Voices of English Language Education Society
| Volume 10, Number 1| | April 2026 | | E-ISSN 2579-7484 |
DOI: https://doi.org/10.29408/veles.v10i1.31762

Can AI Prompts Scaffold Critical Thinking in EFL Academic Writing? A Qualitative Case Study

[*1]Muhammad Deni Irkhamil Maulana, [1]Agus Widyantoro

[1]Universitas Negeri Yogyakarta, Indonesia

Abstract: Generative AI is increasingly used in EFL academic writing, yet its pedagogical value depends on whether students use it as a tool for thinking or merely as a shortcut for text production. This study explored how generative AI prompts functioned as cognitive scaffolding in EFL students' academic writing, particularly in relation to idea development, writing organization, engagement, and critical thinking. Using a qualitative case study design, the study involved 10 undergraduate EFL students enrolled in a second-semester Academic Writing course at a university in Malang, Indonesia. Data were collected through classroom observations, screen-recorded writing sessions, students' writing documents, semi-structured interviews, and reflection journals, and were analyzed using reflexive thematic analysis. The findings revealed that generative AI prompts supported students by activating initial ideas, clarifying concepts, and helping them organize writing into more coherent and logical structures. AI prompts also reduced hesitation and increased students' confidence during the writing process by offering immediate support when they felt stuck. However, the findings also showed a risk of overreliance, as some students copied or minimally revised AI-generated text without sufficient evaluation. More reflective students used AI prompts to compare ideas, identify weaknesses, evaluate paragraph connections, and revise arguments more purposefully. These findings suggest that generative AI prompts can serve as cognitive scaffolding when they encourage planning, evaluation, and reflective revision, but they may weaken students' intellectual ownership when used passively. To maximize their value in EFL academic writing, AI-assisted instruction should include explicit prompt literacy, critical evaluation of AI responses, revision-based tasks, and teacher-guided reflection so that students write with AI rather than by AI.

Keywords: generative AI, AI prompts, cognitive scaffolding, EFL academic writing, critical thinking, student engagement.

1. Introduction

The rapid emergence of generative artificial intelligence in education has created a new tension in academic writing instruction: AI can support students' thinking, yet it can also encourage dependence on machine-generated text. In EFL academic writing, this tension is especially significant because students often face difficulties not only in grammar and vocabulary but also in generating ideas, organizing arguments, developing coherence, and expressing a critical academic stance. When used productively, generative AI may help students overcome initial writing barriers by offering prompts, examples, outlines, explanations, and alternative formulations. However, when used passively, it may reduce students' intellectual engagement, weaken authorship, and turn writing into a process of accepting ready-made responses. Therefore, the central issue is no longer whether generative AI should be present in EFL writing classrooms, but how

***Correspondence:** mdeni6539@gmail.com

Submitted: November 9, 2025 Revised: Feberuary 1, 2026 Accepted, March 3, 2026

students engage with it during the writing process and whether its support strengthens or replaces their own thinking.

Academic writing in EFL contexts has long been recognized as a demanding intellectual practice. Students are required to construct ideas, justify claims, organize arguments, and develop coherent texts rather than merely produce grammatically acceptable sentences (Altınmakas & Bayyurt, 2018; Ariyanti & Fitriana, 2017; Chuang & Yan, 2022). Many EFL students continue to produce writing that is descriptive, weakly structured, and insufficiently argumentative, suggesting that their difficulties are not only linguistic but also cognitive and rhetorical (Bulqiyah et al., 2021; Nejmaoui, 2019; Toba et al., 2019). In argumentative and academic writing, students must move from general ideas to defensible claims, connect evidence to reasoning, and sustain logical progression across paragraphs (Fajaryani et al., 2021; Ozfidan & Mitchell, 2020; Setyowati et al., 2020). These processes require critical thinking because effective academic writing depends on the ability to analyze issues, evaluate options, make rhetorical decisions, and revise ideas deliberately. Thus, writing instruction should not treat critical thinking as an additional skill separate from writing, but as a core process through which students develop more meaningful academic texts.

Generative AI has entered this landscape as a powerful but pedagogically ambiguous resource. Unlike earlier automated writing evaluation systems that mainly provided corrective or evaluative feedback, generative AI can produce ideas, outlines, explanations, paragraphs, counterarguments, and revisions in response to student prompts. This capacity makes AI potentially useful for EFL students who struggle with idea generation, organization, and revision. Recent studies suggest that AI-supported writing tools can reduce cognitive pressure, provide immediate and personalized support, and assist students in brainstorming, outlining, proofreading, and reflecting on their writing (Gayed et al., 2022; Tiandem-Adamou, 2025; Wang, 2024). Yet these same affordances also create risks. Students may overtrust AI responses, adopt a proofreading-only orientation, copy AI-generated text, or rely on AI to make decisions that should involve their own reasoning (Ranalli, 2021; Zhang & Hyland, 2018; Zhang, 2020). In this sense, generative AI is not automatically empowering. Its value depends on whether students use it to question, compare, select, revise, and reflect, or whether they use it simply to outsource the writing process.

The concept of cognitive scaffolding offers a useful lens for understanding this issue. Scaffolding refers to support that helps learners perform tasks they cannot yet complete independently while gradually developing greater control, awareness, and autonomy. In EFL academic writing, cognitive scaffolding may help students manage idea development, argument organization, paragraph connection, and revision decisions. Generative AI prompts can function as scaffolding when they help students clarify concepts, explore possible structures, evaluate alternative arguments, and identify weaknesses in their drafts. This perspective shifts the focus from AI as a text generator to AI as a thinking support. However, scaffolding is meaningful only when learners remain cognitively active. If students accept AI output without evaluation, the scaffold becomes a shortcut rather than a learning support. Therefore, AI-assisted writing needs to be examined not only through final written products, but through students' actual interactions with prompts, their revision behavior, and their reflections on how AI shaped their thinking.

Existing research has increasingly examined AI-supported writing, but several gaps remain. Many studies focus on writing outcomes, automated feedback, product improvement, or general perceptions of AI tools. Less attention has been given to how EFL students interact with generative AI prompts during the writing process itself, especially in authentic academic writing classrooms. Qualitative evidence is still needed to explain how students use AI prompts to activate ideas, organize writing, evaluate suggestions, revise arguments, and develop critical awareness. It is also important to understand the tension between support and dependence: the same AI tool may help one student think more clearly while encouraging another student to copy generated text with minimal reflection. Without examining these process-level experiences, it remains difficult to determine whether generative AI functions as cognitive scaffolding or merely as a convenient writing shortcut.

To address this gap, the present study explores how undergraduate EFL students use generative AI prompts during academic writing and how these prompts shape their engagement, writing organization, and critical thinking. The study is situated in a second-semester Academic Writing course at a university in Malang, Indonesia, involving 10 undergraduate EFL students. Data were collected through classroom observations, screen-recorded writing sessions, students' writing documents, semi-structured interviews, and reflection journals, allowing the study to trace students' AI-supported writing processes in detail. The novelty of this study lies in its process-oriented qualitative focus: instead of measuring whether AI improves writing scores, it examines how students negotiate AI support while generating ideas, structuring arguments, revising drafts, and reflecting on their own thinking. This focus is consistent with the article's findings that AI prompts helped students activate ideas, organize writing, increase confidence, and develop metacognitive awareness, while also creating risks of overreliance and uncritical adoption of AI-generated text.

Based on the gaps identified above, this study aims to explore how undergraduate EFL students use generative AI prompts during academic writing and how these prompts function as cognitive scaffolding in the development of ideas, organization of writing, and reflective revision. It also examines the forms of engagement and risks that emerge when students interact with AI-generated responses during the writing process. By focusing on students' actual AI-assisted writing practices rather than only on final written products, this study seeks to provide a more contextualized understanding of when generative AI supports students' thinking and when it may lead to dependency or uncritical text adoption. Accordingly, the study is guided by the following research questions: (1) How do EFL students use generative AI prompts during the academic writing process? (2) How do generative AI prompts scaffold students' idea development, writing organization, and reflective revision? and (3) What forms of engagement and risk emerge from students' use of generative AI prompts in academic writing?

2. Method

2.1 Research Design

This study employed a descriptive qualitative case study design to explore how EFL students used generative AI prompts during the academic writing process and how such prompts functioned as cognitive scaffolding for idea development, writing organization, reflective revision, engagement, and critical thinking. A qualitative case study was considered appropriate because the study focused on a contemporary phenomenon, AI-assisted academic writing, within a bounded instructional context, namely one Academic Writing course in an English Education program at a university in Malang, Indonesia. Rather than measuring the effectiveness of AI prompts through test scores or establishing causal relationships, the study sought to understand students' actual writing practices, decision-making processes, experiences, and reflections while interacting with generative AI. This design allowed the researcher to examine not only what students produced, but also how they generated prompts, interpreted AI responses, revised drafts, and negotiated the tension between assistance and dependency during composing. Qualitative case study research is particularly suitable for investigating complex classroom-based practices in depth and for capturing participants' perspectives within their natural learning context (Creswell & Poth, 2018; Merriam & Tisdell, 2016).

2.2 Research Context and Participants

The study was conducted in a second-semester Academic Writing course in an English Education program at a university in Malang, Indonesia. The course introduced undergraduate students to foundational academic writing practices, including paragraph development, essay organization, argument construction, drafting, revision, and reflection. Generative AI-assisted writing activities were integrated into selected writing sessions to support students in exploring ideas, organizing arguments, revising drafts, and reflecting on writing decisions.

The participants were 10 first-year undergraduate EFL students enrolled in the course. They were selected through purposive sampling because the study required participants who had direct and relevant experience with the phenomenon under investigation. Three selection criteria were applied: the students had prior experience using ChatGPT or similar generative AI tools for writing-related tasks, they actively participated in AI-assisted writing activities during the course, and they were willing to share their writing documents, reflection journals, screen-recorded writing sessions, and interview responses. Purposive sampling was appropriate because it enabled the researcher to select information-rich participants who could provide detailed insights into AI-supported writing practices (Patton, 2015; Merriam & Tisdell, 2016). The original manuscript reports that the study involved 10 undergraduate EFL students in an Academic Writing course and used multiple sources of data, including observations, screen recordings, writing documents, interviews, and reflection journals.

2.3 Data Sources and Instruments

This study used multiple qualitative data sources to obtain a rich and triangulated understanding of students' interaction with generative AI prompts. The data sources included classroom observations, screen-recorded writing sessions, students' writing documents, AI interaction records, semi-structured interviews, and reflection journals.

Classroom observations were conducted to examine how students used generative AI prompts during different stages of writing, including pre-writing, drafting, revising, and finalizing their texts. Observation notes focused on students' prompt formulation, responses to AI-generated suggestions, writing behavior, revision decisions, and visible forms of engagement during AI-assisted writing activities. Screen recordings were used to capture students' real-time interactions with AI tools more closely, including the prompts they entered, the responses they received, the parts they accepted or rejected, and the ways they modified their drafts. This source was important because it allowed the researcher to trace writing behavior that might not be fully captured through interviews alone.

Students' writing documents were also collected and analyzed. These documents included two to three versions of each student's text, namely initial drafts, revised drafts, and final drafts. When available, prompt histories, AI-generated responses, and revised text segments were also collected to examine how students transformed AI input into their own writing. Document analysis was useful for identifying changes in idea development, paragraph organization, argument structure, transitions, and revision quality across drafts (Bowen, 2009).

Semi-structured interviews were conducted with all participants after the writing project had been completed. The interview protocol explored students' experiences of using generative AI prompts, including how AI helped them generate ideas, organize writing, manage difficulties, evaluate suggestions, revise drafts, and reflect on their thinking. Semi-structured interviews were appropriate because they allowed the researcher to ask focused questions while also giving participants space to elaborate on their experiences and reasoning (Galletta, 2013). Reflection journals were collected at the end of each writing session. In these journals, students reflected on the challenges they faced, the strategies they used, the role of AI in their writing process, and their awareness of changes in their thinking and writing behavior. Reflection journals were included to capture students' internal experiences and evolving perceptions across the writing process (Hyers, 2018).

2.4 Data Collection Procedures

Data were collected over four instructional sessions conducted across four weeks, with each session lasting approximately 90 minutes. The data collection process followed the sequence of the academic writing activities implemented in the course. In the first stage, students were introduced to the writing task and were allowed to use generative AI prompts to support idea generation and initial planning. They were encouraged to formulate prompts related to topic exploration, possible arguments, essay outlines, and paragraph

organization. During this stage, classroom observations and screen recordings were used to document how students interacted with AI at the pre-writing stage.

In the second stage, students developed their initial drafts by using their own ideas and, where relevant, by referring to AI-generated suggestions. The researcher collected students' early drafts and recorded how they used AI responses to shape their writing. In the third stage, students revised their drafts. They used AI prompts to ask for feedback on organization, coherence, argument development, paragraph transitions, and language clarity. Their screen-recorded sessions and writing documents were examined to identify whether they accepted, adapted, questioned, or rejected AI-generated suggestions. In the fourth stage, students finalized their drafts and completed reflection journals describing how AI had supported or complicated their writing process.

After the writing project was completed, semi-structured interviews were conducted with all participants. The interviews focused on students' cognitive, emotional, and reflective experiences while using generative AI prompts. They were asked to explain how they used AI, which AI responses they found useful, how they decided whether to accept or revise AI suggestions, whether AI made them more confident or dependent, and how the experience affected their thinking about academic writing. This sequence of observation, document collection, journal writing, and interviews enabled the researcher to examine AI-assisted writing both as an observable classroom practice and as a reflective student experience.

2.5 Data Analysis

The data were analyzed using reflexive thematic analysis, following Braun and Clarke's approach and its later operationalization in qualitative research (Braun & Clarke, 2022; Byrne, 2022). This method was selected because the study aimed to identify patterns of meaning across multiple qualitative data sources and interpret how students experienced generative AI prompts as cognitive scaffolding during academic writing.

The analysis began with data familiarization. The researcher repeatedly read classroom observation notes, interview transcripts, reflection journals, writing drafts, and AI interaction records to develop an overall understanding of the dataset. Screen-recorded sessions were reviewed to identify students' prompting behavior, responses to AI output, revision decisions, and signs of dependence or critical engagement. In the second stage, initial codes were generated manually. The codes captured recurring actions and experiences such as *starting from AI-generated ideas*, *asking AI for outlines*, *reorganizing paragraphs*, *copying AI text*, *modifying AI suggestions*, *feeling more confident*, *evaluating AI responses*, *noticing weak arguments*, and *reflecting on paragraph connection*.

In the third stage, related codes were grouped into preliminary themes. These themes were reviewed across data sources to ensure that they reflected consistent patterns rather than isolated responses. For example, codes related to brainstorming, topic narrowing, and identifying writing directions were grouped under the broader theme of initial idea activation. Codes related to outlining, sequencing, transitions, and essay structure were grouped under writing workflow and organization. Codes related to confidence, reduced anxiety, and emotional reassurance were grouped under affective engagement. Codes related to copy-paste behavior, minimal modification, and heavy reliance on AI were grouped under AI dependency and passive uptake. Codes related to evaluating suggestions, comparing arguments, and revising ideas were grouped under reflective thinking and metacognitive awareness.

In the fourth stage, the themes were reviewed, refined, and named to align with the research questions. The final themes were then interpreted in relation to the concept of cognitive scaffolding, student engagement, and critical thinking in EFL academic writing. Evidence from interviews, reflection journals, screen recordings, and writing documents was compared to strengthen the credibility of the interpretation. The analysis was conducted manually to maintain interpretive sensitivity, transparency, and close engagement with the data (Miles et al., 2020; Saldaña, 2021).

2.6 Trustworthiness

Several strategies were used to enhance the trustworthiness of the study. First, data triangulation was applied by using multiple sources of evidence, including classroom observations, screen recordings, writing documents, interview transcripts, and reflection journals. This allowed the researcher to compare students' reported experiences with their observable writing behavior and actual written drafts. Second, methodological triangulation was used by combining observation, document analysis, interviews, and reflective writing. This helped provide a more comprehensive understanding of how students used generative AI prompts during writing.

Third, an audit trail was maintained throughout the research process. The researcher documented data collection procedures, coding decisions, theme development, and analytic reflections to make the research process more transparent and traceable. Fourth, thick description was used in reporting the findings by providing detailed contextual information, participant accounts, and examples from writing practices. This helped readers understand how AI-assisted writing occurred in the specific EFL academic writing context. Fifth, constant comparison was used during analysis to check whether emerging themes were supported across different participants and data sources. These procedures were intended to strengthen credibility, dependability, and confirmability in the qualitative analysis.

3. Findings

3.1 Students' Use of Generative AI Prompts during the Academic Writing Process

The findings show that students used generative AI prompts across different stages of academic writing, particularly during idea generation, outlining, drafting, organization, and revision. At the beginning of the writing process, several students reported difficulty transforming a broad topic into specific ideas. This difficulty was reflected in AL's statement: **"***I'm usually confused about where to start. Sometimes I already know the theme, but I don't know what specific ideas to bring up***."** For students who experienced this kind of initial writing block, AI prompts helped provide possible angles, topic directions, and starting points for writing.

Students also used AI prompts to generate rough outlines and organize their early writing plans. NY explained, **"***I immediately get a rough structure. It helps me see the order of ideas***."** This indicates that AI was used not only to produce content but also to help students visualize the sequence of ideas before drafting. RF similarly stated, **"***Before using AI, it took me more than an hour to compose an opening paragraph… After using AI prompts…the writing process became more focused and efficient***."** These responses suggest that AI prompts reduced uncertainty in the early stages of writing and helped students enter the composing process with greater clarity. During drafting, AI was used to expand and develop ideas. AJ described AI as **"***like a sparring partner in critical thinking***,"** while RS stated, **"***AI encouraged me to develop arguments, not just descriptions." BY added, "I can explore many possibilities before choosing one main idea***,"** and FR explained, **"***AI helped me combine various ideas into a cohesive new idea***."** These accounts show that some students used AI prompts as a dialogic thinking tool. Rather than immediately accepting AI-generated responses, they used AI to explore alternatives, compare possible arguments, and decide which ideas were most relevant to their writing purpose.

The screen-recorded writing sessions and document analysis further showed that students used AI prompts to improve structure and flow. In several pre-AI drafts, particularly those written by AL, AJ, RS, FR, and BY, the essays displayed weak openings, loosely connected ideas, and limited transitions. After using prompts such as **"***Can you give me a structure outline***…?"**, their later drafts showed clearer introductions, more visible thesis statements, more logical sequencing of arguments, and stronger transitions between paragraphs. DW's draft showed a similar shift: before AI use, the essay lacked clear organization; after AI-assisted writing, it adopted a five-paragraph structure and used connectors such as **"***This shows that***…"** and **"***Therefore***…"** to

strengthen textual progression. These findings indicate that students used generative AI prompts as writing support across multiple stages, not merely as a grammar correction tool. AI helped them begin writing, develop ideas, organize arguments, and revise structure. However, students' use of AI varied in quality. Some students treated AI as a source of options to be evaluated and adapted, while others used it more directly as a source of ready-made text.

3.2 Generative AI Prompts as Scaffolding for Idea Development, Organization, and Reflective Revision

The second research question examined how generative AI prompts functioned as cognitive scaffolding in students' academic writing. The findings indicate that AI prompts acted as scaffolding when they helped students perform writing tasks that they found difficult, particularly generating ideas, structuring essays, developing arguments, improving coherence, and evaluating draft weaknesses. Table 1 summarizes the main scaffolding functions identified from the data.

Table 1. Generative AI Prompts as Cognitive Scaffolding in Students' Academic Writing

Scaffolding function	**Selected evidence**	**Interpretation**
Activating initial ideas	"I'm usually confused about where to start… I don't know what specific ideas to bring up." — AL	AI helped students overcome initial writing blocks and identify possible directions for writing.
Creating a writing structure	"I immediately get a rough structure. It helps me see the order of ideas." — NY	AI made essay structure more visible and manageable for students.
Supporting argument development	"AI encouraged me to develop arguments, not just descriptions." — RS	AI helped students move beyond descriptive writing toward more argumentative thinking.
Encouraging idea selection	"I can explore many possibilities before choosing one main idea." — BY	AI gave students options that they could compare before deciding on a writing direction.
Improving paragraph connection	"After seeing AI's suggestions, I understand better how to make connections between one paragraph and another." — AJ/RF	AI supported students' awareness of coherence and paragraph relationships.
Supporting reflective revision	"AI helped me realize my paragraphs were too long, arguments were weak, or transitions were unclear." — FR/BY	AI helped students notice weaknesses and revise their drafts more purposefully.

As shown in Table 1, one of the clearest scaffolding functions of AI prompts was idea activation. Students who initially struggled to begin writing were able to use AI prompts to generate possible directions and identify more specific points to develop. This support was especially useful during the pre-writing stage because it helped students move from vague topic awareness to a more workable writing plan. In this sense, AI did not simply supply answers; it helped students identify what could be written and how their ideas might be developed.

AI prompts also scaffolded writing organization. Students used AI-generated outlines to understand possible essay structures, arrange ideas, and create clearer paragraph sequences. The development of more explicit introductions, thesis statements, and transitions in later drafts suggests that AI made organizational patterns more visible. However, the data also show that this scaffolding was more meaningful when students adapted AI-generated structures to fit their own writing purposes. NY and RS, for example, reported that they adjusted AI suggestions rather than copying them directly, indicating that AI worked as a support for organization rather than a fixed template.

Another important scaffolding function appeared in argument development and reflective revision. RS's statement that AI encouraged him to develop "*arguments, not just descriptions*" suggests that AI helped

students recognize the argumentative demands of academic writing. BY's statement about exploring **"***many possibilities before choosing one main idea***"** also shows that AI prompted comparison and selection. This process is important because critical writing requires students to evaluate alternatives, decide which argument is most suitable, and organize their position coherently.

The strongest evidence of cognitive scaffolding appeared in students' reflective use of AI during revision. AJ and RF stated, **"***Usually, I write just following what I think at the time. But after seeing AI's suggestions, I understand better how to make connections between one paragraph and another***."** This suggests that AI helped students become more aware of coherence and paragraph relationships. AL and NY described a more deliberate decision-making process: **"***I don't write right away. I ask the AI to give me some arguments, then I consider which one best fits my position. From there, I organize my own ideas***."** FR and BY further noted, **"***AI helped me realize my paragraphs were too long, arguments were weak, or transitions were unclear***."** These responses indicate that AI prompts supported metacognitive awareness by helping students notice weaknesses, evaluate options, and revise more deliberately. The findings, therefore, show that generative AI prompts functioned as cognitive scaffolding when students remained active in interpreting and adapting AI responses. AI-supported idea development, organization, and revision are most effective when students use it to think through choices rather than replace their own writing decisions.

3.3 Forms of Engagement and Risks in Students' Use of Generative AI Prompts

The third research question focused on the forms of engagement and risks that emerged from students' use of generative AI prompts. The findings revealed two contrasting patterns. On the one hand, AI supported students' affective and cognitive engagement by reducing hesitation, increasing confidence, and helping them continue writing when they felt stuck. On the other hand, AI also created risks of dependency, particularly when students copied AI-generated text or accepted AI suggestions without critical evaluation. Table 2 summarizes these forms of engagement and risk.

Table 2. Forms of Engagement and Risks in Students' Use of Generative AI Prompts

Category	Evidence	Interpretation
Affective engagement	"I feel more confident because when I'm confused, I can ask AI first. It feels like having a friend who is always ready to help." — BY	AI reduced uncertainty and increased students' confidence during writing.
Affective engagement	"I used to be afraid of making mistakes… Now I'm calmer, because I have a 'backup of ideas' when I get stuck." — DW	AI helped students manage fear of error and writing anxiety.
Cognitive engagement	"AI encouraged me to develop arguments, not just descriptions." — RS	AI stimulated students to think more deeply about argument development.
Productive persistence	"I used to only be able to write one paragraph, but now I can write three to five paragraphs because I have references." — RF	AI helped students sustain writing and expand their drafts.
Reflective engagement	"I ask the AI to give me some arguments, then I consider which one best fits my position." — AL/NY	AI supported evaluation and selective use of ideas.
Dependency risk	"If I've been given a good text from the AI, I just copy-paste it and change it slightly. Sometimes I don't change it at all." — RF	AI use became problematic when students treated generated text as a finished product.
Dependency risk	DW admitted relying completely on AI for important sections such as the introduction and conclusion.	AI could reduce student authorship when used as a substitute for independent composing.

The first pattern was affective engagement. Several students described AI as a source of reassurance during writing. BY stated, **"***I feel more confident because when I'm confused, I can ask AI first. It feels like having a*

friend who is always ready to help." DW similarly explained, "*I used to be afraid of making mistakes, afraid that my writing would be bad. Now I'm calmer, because I have a 'backup of ideas' when I get stuck*." These responses indicate that AI helped reduce writing-related hesitation and gave students emotional support when they experienced uncertainty. For students who often felt blocked at the beginning of writing, immediate AI responses created a sense of security and encouraged them to continue. AI also supported cognitive engagement and writing persistence. RF explained, "*I used to only be able to write one paragraph, but now I can write three to five paragraphs because I have references*." This suggests that AI helped students sustain writing by providing ideas, examples, or references that could be developed further. Other students, including AL, AJ, and RS, reported greater willingness to try new styles and express opinions more openly. These findings suggest that AI prompts encouraged students to engage more actively with the writing process by exploring ideas, extending arguments, and revising their drafts.

However, the same tool also produced clear risks. Some students used AI-generated responses passively and showed signs of dependency. RF admitted, "*If I've been given a good text from the AI, I just copy-paste it and change it slightly. Sometimes I don't change it at all*." Screen-recorded writing sessions supported this finding, showing that in some unsupervised sessions, students copied AI-generated passages directly into their essays without meaningful modification. DW also admitted relying completely on AI for important sections such as the introduction and conclusion. These examples show that AI sometimes became a substitute for students' own composing rather than a scaffold for learning. The findings reveal that engagement and risk existed side by side in AI-assisted writing. AI promoted engagement when students used it to ask questions, compare options, revise ideas, and improve organization. It became risky when students treated AI as an authority or a shortcut. The distinction depends largely on whether students remained active decision-makers in the writing process. When students evaluated and adapted AI responses, prompts functioned as cognitive scaffolding. When students copied or minimally modified AI output, the scaffolding potential weakened, and students' intellectual ownership became less visible.

4. Discussion

The findings of this study show that generative AI prompts played a meaningful but ambivalent role in EFL academic writing. AI prompts helped students activate initial ideas, organize arguments, revise drafts, and reflect on writing decisions, but they also created risks of dependency when students treated AI-generated responses as ready-made text. This duality is central to understanding AI-assisted writing: generative AI is neither automatically empowering nor inherently harmful. Its educational value depends on how students engage with it during the writing process. When students used AI to explore alternatives, compare ideas, identify weak transitions, and revise arguments, AI functioned as a cognitive scaffold. However, when they copied AI-generated text with minimal modification, AI shifted from being a scaffold to becoming a substitute for students' own thinking and authorship. This pattern is consistent with the study's classroom observations, screen-recorded sessions, writing documents, interviews, and reflection journals, which showed both productive engagement and passive dependence in students' AI-assisted writing practices.

One important contribution of AI prompts was their ability to support students at the early stages of writing. Several students struggled not because they lacked a general understanding of the topic, but because they did not know how to transform broad ideas into specific, developable points. AI prompts helped reduce this initial writing blockage by providing possible directions, outlines, and argumentative angles. This finding reinforces the view that EFL academic writing difficulties are not limited to grammar and vocabulary, but also involve idea generation, argument development, organization, and rhetorical control (Altınmakas & Bayyurt, 2018; Bulqiyah et al., 2021; Chuang & Yan, 2022; Toba et al., 2019). In this study, AI was useful because it made writing pathways visible. Students could see how ideas might be sequenced, how thesis statements could be framed, and how transitions could connect paragraphs. Nevertheless, this support became educationally meaningful only when students adapted AI suggestions to their own writing purposes rather than accepting them as fixed templates.

The findings also suggest that generative AI prompts can scaffold metacognitive awareness and critical thinking when students remain active decision-makers. Some participants used AI to compare arguments, select ideas that matched their position, recognize long paragraphs, identify weak reasoning, and revise unclear transitions. These behaviors indicate that AI-supported planning, monitoring, evaluation, and revision, which are central to metacognitive writing development (Huang & Zhang, 2022; Qin et al., 2022; Sun & Zhang, 2022; Teng, 2020). Critical thinking in this study was not treated as an abstract ability or test score, but as a process visible in students' writing decisions: what to include, what to reject, how to organize arguments, and how to revise weak ideas. From this perspective, AI prompts supported critical thinking not by producing better answers, but by giving students options to question, evaluate, and transform. The strongest evidence of scaffolding appeared when students used AI-generated responses as provisional input for reflection, not as final text.

Another important finding concerns students' affective engagement. AI prompts helped some students feel calmer, more confident, and more willing to continue writing when they felt stuck. Students described AI as a "friend" or a "backup of ideas," suggesting that AI reduced hesitation and made writing feel more manageable. This affective support matters because EFL academic writing often creates anxiety, especially when students must develop ideas, organize arguments, and manage language accuracy simultaneously. However, confidence alone is not sufficient. Emotional comfort becomes pedagogically valuable only when it leads to deeper engagement with planning, argumentation, revision, and self-evaluation. If confidence results only in copying AI-generated text, then AI may reduce anxiety while weakening learning. Therefore, the affective benefit of AI should be directed toward reflective writing practices rather than passive reliance.

The risk of overreliance remains the most serious challenge identified in this study. Some students copied AI-generated passages directly, made only minor modifications, or depended heavily on AI for important sections such as introductions and conclusions. This finding aligns with concerns that learners may overtrust automated responses, adopt a surface-level proofreading orientation, or accept AI output without sufficient evaluation (Ranalli, 2021; Wang, 2024; Zhang & Hyland, 2018; Zhang, 2020). The issue is not merely academic dishonesty, but the possible weakening of students' intellectual ownership. When AI makes rhetorical decisions on behalf of students, they may lose opportunities to practice argument construction, organization, and revision. For this reason, AI-assisted writing instruction should require students to make their decision-making visible. Teachers can ask students to submit prompts, selected AI responses, annotated revisions, and short reflections explaining what they accepted, rejected, or modified. Such practices can shift AI use from hidden dependency to accountable learning.

These findings imply that generative AI should be integrated into EFL academic writing through explicit instructional guidance. Students need prompt literacy, including how to write context-rich prompts, ask for alternatives, request explanations, evaluate accuracy, and revise AI output into student-owned writing. Teachers should frame AI as a scaffold for thinking, not as a source of finished text. AI-supported writing tasks should also emphasize process rather than product by requiring drafting, reflection, peer discussion, and revision evidence. Because students with stronger digital literacy and motivation appeared more capable of using AI critically, unguided AI use may widen learning differences. The study is limited by its small sample of 10 students, its single-course context, and its short four-session duration, so the findings should be interpreted as context-specific rather than generalizable. Future studies should examine AI-assisted writing longitudinally, compare guided and unguided AI use, and investigate how prompt literacy instruction affects students' writing quality, metacognitive awareness, and critical thinking. The central contribution of this study is its process-oriented insight: generative AI prompts can support EFL academic writing when they help students think, organize, evaluate, and revise, but they become problematic when they replace students' own reasoning and authorship.

5. Conclusion

The use of generative AI in EFL academic writing should be understood less as a question of whether students may use AI and more as a question of how they learn to think with it. This study shows that generative AI prompts can support students' academic writing when they function as cognitive scaffolding rather than as a substitute for authorship. Students used AI prompts to activate initial ideas, generate possible writing directions, organize arguments, improve paragraph flow, and identify weaknesses in their drafts. For several participants, AI also reduced hesitation and increased confidence during writing, especially when they felt stuck or uncertain. These findings indicate that AI prompts can make the writing process more manageable by supporting idea development, organization, and reflective revision. However, the study also revealed a critical tension: some students used AI-generated responses passively, copied text with minimal revision, or relied heavily on AI for important essay sections. This suggests that the value of generative AI depends not on the tool itself, but on students' ability to evaluate, adapt, and take ownership of AI-supported writing decisions.

The study therefore, calls for AI-assisted writing instruction that foregrounds student agency, prompt literacy, and critical reflection. Teachers should guide students to use AI as a source of questions, alternatives, outlines, and feedback rather than as a provider of finished text. Writing tasks should require students to document prompts, explain how they used AI responses, justify revisions, and reflect on what remained their own intellectual contribution. Although this study offers rich process-based insights, it was limited to 10 undergraduate EFL students in one Academic Writing course and was conducted across four instructional sessions. Future research should involve larger and more diverse participants, examine AI-assisted writing over longer periods, and compare guided and unguided AI use to determine how prompt literacy and teacher mediation influence writing quality, critical thinking, and learner autonomy. The main contribution of this study lies in showing that generative AI prompts can become meaningful scaffolds for EFL academic writing only when they help students develop, not bypass, their own thinking.

6. Acknowledgement

The authors gratefully acknowledge the participation of the undergraduate EFL students who took part in this study and shared their writing processes, reflection journals, AI interaction records, and interview responses. Appreciation is also extended to the Academic Writing course instructor and the English Education program at the university in Malang, Indonesia, for supporting the implementation of this classroom-based research. The authors value the participants' openness in discussing both the benefits and challenges of using generative AI prompts during academic writing, as their contributions made it possible to understand AI-assisted writing as a real learning process rather than merely a technological practice.

7. Declaration of AI Use

This manuscript was prepared with limited support from ChatGPT. ChatGPT was used to assist with language refinement, academic phrasing, and improving the flow of selected sections. The use of this tool was restricted to editing and presentation support. The authors designed the study, collected the data, analyzed the findings, interpreted the results, and developed the scholarly arguments independently. AI-assisted suggestions were reviewed, revised, and approved by the authors to ensure that the final manuscript reflects the authors' own academic judgment, maintains the integrity of the data, and preserves the originality of the work.

8. References

Altınmakas, D., & Bayyurt, Y. (2018). An exploratory study on factors influencing undergraduate students' academic writing practices in Turkey. *Journal of English for Academic Purposes*, *37*, 88–103. https://doi.org/10.1016/j.jeap.2018.11.006

Andewi, W., Waziana, W., Wibisono, D., Putra, K. A., Hastomo, T., & Oktarin, I. B. (2025). From Prompting to Proficiency: A Mixed-Methods Analysis of Prompting with ChatGPT Versus Lecturer Interaction in an EFL Classroom. *Journal of Studies in the English Language*, *20*(2), 210–238. https://doi.org/10.64731/jsel.v20i2.282318

Angrosino, M. (2007). *Doing ethnographic and observational research*. SAGE.

Ariyanti, A., & Fitriana, R. (2017). EFL Students' Difficulties and Needs in Essay writing. *Proceedings of the International Conference on Teacher Training and Education 2017 (ICTTE 2017)*. https://doi.org/10.2991/ictte-17.2017.4

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.

Bulqiyah, S., Mahbub, M. A., & Nugraheni, D. A. (2021). Investigating writing difficulties in essay writing: Tertiary students' perspectives. *English Language Teaching Educational Journal, 4*(1), 61–73. https://doi.org/10.12928/eltej.v4i1.2371

Byrne, D. (2022). A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & Quantity, 56*(3), 1391–1412. https://doi.org/10.1007/s11135-021-01182-y

Chuang, P.-L., & Yan, X. (2022). An investigation of the relationship between argument structure and essay quality in assessed writing. *Journal of Second Language Writing, 56*, 100892. https://doi.org/10.1016/j.jslw.2022.100892

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE.

Fajaryani, N., Mukminin, A., Hidayat, M., Muhaimin, M., Haryanto, E., Nazurty, N., Marzulina, L., Harto, K., & Habibi, A. (2021). Cultural capital and Argumentative writing in English: Challenges and strategies used by EFL student teachers. *The Qualitative Report*. https://doi.org/10.46743/2160-3715/2021.4784

Galletta, A. (2013). *Mastering the semi-structured interview and beyond: From research design to analysis and publication*. NYU Press.

García-Moro, F. J., Gómez-Baya, D., Muñoz-Silva, A., & Martín-Romero, N. (2021). A qualitative and quantitative study on critical thinking in social education degree students. *Sustainability, 13*(12), 6865. https://doi.org/10.3390/su13126865

Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence, 3*, 100055. https://doi.org/10.1016/j.caeai.2022.100055

Ghanizadeh, A. (2017). The interplay between reflective thinking, critical thinking, self-monitoring, and academic achievement in higher education. *Higher Education, 74*(1), 101–114. https://doi.org/10.1007/s10734-016-0031-y

Huang, Y., & Zhang, L. J. (2022). Facilitating L2 writers' metacognitive strategy use in argumentative writing using a process-genre approach. *Frontiers in Psychology, 13*, Article 1036831. https://doi.org/10.3389/fpsyg.2022.1036831

Hyers, L. L. (2018). *Diary methods: Understanding qualitative research*. Oxford University Press.

Ka-Kan-Dee, M., & Kaur, S. (2015). Teaching strategies used by Thai EFL lecturers to teach argumentative writing. *Procedia - Social and Behavioral Sciences*, *208*, 143–156. https://doi.org/10.1016/j.sbspro.2015.11.191

Kochmar, E., Vu, D. D., Belfer, R., Gupta, V., Serban, I. V., & Pineau, J. (2022). Automated data-driven generation of personalized pedagogical interventions in intelligent tutoring systems. *International Journal of Artificial Intelligence in Education, 32*(2), 323–349. https://doi.org/10.1007/s40593-021-00267-x

Li, H. H., & Zhang, L. J. (2022). Investigating the effects of Small-Group Student talk on the quality of argument in Chinese Tertiary English as a foreign language learners' argumentative writing. *Frontiers in Psychology*, *13*, 868045. https://doi.org/10.3389/fpsyg.2022.868045

Li, J., Link, S., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing, 27*, 1–18. https://doi.org/10.1016/j.jslw.2014.10.004

Lim, F. V., & Phua, J. (2019). Teaching writing with language feedback technology. *Computers and Composition, 54*, Article 102518. https://doi.org/10.1016/j.compcom.2019.102518

Liu, M., Zhang, L. J., & Biebricher, C. (2023). Investigating students' cognitive processes in generative AI-assisted digital multimodal composing and traditional writing. *Computers & Education*, *211*, 104977. https://doi.org/10.1016/j.compedu.2023.104977

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE.

Nejmaoui, N. (2019). Improving EFL learners' critical thinking skills in argumentative writing. *English Language Teaching, 12*(1), 98–109. https://doi.org/10.5539/elt.v12n1p98

Ozfidan, B., & Mitchell, C. (2020). Detected difficulties in argumentative writing: the case of culturally and linguistically Saudi backgrounded students. *Journal of Ethnic and Cultural Studies*, 15–29. https://doi.org/10.29333/ejecs/382

Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). SAGE.

Petchprasert, A. (2021). Utilizing an automated tool analysis to evaluate EFL students' writing performances. *Asian-Pacific Journal of Second and Foreign Language Education*, *6*(1). https://doi.org/10.1186/s40862-020-00107-w

Qin, C., Zhang, R., & Xiao, Y. (2022). A questionnaire-based validation of metacognitive strategies in writing and their predictive effects on the writing performance of English as foreign language student writers. *Frontiers in Psychology, 13*, Article 1071907. https://doi.org/10.3389/fpsyg.2022.1071907

Ranalli, J. (2021). L2 student engagement with automated feedback on writing: Potential for learning and issues of trust. *Journal of Second Language Writing, 52*, Article 100816. https://doi.org/10.1016/j.jslw.2021.100816

Robillos, R. J., & Thongpai, J. (2022). Computer-aided argument mapping within a metacognitive approach: Its impact on students' argumentative writing performance and self-regulated learning. *LEARN Journal: Language Education and Acquisition Research Network, 15*(2), 160–186.

Rusfandi. (2015). Argument-Counterargument structure in Indonesian EFL Learners' English Argumentative Essays: A Dialogic Concept of Writing. *RELC Journal*, *46*(2), 181–197. https://doi.org/10.1177/0033688215587607

Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). SAGE.

Setyowati, L., Agustina, F., Sukmawan, S., El-Sulukiyyah, A. A., & Mabaroh, B. (2020). The students' problems and solutions in writing an argumentative essay on gender issue. *Journal of English Language Teaching and Linguistics*, *5*(3), 279. https://doi.org/10.21462/jeltl.v5i3.424

Sun, Q., & Zhang, L. J. (2022). Understanding learners' metacognitive experiences in learning to write in English as a foreign language: A structural equation modeling approach. *Frontiers in Psychology, 13*, Article 986301. https://doi.org/10.3389/fpsyg.2022.986301

Sundari, H., & Febriyanti, R. H. (2021). The Analysis of Indonesian EFL Argumentative Writing Using Toulmin's Model: The Structure and Struggles from the Learners. *Scope Journal of English Language Teaching*, *5*(2), 67. https://doi.org/10.30998/scope.v5i2.8544

Teng, F. (2019). Tertiary-level students' English writing performance and metacognitive awareness: A group metacognitive support perspective. *Scandinavian Journal of Educational Research, 64*(4), 551–568. https://doi.org/10.1080/00313831.2019.1595712

Teng, F. (2020). The role of metacognitive knowledge and regulation in mediating university EFL learners' writing performance. *Innovation in Language Learning and Teaching, 14*(5), 436–450. https://doi.org/10.1080/17501229.2019.1615493

Tiandem-Adamou, Y. (2025). Using generative artificial intelligence to support EFL students' writing proficiency in university in China. *Journal of Educational Technology and Innovation*, *6*(4). https://doi.org/10.61414/jeti.v6i4.213

Toba, R., Noor, W. N., & Sanu, L. O. (2019). The current issues of Indonesian EFL students' writing skills: Ability, problem, and reason in writing comparison and contrast essay. *Dinamika Ilmu, 19*(1), 57–73. https://doi.org/10.21093/di.v19i1.1506

Wang, Y. (2024). Cognitive and sociocultural dynamics of self-regulated use of machine translation and generative AI tools in academic EFL writing. *System*, *126*, 103505. https://doi.org/10.1016/j.system.2024.103505

Wu, Y., & Schunn, C. D. (2021). The effects of providing and receiving peer feedback on writing performance and learning of secondary school students. *American Educational Research Journal, 58*(3), 492–526. https://doi.org/10.3102/0002831220945266

Xu, C. (2018). Understanding online revisions in L2 writing: A computer keystroke-log perspective. *System, 78*, 104–114. https://doi.org/10.1016/j.system.2018.08.007

Zhang, Z. V. (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assessing Writing, 43*, 100439. https://doi.org/10.1016/j.asw.2019.100439

Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing, 36*, 90–102. https://doi.org/10.1016/j.asw.2018.02.004

Zhao, C. G., & Liao, L. (2021). Metacognitive strategy use in L2 writing assessment. *System*, *98*, 102472. https://doi.org/10.1016/j.system.2021.102472