic writing performance. Across experimental and quasi-experimental studies, AI-assisted interventions have consistently generated significant improvements in overall writing quality, especially in grammatical accuracy, lexical sophistication, coherence, and organization (Najmiah & Al-Anbari, 2025; Shen, 2025; Tran, 2025).

Furthermore, AI-powered writing assistants, including Grammarly, ChatGPT, and QuillBot, have been found to facilitate both lower-order language revision and higher-order writing development by providing immediate and iterative feedback that enhances clarity, syntactic complexity, cohesion, and revision quality (El-Garawany, 2024; Mahapatra, 2024; Song & Song, 2023). Recent studies have extended this evidence by evaluating the pedagogical value of AI-assisted learning tools in different research designs and contexts. Using a mixed-methods design with 60 EFL postgraduate students, [Mekheimer \(2025\)](#) found that Grammarly-assisted feedback not only improved writing proficiency but also promoted students' engagement by reducing frustration and boosting confidence and sense of accomplishment. Likewise, employing a prospective, double-blinded, paired-comparison design, [Tsai et al. \(2024\)](#) demonstrated that ChatGPT-assisted revision significantly improved the quality of essays produced by 44 EFL English majors across all four assessment dimensions, with the highest gain observed in vocabulary, followed by grammar, organization, and content.

Despite the growing body of research on AI-supported writing instruction and cognitive scaffolding in EFL contexts, several gaps remain. First, rigorous experimental research is still insufficient, particularly studies employing pretest–posttest control-group designs that can establish a causal relationship between an AI intervention and writing outcomes. Second, prior research has focused on overall writing performance, with little evaluation of the extent to which AI-based cognitive scaffolding affects specific components of writing, such as content development, organization, lexical choice, and grammatical accuracy. These gaps highlight the need for empirically grounded studies that conceptualize AI as cognitive scaffolding and systematically investigate its impact on the academic writing performance of EFL learners.

Therefore, two research questions (RQs) were formulated:

RQ1. How was students' academic writing performance affected by the application of AI-based cognitive scaffolding?

RQ2. How did the writing components of the experimental group students change after the intervention?

Based on these research questions, three hypotheses (H) were proposed:

H1. AI-based scaffolding positively affected the experimental group students' academic writing performance.

H2. The changes among the writing criteria of the experimental group participants were similar.

H3. The changes in writing components were similar between the experimental and control groups.

Academic Writing in International EFL Contexts and in Vietnam

Academic writing in English remains a persistent challenge for EFL learners worldwide and is a critical concern within the Vietnamese education system, spanning secondary, tertiary, and higher education. Despite differences in educational contexts, a substantial body of research points to a convergence of difficulties across settings, encompassing linguistic, cognitive, affective, and institutional dimensions, alongside context-specific pedagogical and cultural constraints.

At the global level, EFL learners consistently encounter challenges related to limited lexical resources, grammatical inaccuracy, and insufficient control of cohesion and coherence, which collectively hinder the production of academically appropriate texts (Alzahrani, 2025;

Esperanzate, 2025; Singh, 2019). In addition to these linguistic constraints, students frequently struggle with higher-order academic skills, including paraphrasing, argument development, and the integration of sources, all of which are essential for participation in academic discourse communities (Al-Khulaidi & Abdulkhalek, 2022; Pandit, 2025). These difficulties are often compounded by affective factors, such as writing anxiety and low self-efficacy, which further impede students' effective engagement in the writing process.

Empirical studies conducted in various contexts, including Saudi Arabia, Malaysia, and China, report similar patterns of difficulty, especially in relation to adherence to academic conventions, organizational structure, and referencing practices (Alzahrani, 2025; Singh, 2019; Wei et al., 2024). These challenges are frequently attributed to limited opportunities for sustained writing practice and insufficient instructional support tailored to the demands of academic writing. Such findings suggest that the challenges faced by EFL learners are not merely individual but are also shaped by broader institutional and pedagogical conditions.

In response to these challenges, a range of instructional approaches have been proposed and empirically validated. Among the most prominent are genre-based instruction, which explicitly teaches the rhetorical structures of academic texts; collaborative writing, which facilitates peer interaction and shared knowledge construction; and the integration of technology-enhanced tools, such as online collaborative platforms and automated systems (Al-Khulaidi & Abdulkhalek, 2022; Esperanzate, 2025; Hoang & Hoang, 2022; Pandit, 2025). The nature and intensity of these challenges also vary across educational levels. At the secondary level, writing instruction tends to emphasize product-oriented outcomes, often prioritizing grammatical accuracy over rhetorical development (Alzahrani, 2025; Esperanzate, 2025; Singh, 2019). In contrast, higher education imposes more complex demands, requiring students to engage in extended academic writing, including research-based assignments and thesis writing, which necessitate advanced skills in argumentation, synthesis, and critical evaluation (Hang, 2021; Pham & Bui, 2021).

In the Vietnamese EFL context, academic writing has consistently been identified as one of the most challenging language skills for learners at all educational levels. A substantial body of research indicates that students' writing development is constrained by a combination of pedagogical, institutional, and sociocultural factors. Among these, the predominance of grammar-oriented instruction and exam-driven assessment practices has been shown to limit opportunities for meaningful writing practice and individualized feedback (Hang, 2021; Nguyen et al., 2023; Pham & Truong, 2021).

In addition to these structural constraints, writing practices in Vietnam are shaped by broader cultural norms associated with Confucian heritage traditions, including high respect for authority, sensitivity to face, and hierarchical teacher–student relationships. These cultural dimensions can influence both instructional practices and learner behaviors, often resulting in limited peer interaction, reduced critical engagement, and reluctance to provide or receive evaluative feedback (Pham et al., 2020; Trinh & Pho, 2024). Consequently, writing instruction may remain predominantly teacher-centered, with limited emphasis on collaborative or dialogic approaches.

From a pedagogical perspective, research suggests that Vietnamese writing instruction continues to be dominated by form-focused and product-oriented approaches, despite growing awareness among educators of process-based and socially situated models of writing (Pham & Bui, 2021). This tension between traditional practices and contemporary pedagogical orientations highlights a gap between theoretical understanding and classroom implementation, which may further constrain the development of higher-order writing skills. At more advanced levels, including pre-service teacher education and postgraduate study, learners face additional challenges related to academic discourse conventions, such as the use of discipline-specific vocabulary, adherence to genre structure, and maintenance of an appropriate academic tone (Hang, 2021). To address these demands, students tend to employ a range of cognitive, metacognitive, and rhetorical strategies, and there is increasing evidence that self-efficacy and self-regulation support successful writing performance (Bui et al., 2023; Kies & Lambert, 2023; Le & Trinh, 2025).

Several innovative pedagogical approaches have emerged in the Vietnamese context in response to these challenges. They include genre-based instruction, scaffolding techniques, analytic assessment rubrics, and technology-enhanced practices such as electronic peer feedback and online collaborative writing platforms, which are expected to foster improvement in task achievement, content development, language use, and reflective thinking (Bui et al., 2023; Hoang & Hoang, 2022; Pham & Bui, 2021; Wei et al., 2024). However, improvements have not been uniformly distributed across all dimensions of writing; aspects such as grammatical accuracy and textual coherence often show more limited gains (Kies & Lambert, 2023; Phuong et al., 2023).

Cognitive Scaffolding in Writing

Cognitive scaffolding in academic writing refers to forms of instructional support that directly assist learners in processing ideas, managing information, and constructing coherent arguments, rather than merely regulating or motivating their learning behaviors. Grounded in Vygotsky's (1978) Zone of Proximal Development (ZPD) and Sweller's (1988) Cognitive Load Theory (CLT), cognitive scaffolding aims to reduce the complexity of writing tasks by providing structured guidance that enables learners to perform beyond their current level of independent competence. In academic writing contexts, cognitive scaffolding typically takes the form of explicit, task-oriented supports that guide learners through the key stages of the writing process, including guiding prompts and questions, templates, models, and visual tools such as concept maps (Allagui, 2024; Chairinkam & Yawiloeng, 2024; Mitina et al., 2025). All of these serve to facilitate idea generation, organization, and elaboration (Hassen et al., 2024; Khojasteh et al., 2021). Such supports are particularly valuable in EFL contexts, in which learners must simultaneously manage linguistic encoding and higher-order reasoning, which places substantial demands on cognitive resources.

A critical distinction in the literature lies between cognitive scaffolding and other forms of instructional support. According to Allagui (2024), cognitive scaffolding primarily targets problem-solving and content processing, whereas metacognitive scaffolding focuses on learners' ability to plan, monitor, and evaluate their own learning, and motivational scaffolding addresses affective dimensions such as confidence, persistence, and engagement. This differentiation is essential for understanding the specific contribution of cognitive scaffolding to writing development, especially in relation to the construction of meaning and argumentation.

More recently, the concept of cognitive-linguistic scaffolding has gained attention, emphasizing the integration of language forms with disciplinary reasoning practices. In this approach, learners are supported not only in generating ideas but also in articulating them through appropriate rhetorical and linguistic structures, such as claim-evidence-warrant patterns and the use of logical connectives (Mitina et al., 2025). This integration is especially relevant in academic writing, where effective communication depends on the alignment of cognitive and linguistic competence.

Additionally, with the rapid advancement of educational technologies, cognitive scaffolding is increasingly implemented through AI-based systems. These systems provide a range of supports, including automated feedback on grammar and organization, guided prompts for idea development, and interactive co-writing assistance. By offloading lower-level processing demands such as surface-level linguistic corrections, AI-based scaffolding can help reduce cognitive load and enable learners to allocate more attention to higher-order composing processes, including argumentation and critical reasoning (Li & Wilson, 2025).

AI-Based Cognitive Scaffolding

AI-driven tools are conceptualized not merely as assistive technologies but also as cognitive partners that actively support learners throughout multiple stages of the writing process, including planning, idea generation, drafting, and revision (Li & Wilson, 2025). This shift is particularly evident in EFL and ESL contexts across secondary and tertiary education,

where learners often require sustained and individualized support to manage the cognitive demands of writing in a second language (Salleh et al., 2025).

Furthermore, according to Dhillon et al. (2024), AI-based scaffolding provides learners with adaptive and interactive support, enabling them to engage in iterative writing processes that would otherwise be difficult to sustain in traditional classroom settings. By reducing the cognitive burden associated with idea development and organization, such tools allow learners to allocate greater cognitive resources to higher-order composing processes. However, this form of high-level scaffolding also raises significant concerns about learner agency and authorship, as excessive reliance on AI-generated content may diminish students' sense of ownership over their writing (Salleh et al., 2025).

METHOD

Research Design

This study adopted a quasi-experimental pretest–posttest design to evaluate the effects of AI-based cognitive scaffolding on EFL university students' academic writing performance. This design is widely employed in educational research where random assignment of individual students is not feasible, and it allows inferences about intervention effects to be drawn through group comparison. Two intact classes were assigned as an experimental group, which received AI-based cognitive scaffolding, and a control group, which received conventional instruction.

A pre-test was administered during the first three weeks, after some input on foundational knowledge of various academic writing genres, especially the argumentative essay, in order to establish baseline equivalence between the two groups and to control for potential initial differences in writing proficiency. The instructional intervention was implemented from week 4 to week 13. Following the 10-week intervention, a post-test was conducted in week 14 to assess both within-group development and between-group differences.

To enhance internal validity, the same instructional duration, writing tasks, and assessment procedures were applied to both groups, with the mode of instructional support (AI-based versus conventional) being the only systematic variation.

Participants

The participants were 53 first-year EFL students aged 18–19 years (18 males and 35 females). Academic writing was one of the core subjects in their program, with a 2.5-hour academic writing class per week over 14 weeks. The participants were selected through intact-class sampling, a common approach in quasi-experimental studies in educational settings. All participants had a relatively homogeneous level of English proficiency (approximately B2 on the CEFR), as determined by the admission marks for their chosen major, which helped reduce variability attributable to proficiency differences. The experimental class included 33 students, and the control class had 20 students. All participants provided informed consent, and ethical considerations regarding data confidentiality and voluntary participation were strictly observed.

Instructional intervention

AI-Based Cognitive Scaffolding

The experimental group engaged in AI-supported writing tasks designed to function as cognitive scaffolding, facilitating learners' idea generation, organization, and linguistic development during the writing process. The intervention was grounded in Vygotsky's (1978) ZPD, in which temporary support is provided to help learners perform tasks beyond their current level of independent capability. In this study, AI functioned as a responsive scaffold that offered immediate and individualized support. Students were therefore instructed to use AI tools in four key cognitive processes, which were assessed through their academic writing performance: (1) Idea generation: creating relevant arguments and supporting ideas. (2) Organizational structuring: developing outlines and a logical sequence of ideas. (3) Content

elaboration: expanding arguments with explanations and examples. (4) Language enhancement: improving lexical resources and grammatical structures.

To ensure that AI served as a learning aid rather than a content generator, students were instructed to critically evaluate and revise AI-generated suggestions. This approach is aligned with the conceptualization of AI as a cognitive assistant rather than a substitute for learning effort.

Control Condition

The control group received conventional writing instruction, including teacher-led explanations, peer discussions, and guided practice, without the assistance of AI tools. This condition allowed a comparison between technology-mediated scaffolding and conventional pedagogical support.

Instructional Procedure

Table 1 summarizes the focus, activities, and AI tools used in each phase of the 14-week program.

Table 1. Instructional Procedure of The Experimental Group

Week	Focus	Activities	AI tools used
1-3	Orientation and baseline writing	• Introduction to critical thinking and academic writing. • Introduction to the argumentative essay and assessed criteria. • Pre-test administered.	Google Docs
4-6	Analyze and reflect	• Introduction to 3R (Role-Request-Result) prompts. • Students' writing using AI feedback.	ChatGPT, QuillBot, Google Docs
7-9	Evaluate and argue	• Peer- and AI-supported critique. • Students' writing using AI-supported critique.	ChatGPT
10-13	Rebut and strengthen	• Introduction to AI-critique prompts to support counterarguments. • Students' writing using AI-critique prompts.	ChatGPT, teacher-led debriefing
14	Final revision	• Revision. • Post-test administered.	AI logs, Grammarly, reflective templates

Weekly Instruction And AI-Based Activities

The AI-based cognitive scaffolding intervention was implemented over 10 weeks (weeks 4-13), following a three-week orientation in which students were introduced to academic writing, the argumentative essay, and the marking rubric. During the intervention, students attended a 2.5-hour teacher-guided writing lesson per week, supplemented by one AI-supported homework assignment designed to reinforce the writing strategies introduced in class. Consequently, the experimental group engaged in approximately 20 scaffolded writing activities during the intervention.

The scaffolding was progressively structured from lower-order to higher-order writing processes. In weeks 4-6, students were instructed to use the 3R (Role-Request-Result) prompting technique to generate ideas, refine thesis statements, and receive feedback on grammar and vocabulary. In weeks 7-9, the focus shifted to peer-assisted and AI-assisted critique, in which students evaluated draft essays using AI-generated feedback together with peer discussion to foster reflective revision. During weeks 10-13, students were trained to use AI-critique prompts to generate counterarguments and strengthen their supporting ideas. Week 14 was devoted to final revision and post-test administration. The time allowed for the intervention in each writing session was 15 to 20 minutes.

Throughout the intervention, AI-assisted learning tools functioned as a cognitive scaffold rather than an automated writing tool. Students were required to interact iteratively with ChatGPT and Grammarly by generating multiple prompts, evaluating AI-generated

suggestions, revising their prompts, and justifying whether to accept, modify, or reject the feedback. Teachers continuously monitored the classroom, modelled effective prompts, provided explicit instruction on critically evaluating AI-generated responses, and interviewed participants about the reasons for their choices in order to prevent overreliance on AI-generated results. This guided approach was expected to develop students' cognitive regulation and critical AI literacy while ensuring that students remained responsible for planning, organizing, and revising their own writing.

Data Collection Instruments

Writing Assessment Tasks

Participants completed two argumentative essay tasks as pre-test and post-test measures. The tasks were designed to be equivalent in topic complexity, rhetorical demand, and length requirements to ensure measurement consistency. Argumentative writing was selected on the basis of the participants' curriculum and its cognitive demands, including idea development, logical reasoning, and linguistic accuracy, and it was considered particularly suitable for evaluating the effects of AI-based cognitive scaffolding.

Reliability And Rater Training

Because of the similarities in the marking and evaluation features of IELTS Writing Task 2, the IELTS writing criteria were employed to assess students' writing performance across four dimensions: content, organization, vocabulary, and grammar. The analytic rubric enabled a fine-grained analysis of specific writing components and a more differentiated understanding of how AI-based cognitive scaffolding influenced different aspects of writing performance.

Table 2 shows the internal consistency of the adapted marking rubric, with a Cronbach's alpha coefficient of $\alpha = 0.978$ across 10 items. Such a high coefficient indicates that the components of the scoring rubric were strongly interrelated and functioned cohesively as a measurement instrument.

Table 2. Reliability of the Marking Rubric

Statistic	Value
Cronbach's alpha	0.978
Number of items	10

Additionally, to ensure the reliability of the scoring process, three experienced raters who were qualified as VSTEP (Vietnamese Standardized Test of English Proficiency) examiners participated in the assessment. Each test was marked independently by two raters. When the score discrepancy exceeded 0.5, the third rater was consulted, and the final score was determined on the basis of the adjudicated ratings. Prior to rating, the raters underwent a calibration session to ensure consistency in rubric interpretation. Before the pre-test and post-test were implemented, a pilot study involving 10 volunteer participants, who were not included in the final sample, was conducted to examine the tests and inter-rater reliability. These steps were undertaken to ensure that the rating process achieved acceptable reliability thresholds.

Table 3 presents inter-rater reliability as assessed by the intraclass correlation coefficient (ICC). The single-measures ICC indicated a high level of agreement between individual raters, suggesting that each rater produced reliable test scores. More importantly, the average-measures ICC demonstrated excellent reliability, implying that when the raters' scores were combined, the results were highly stable and dependable.

Table 3. Intraclass Correlation Coefficient (ICC)

Measure type	ICC	95% CI lower	95% CI upper	F	df1	df2	Sig.
Single measures	0.813	0.650	0.938	44.614	9	81	0.000
Average measures	0.978	0.949	0.993	44.614	9	81	0.000

Data Collection Procedure

Data collection was conducted over four stages: (1) **Pilot testing.** Prior to the pre-test, pilot testing was conducted with 15 volunteer students to examine the validity of both the pre-test and post-test. The pilot sessions were implemented in weeks 1 and 2 with students who were not involved in the main study. These students were adequately familiar with various forms of argumentative essays, ensuring that their performance would provide a meaningful evaluation of the test instruments. (2) Pre-test phase. All participants completed an argumentative writing task under standardized conditions to establish baseline performance in week 3. (3) Intervention phase. Over 10 weeks, the experimental group engaged in AI-based writing activities, while the control group followed traditional instructional practices. (4) Post-test phase. Participants completed a second argumentative writing task under comparable conditions to assess changes in writing performance in week 14.

All writing tasks were administered in a controlled classroom setting to ensure consistency between the two tests.

Data Analysis

Preliminary Analysis

Descriptive statistics, including means (M) and standard deviations (SD), were calculated to summarize writing performance across groups and testing occasions. Prior to inferential analysis, the assumptions of normality, homogeneity of variance, and linearity were examined to ensure the appropriateness of the analyses.

Inferential Analysis

To evaluate the effectiveness of the intervention, the following analyses were implemented: (1) One-way ANOVA was conducted to compare pre-test performance between the experimental and control groups and to determine whether there were statistical differences in the overall pre-test scores and the assessed components. (2) Because of discrepancies between the groups in pre-test results, one-way multivariate analysis of covariance (MANCOVA) was employed to compare post-test scores on the four writing components while controlling for pre-test scores, followed by univariate tests for each dependent variable. (3) Partial eta squared (partial η^2) was calculated to identify the magnitude of the intervention effects, providing an interpretation that goes beyond statistical significance.

Ethical approval was obtained prior to data collection. Participants were informed of the purposes of the study, their participation was voluntary, and all data were anonymized to ensure confidentiality.

FINDINGS AND DISCUSSIONS

Pre-Intervention Evaluations

The purpose of the pre-intervention comparison was to assess students' academic writing proficiency and related writing components before the intervention. At this stage, the equivalence of overall writing performance, the equality of variances, and the differences in writing components between the two groups were examined. A series of one-way ANOVAs was conducted to evaluate pre-test differences between the experimental and control groups.

Table 4. Descriptive Statistics of The Pre-Test By Group

Variable	Group	N	Mean (M)	Std. deviation (SD)
Idea generation	Control	30	6.77	0.94
	Experimental	23	7.00	1.04
Organizational structuring	Control	30	6.13	0.82
	Experimental	23	6.39	1.23
Content elaboration	Control	30	5.75	0.99
	Experimental	23	6.17	1.11
Language enhancement	Control	30	6.10	0.92
	Experimental	23	5.57	0.90
Total score	Control	30	6.19	0.78
	Experimental	23	6.28	0.97

Note. Group labels follow the original data file (Group 1 = control, Group 2 = experimental).

Table 4 presents the descriptive statistics of the pre-test scores for the control group (n = 30) and the experimental group (n = 23) across the four dimensions of academic writing and the overall writing score. Overall, both groups showed comparable baseline performance before the intervention, with some modest variations in mean scores across the assessed components.

Specifically, the experimental group obtained slightly higher mean scores in idea generation (M = 7.00, SD = 1.04 vs. M = 6.77, SD = 0.94), organizational structuring (M = 6.39, SD = 1.23 vs. M = 6.13, SD = 0.82), and content elaboration (M = 6.17, SD = 1.11 vs. M = 5.75, SD = 0.99), whereas the control group obtained a marginally higher mean score in language enhancement (M = 6.10, SD = 0.92) than the experimental group (M = 5.57, SD = 0.90). Despite these minor differences, the overall pre-test means were highly comparable between the control group (M = 6.19, SD = 0.78) and the experimental group (M = 6.28, SD = 0.97). The SDs across all writing dimensions ranged from 0.78 to 1.23, indicating a relatively similar degree of variability in writing performance within the two groups.

Table 5. Test Of Equality of Variances (Pre-Test)

Variable	F	p-value	Interpretation
Idea generation	0.169	0.683	Assumption met
Organizational structuring	4.167	0.046	Slight violation
Content elaboration	0.398	0.531	Assumption met
Language enhancement	0.010	0.921	Assumption met
Total score	0.680	0.413	Assumption met

Table 5 presents the results of the test of equality of variances. The assumption of homogeneity of variance was met for most variables ($p > 0.05$), indicating comparable variances across the groups. For organizational structuring, however, the result ($p = 0.046$) suggested a difference in variability between the groups on this component.

Table 6. Tests of Between-Subject Effects (Pre-Test)

Variable	F	p-value	Interpretation
Idea generation	0.732	0.396	Not significant
Organizational structuring	0.835	0.365	Not significant
Content elaboration	2.143	0.149	Not significant
Language enhancement	4.484	0.039	Significant
Total score	0.155	0.695	Not significant

Table 6 shows the pre-test differences between the control and experimental groups across the five measures. Overall, the analysis revealed no statistically significant differences

between the two groups on four of the five measures. However, one statistically significant difference was found in language enhancement, $F(1, 51) = 4.484$, $p = 0.039$, suggesting an initial imbalance in this specific component. This finding provided the basis for employing multivariate analysis of covariance (MANCOVA) to examine the differences between the experimental and control groups after the intervention.

Post-Intervention Evaluations

In this phase, a series of tests was implemented and analyzed to identify within-group and between-group changes.

Table 7. Descriptive Statistics By Group (M $\pm$ SD)

Criterion	Control group			Experimental group		
	Pre	Post	Gain	Pre	Post	Gain
Idea generation	6.70 $\pm$ 0.94	6.98 $\pm$ 0.78	+0.28	7.00 $\pm$ 1.04	7.57 $\pm$ 0.65	+0.57
Organizational structuring	6.09 $\pm$ 0.82	6.78 $\pm$ 0.75	+0.69	6.39 $\pm$ 1.23	7.00 $\pm$ 0.86	+0.64
Content elaboration	5.59 $\pm$ 0.99	6.33 $\pm$ 0.92	+0.74	6.17 $\pm$ 1.11	6.89 $\pm$ 0.80	+0.72
Language enhancement	5.96 $\pm$ 0.92	6.44 $\pm$ 0.73	+0.48	5.57 $\pm$ 0.90	6.80 $\pm$ 0.58	+1.32
Total	6.08 $\pm$ 0.78	6.63 $\pm$ 0.69	+0.55	6.28 $\pm$ 0.97	7.09 $\pm$ 0.58	+0.81

Note. Gain = post-test mean – pre-test mean.

The descriptive statistics in Table 7 indicate that both the control and experimental groups improved on all writing criteria from pre-test to post-test; however, the magnitude of the gains differed between the groups. The experimental group outperformed the control group in most dimensions, particularly in idea generation (+0.57 vs. +0.28) and language enhancement, for which a markedly larger improvement was observed (+1.32 vs. +0.48). These substantial gains in language suggest that the intervention had a relatively strong effect on linguistic accuracy and expression. Interestingly, organizational structuring and content elaboration showed relatively similar gains across the two groups (+0.64 vs. +0.69 and +0.72 vs. +0.74, respectively), implying that these skills may be less sensitive to the intervention.

Table 8. Multivariate Tests of The Group Effect

Effect	Pillai's trace	F	df (hyp., error)	p	Partial η^2
Group	0.447	8.902	(4, 44)	< 0.001	0.447

A one-way MANCOVA was conducted to examine the effect of group on post-test writing performance while controlling for pre-test scores. The multivariate test indicated a significant effect of group, Pillai's trace = 0.447, $F(4, 44) = 8.902$, $p < 0.001$, partial $\eta^2 = 0.447$. This means that 44.7% of the variance in the combined dependent variables (idea generation, organizational structuring, content elaboration, and language enhancement) could be attributed to group membership when baseline differences were accounted for, which represents a large effect according to conventional benchmarks. The significant multivariate effect indicated that the intervention had a meaningful and systematic impact on learners' overall academic writing performance compared with the control condition, and it justified follow-up univariate tests to identify which writing components contributed to the multivariate difference.

Table 9. Tests Of Between-Subject Effects (Post-Test, Controlling for Pre-Test)

Dependent variable	F	df	p	Partial η^2
Idea generation	10.706	(1, 47)	0.002	0.186
Organizational structuring	1.066	(1, 47)	0.307	0.022
Content elaboration	0.211	(1, 47)	0.648	0.004
Language enhancement	20.590	(1, 47)	< 0.001	0.305
Total score	9.077	(1, 47)	0.004	0.162

Table 9 reveals a differentiated pattern of group effects across the writing components. After controlling for the pre-test score, a statistically significant effect of group was found for idea generation, $F(1, 47) = 10.706$, $p = 0.002$, partial $\eta^2 = 0.186$, indicating a large effect. Similarly, a highly significant effect emerged for language enhancement, $F(1, 47) = 20.590$, $p < 0.001$, partial $\eta^2 = 0.305$, also a large effect. In addition, the total writing score differed significantly between the groups, $F(1, 47) = 9.077$, $p = 0.004$, partial $\eta^2 = 0.162$, suggesting an overall advantage for the experimental group.

In contrast, no statistically significant group differences were observed for organizational structuring, $F(1, 47) = 1.066$, $p = 0.307$, partial $\eta^2 = 0.022$, or content elaboration, $F(1, 47) = 0.211$, $p = 0.648$, partial $\eta^2 = 0.004$. The effect sizes for these variables were small, indicating that the intervention did not meaningfully influence these aspects of writing.

In general, these findings suggest that the intervention was particularly effective in enhancing idea generation and language-related performance, while its impact on organizational structure and content elaboration was limited. The significant improvement in the total score further confirmed the overall effectiveness of the intervention despite its domain-specific variation in outcomes.

Discussions

The present study investigated the effects of AI-based cognitive scaffolding on EFL students' academic writing performance, with particular attention to both overall outcomes and changes in specific writing components.

Overall Effects of AI-Based Cognitive Scaffolding on Academic Writing Performance (RQ1 And H1)

This study examined whether AI-based cognitive scaffolding could enhance EFL students' academic writing performance. The findings provided strong empirical support for H1, demonstrating that AI-based cognitive scaffolding significantly improved students' writing outcomes. The MANCOVA results indicated a strong multivariate effect of group on post-test performance after controlling for pre-test scores, which was necessary because of the non-equivalent baseline identified in the pre-test. This suggests that the intervention contributed substantially to learners' writing development. The large effect size (partial $\eta^2 = 0.447$) further confirmed that the impact was not only statistically significant but also educationally meaningful.

These findings are consistent with a growing body of research indicating that AI tools can function as cognitive scaffolding, supporting learners in generating, organizing, and refining ideas rather than merely correcting surface-level errors (Cer, 2019). Prior studies (Mahapatra, 2024; Song & Song, 2023) showed that tools such as ChatGPT provide meaningful writing models and immediate feedback that enhance idea development, grammatical accuracy, and structural coherence. Similarly, AI systems such as Grammarly and QuillBot have been found to improve overall writing proficiency, particularly language use and coherence (Liu et al., 2023; Pratama et al., 2025). Nevertheless, whereas previous research often reported general improvement across multiple writing dimensions, the present study revealed a more differentiated pattern. Despite significant improvement in overall performance, the gains were not equally distributed across all writing components. This suggests that the effectiveness of AI-based cognitive scaffolding may depend on how different writing

components interact with the affordances of AI tools and with the cognitive abilities of the writers.

Changes Among the Writing Components of the Experimental Group (RQ2 And H2)

In response to the second research question, the findings indicated that the changes in writing components were not homogeneous, which led to the rejection of H2. Specifically, the experimental group's advantage over the control group was significant for idea generation and language enhancement, whereas it was not significant for organizational structuring and content elaboration (Table 9).

This domain-specific pattern both aligned with and extended previous literature. Consistent with the findings of [Mahapatra \(2024\)](#) and [Mekheimer \(2025\)](#), idea generation and language enhancement showed significantly greater improvement than the other writing dimensions. This differential effect can be attributed to the varying cognitive demands of the writing components and to the affordances of AI-based cognitive scaffolding. Specifically, AI was particularly effective in supporting idea generation and language formulation by providing brainstorming prompts, alternative phrasing, grammatical corrections, and immediate feedback. Such support reduces the cognitive load associated with lexical retrieval and linguistic encoding (Sweller, 1988), while the iterative and conversational nature of the interaction facilitates continuous refinement of ideas within learners' ZPD (Vygotsky, 1978).

In contrast, the absence of significant improvement in organizational structuring and content elaboration suggests that these dimensions impose substantially greater cognitive demands than the others. Although AI-assisted learning tools can provide organizational suggestions and outline templates, learners must still critically evaluate, adapt, and integrate these suggestions into coherent arguments. Such complex decision-making depends heavily on learners' prior knowledge, disciplinary understanding, and metacognitive regulation, which cannot be fully externalized through AI-assisted tools. This finding contrasts with [Hwang et al. \(2023\)](#), who reported significant improvement in organizational structure and content development using the RoamLingo system. From a sociocultural perspective, scaffolding facilitates performance but does not fully compensate for insufficient background knowledge and conceptual understanding (Vygotsky, 1978; Wood et al., 1976).

Between-Group Differences and Differential Gains (H3)

The third hypothesis predicted that the changes in the writing components of the experimental and control groups would be similar. The findings, however, did not support this assumption. Instead, the results indicate differential, domain-specific gains between the groups (Table 9). The experimental group demonstrated significantly greater improvement in idea generation, language enhancement, and overall writing performance than the control group, whereas both groups showed comparable outcomes in organization and content. This pattern suggests that AI-based cognitive scaffolding introduced targeted advantages rather than uniform improvements across all writing dimensions. These findings partially align with previous research reporting large overall gains from AI-assisted writing tools, but they also highlight the refinement that improvements are not necessarily distributed equally (Tuong & Tran, 2025; Zhang et al., 2025; Zhao, 2025). The impact of AI may therefore be selective and contingent on the nature of the writing components.

AI as Cognitive Scaffolding

The findings provide empirical support for conceptualizing AI as a form of cognitive scaffolding that operates across multiple stages of the writing process. In particular, AI tools appeared to facilitate idea generation by offering prompts, examples, and alternative perspectives. They also supported language formulation through feedback on grammar, vocabulary, and phrasing. Finally, AI tools enabled iterative revision, reducing cognitive load and promoting self-regulation. These mechanisms align with the sociocultural and scaffolding theories (Vygotsky, 1978; Wood et al., 1976), which emphasize the role of external supports in extending learners' capacities. However, the limited impact on higher-order skills suggests

that AI-assisted cognitive scaffolding may be insufficient to support complex discourse-level processes without additional mediation. Thus, while AI can effectively scaffold certain cognitive processes, its impact is contingent on the type of cognitive process involved. This highlights the importance of integrating AI tools within a broader pedagogical framework that explicitly addresses higher-order writing skills.

CONCLUSION

Overall, the findings demonstrated that AI-based cognitive scaffolding significantly improved students' academic writing performance in EFL classrooms, particularly in idea generation and language enhancement, while gains in organizational structuring and content elaboration were comparatively modest. These results suggest that AI primarily supports iterative refinement and linguistic assistance, whereas complex discourse-level processes require explicit instructional mediation. The study contributes to AI-assisted language learning by demonstrating how AI can selectively support different stages of academic writing without replacing teachers' instruction. This finding aligns with Vygotsky's (1978) sociocultural perspective, which emphasizes the role of external support in facilitating learning within meaningful instructional contexts. However, the study's small sample from a single public institution limits generalizability. Additionally, it did not examine long-term effects or directly assess learners' cognitive changes and AI interaction behaviors during writing. Future research should investigate long-term outcomes across diverse educational contexts and proficiency levels, incorporating process-oriented data, such as AI interaction logs, to clarify how instructional scaffolding and AI tools jointly support higher-order writing skills.

REFERENCES

- Al-Khulaidi, M. A., & Abdulkhalek, M. M. (2022). Academic writing problems in L2 settings: Realities and need for intervention. *Journal of English Studies in Arabia Felix*, 1(1), 42–51. <https://doi.org/10.56540/jesaf.v1i1.15>
- Allagui, B. (2024). A scaffolding intervention to improve self-efficacy in source-based argumentative writing. *Frontiers in Psychology*, 15, 1–14. <https://doi.org/10.3389/fpsyg.2024.1454104>
- Alzahrani, G. (2025). Academic writing difficulties of EFL learners at Najran University. *Journal of Educational and Social Research*, 15(3), 217–229. <https://doi.org/10.36941/jesr-2025-0093>
- Bui, H. P., Nguyen, L. T., & Nguyen, T. V. (2023). An investigation into EFL pre-service teachers' academic writing strategies. *Heliyon*, 9(2), 1–11. <https://doi.org/10.1016/j.heliyon.2023.e13743>
- Cer, E. (2019). The instruction of writing strategies: The effect of the metacognitive strategy on the writing skills of pupils in secondary education. *SAGE Open*, 9(2), 1–17. <https://doi.org/10.1177/2158244019842681>
- Chairinkam, J., & Yawiloeng, R. (2024). The use of scaffolding strategies to enhance the writing development of EFL students. *Theory and Practice in Language Studies*, 14(9), 2996–3007. <https://doi.org/10.17507/tpls.1409.35>
- El-Garawany, M. S. M. (2024). The effects of a QuillBot-based intervention on English language majors' EFL writing performance, apprehension, and self-efficacy. *Language Teaching Research Quarterly*, 43, 167–189. <https://doi.org/10.32038/ltrq.2024.43.10>
- Esperanzate, M. D. (2025). Teaching academic writing in English to ESL/EFL learners in higher education: A systematic review of strategies and challenges. *Journal of Exceptional Multidisciplinary Research*, 2(2), 45–54. <https://doi.org/10.69739/jemr.v2i2.919>
- Ferris, D. R., & Hedgcock, J. S. (2014). *Teaching L2 composition: Purpose, process, and practice* (3rd ed.). Routledge. <https://doi.org/10.4324/9780203813003>
- Hang, N. T. T. (2021). Vietnamese upper-high school teachers' views, practices, difficulties, and expectations on teaching EFL writing. *Journal on English as a Foreign Language*, 11(1),

- 1–20. <https://doi.org/10.23971/jefl.v11i1.2228>
- Hassen, S. K., Adugna, E. T., & Bogale, Y. N. (2024). EFL students' writing achievement via the execution of scaffolding strategies instruction. *System*, 125, 103439. <https://doi.org/10.1016/j.system.2024.103439>
- Hoang, T. N. D., & Hoang, T. (2022). Enhancing EFL students' academic writing skills in online learning via Google Docs-based collaboration: A mixed-methods study. *Computer Assisted Language Learning*, 37(12), 1504–1526. <https://doi.org/10.1080/09588221.2022.2083176>
- Hong, H., Vate-U-Lan, P., & Viriyavejakul, C. (2025). Cognitive offload instruction with generative AI: A quasi-experimental study on critical thinking gains in English writing. *Forum for Linguistic Studies*, 7(7). <https://doi.org/10.30564/fls.v7i7.10072>
- Hwang, W.-Y., Nurtantyana, R., Purba, S., Hariyanti, U., Indrihapsari, Y., & Surjono, H. (2023). AI and recognition technologies to facilitate English as foreign language writing for supporting personalization and contextualization in authentic contexts. *Journal of Educational Computing Research*, 61(4), 1008–1035. <https://doi.org/10.1177/07356331221137253>
- Khojasteh, L., Hosseini, S. A., & Nasiri, E. (2021). The impact of mediated learning on the academic writing performance of medical students in flipped and traditional classrooms: Scaffolding techniques. *Research and Practice in Technology Enhanced Learning*, 16(17), 1–20. <https://doi.org/10.1186/s41039-021-00165-9>
- Kies, S., & Lambert, O. (2023). Acquiring a scholar's voice: Vietnamese students mastering academic vocabulary in thesis writing. *Literatura y Lingüística*, 46, 251–279. <https://doi.org/10.29344/0717621X.46.3145>
- Le, T. A., & Trinh, N. T. (2025). A student-centered framework for understanding EFL thesis writing difficulties in Vietnam. *International Journal of Learning, Teaching and Educational Research*, 24(11), 314–333. <https://doi.org/10.26803/ijlter.24.11.15>
- Li, M., & Wilson, J. (2025). AI-integrated scaffolding to enhance agency and creativity in K-12 English language learners: A systematic review. *Information*, 16(7), 519. <https://doi.org/10.3390/info16070519>
- Liu, C., Hou, J., Tu, Y., Wang, Y., & Hwang, G.-J. (2023). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments*, 31(9), 5614–5632. <https://doi.org/10.1080/10494820.2021.2012812>
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11, 9. <https://doi.org/10.1186/s40561-024-00295-9>
- Mekheimer, M. (2025). Generative AI-assisted feedback and EFL writing: A study on proficiency, revision frequency and writing quality. *Discover Education*, 4(1), 170. <https://doi.org/10.1007/s44217-025-00602-7>
- Mitina, T., Afanasenkova, I., Bazarnova, N., Shaikhova, B., & Oyamo, M. (2025). Cognitive-linguistic scaffolding as a tool for activating argumentative writing skills in the process of teaching analytical chemistry in English language. *International Journal of Innovative Research and Scientific Studies*, 8(2), 2171–2186. <https://doi.org/10.53894/ijirss.v8i2.5645>
- Najmiah, L., & Al-Anbari, A. H. J. (2025). Narrative explorations of EFL learners' engagement with AI tools in developing writing proficiency. *DUTIES: Education and Humanities International Journal*, 1(2), 42–62. <https://doi.org/10.70152/duties.v1i2.222>
- Nguyen, B. M., Ta, H. H., & Nguyen, L. T. P. (2023). English teachers' perception and attitudes of scaffolding method in teaching academic writing: An exploratory study at universities in Vietnam. *Journal of Education and E-Learning Research*, 10(3), 371–379. <https://doi.org/10.20448/jeelr.v10i3.4717>
- Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education*, 49(5), 847–864. <https://doi.org/10.1080/03075079.2024.2323593>
- Pandit, B. R. (2025). Issues and practices of academic writing in English in the global context. *Journal of NELTA Gandaki*, 8(1-2), 128–142. <https://doi.org/10.3126/jong.v8i1-2.85077>

- Pham, N. T., Lin, M., Trinh, V. Q., & Bui, T. P. L. (2020). Electronic peer feedback, EFL academic writing and reflective thinking: Evidence from a Confucian context. *SAGE Open*, 10(1), 1–20. <https://doi.org/10.1177/2158244020914554>
- Pham, V. P. H., & Bui, T. K. L. (2021). Genre-based approach to writing in EFL contexts. *World Journal of English Language*, 11(2), 95–106. <https://doi.org/10.5430/wjel.v11n2p95>
- Pham, V. P. H., & Truong, M. H. (2021). Teaching writing in Vietnam's secondary and high schools. *Education Sciences*, 11(10), 632. <https://doi.org/10.3390/educsci11100632>
- Phuong, H. Y., Phan, Q. T., & Le, T. T. (2023). The effects of using analytical rubrics in peer and self-assessment on EFL students' writing proficiency: A Vietnamese contextual study. *Language Testing in Asia*, 13(1), 1–19. <https://doi.org/10.1186/s40468-023-00256-y>
- Pratama, Y., Fridolini, F., & Pitaloka, R. M. (2025). The impact of AI QuillBot in improving student writing ability to write argumentative essays. *English Review: Journal of English Education*, 13(1), 215–222. <https://doi.org/10.25134/erjee.v13i1.9746>
- Salleh, A. S., Ismail, N. L. M., & Rahman, E. A. (2025). Scaffold ED AI: Harnessing AI platforms for scaffolding academic writing in secondary and tertiary ESL classroom. *International Journal of Research and Innovation in Social Science*, 9(25), 300–305. <https://doi.org/10.47772/IJRISS.2025.925ILEIID000054>
- Shen, Y. (2025). AI-driven scaffolding and affective support in ESL argumentative writing: A multimodal analytics approach. In *Proceedings of the 2025 6th International Conference on Education, Knowledge and Information Management* (pp. 604–608). <https://doi.org/10.1145/3756580.3756680>
- Singh, M. (2019). Academic reading and writing challenges among international EFL Master's students in a Malaysian university: The voice of lecturers. *Journal of International Students*, 9(3), 972–992. <https://doi.org/10.32674/jis.v9i3.934>
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285.
- Tran, T. T. T. (2025). Enhancing EFL writing revision practices: The impact of AI- and teacher-generated feedback and their sequences. *Education Sciences*, 15(2), 232. <https://doi.org/10.3390/educsci15020232>
- Trinh, K. N., & Pho, P. D. (2024). English academic writing through the lens of culture: Implications for current practices in Vietnam. *Science and Technology Development Journal*, 27(SI), 56–63. <https://doi.org/10.32508/stdj.v27isi.4409>
- Tsai, C.-Y., Lin, Y.-T., & Brown, I. K. (2024). Impacts of ChatGPT-assisted writing for EFL English majors: Feasibility and challenges. *Education and Information Technologies*, 29(17), 22427–22445. <https://doi.org/10.1007/s10639-024-12722-y>
- Tuong, T. P. L., & Tran, T. T. (2025). The differential impact of AI tools among EFL university learners: A process writing approach. *International Journal of Learning, Teaching and Educational Research*, 24(5), 452–471. <https://doi.org/10.26803/ijlter.24.5.24>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wei, Y., Sulaiman, N. A., & Ismail, H. H. (2024). Overcoming academic writing challenges: English language learning strategies for Chinese international high school students in blended learning environments. *Theory and Practice in Language Studies*, 14(10), 3285–3294. <https://doi.org/10.17507/tpls.1410.32>
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- Zhang, H., Chen, C., Liu, Y., & Peng, L. (2025). From dependence to development: The role of AI in enhancing writing proficiency among Chinese English majors. *Journal of Posthumanism*, 5(8), 293–300. <https://doi.org/10.63332/joph.v5i8.3159>
- Zhao, D. (2025). 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*, 30(6), 8055–8086. <https://doi.org/10.1007/s10639-024-13145-5>

Zulaikha, Z., Setyawan, C. E., Maburi, M., & Rauhillah, S. (2025). The effectiveness of artificial intelligence and deep learning tools in enhancing academic journal writing: A mixed methods study of Arabic language education students in Indonesia. *Transformasi*, 21(1). <https://doi.org/10.56357/jt.v21i1.429>