International Journal of Instruction *January 2027 ● Vol.20, No.1*
e-ISSN: 1308-1470 ● www.e-iji.net *p-ISSN: 1694-609X*
pp. 289-316

Article submission code: 20260302022149
Received: 02/03/2026 Revision: 24/06/2026
Accepted: 03/07/2026 OnlineFirst: 01/10/2026

Generative AI Use and Creativity in EFL Academic Writing: A Correlational Study Among University Students

Trang Tran Van Pham
Faculty of Foreign Languages, Ton Duc Thang University, Ho Chi Minh City, Vietnam, *mia.trangpham@gmail.com*

Khang Le Minh Tran
Faculty of Foreign Languages, Ton Duc Thang University, Ho Chi Minh City, Vietnam, *joey.khangtran@gmail.com*

Tri Thien Thuong
Faculty of Foreign Languages, Ton Duc Thang University, Ho Chi Minh City, Vietnam, *thuongthientri@gmail.com*

Danh Cong Vu
Corresponding author, Faculty of Foreign Languages, Ton Duc Thang University, Ho Chi Minh City, Vietnam, *vucongdanh@tdtu.edu.vn*

Generative artificial intelligence (AI) tools are increasingly integrated into English as a Foreign Language (EFL) academic writing, yet their relationship with students' creative thinking remains empirically underexamined. This cross-sectional correlational study investigated the association between generative AI use patterns and creativity among 78 English-major undergraduates at a Vietnamese university. Creativity was assessed independently of students' written texts through a modified Alternative Uses Task measuring divergent thinking across fluency, flexibility, and originality. AI use was captured via a seven-dimension questionnaire covering tool selection, frequency and duration, context of use, purpose of use, interaction patterns, content management, and technical aspects. Descriptive results indicated that ChatGPT and Grammarly were the most frequently used tools, employed primarily for grammar correction, vocabulary enhancement, and idea development. Students reported actively modifying and verifying AI-generated content rather than accepting it passively. Correlation analyses revealed that most AI use dimensions were not associated with any creativity measure. A small number of weak positive correlations emerged between originality and usage frequency, content modification, proofreading, and content verification. Flexibility showed no significant association with any AI use variable. These findings suggest that the link between generative AI use and creativity in EFL academic writing is not clear, though post-generation engagement with AI output may be weakly linked to originality.

Keywords: virtual simulation, web-based application, multimedia, teaching, learning

Citation: Pham, T. T. V., Tran, K. L. M., Thuong, T. T., & Vu, D. C. (2027). Generative AI use and creativity in EFL academic writing: A correlational study among university students. *International Journal of Instruction*, *20*(1), 289-316.

INTRODUCTION

The proliferation of generative artificial intelligence (AI) tools, most notably ChatGPT, Grammarly, and QuillBot, has transformed how university students approach academic writing tasks. These tools can produce entire drafts, paraphrase text, correct grammar, and propose organizational structures, fundamentally altering the cognitive demands of writing. Alshreef and Saaty (2025) found that the advent of generative AI disrupted writing classroom dynamics and changed both how students write and what they write. A systematic review by Imran and Almusharraf (2023) concluded that ChatGPT improves writing efficiency, aids translation, and enhances content accuracy in academic writing tasks. For English as a Foreign Language (EFL) students, who often struggle with vocabulary limitations, grammatical accuracy, and idea organization (Dang, 2019; Moses & Mohamad, 2019), generative AI presents both an opportunity for linguistic support and a potential risk to independent cognitive engagement.

Creativity occupies a central position in the writing process. Writing is widely viewed as an iterative cognitive task involving planning, translating ideas into language, and revising (Jurecic, 2006). During the planning phase, writers generate and organize ideas, an inherently creative process. Kormos et al. (2024) demonstrated that L2 learners who exhibit adaptive creativity, the ability to elaborate ideas flexibly, tend to produce higher-quality argumentative writing, whereas excessive emphasis on sheer originality did not necessarily improve writing scores. In other words, creativity, particularly the flexible generation of relevant ideas, plays a significant role in forming coherent academic texts.

EFL students often struggle to write creatively in academic contexts. Numerous studies note that L2 writers face obstacles such as limited vocabulary, weak grammar control, and difficulties organizing ideas, all of which hinder creative expression (Dang, 2019; Moses & Mohamad, 2019). Dang (2019) reported that over half of Vietnamese EFL students cited a lack of ideas as a major difficulty when composing academic texts. Chan and Hu (2023) found that Hong Kong university students viewed generative AI as helpful for idea generation and writing support in higher education. These linguistic constraints can divert cognitive resources away from higher-order creative processes, such as generating novel arguments, exploring unconventional perspectives, or synthesizing ideas across domains, toward surface-level concerns such as grammatical accuracy (Kormos et al., 2024). Generative AI tools may theoretically free up these cognitive resources, enabling students to focus on ideation rather than error correction (Lin, 2025). Alternatively, over-reliance on AI-generated suggestions may reduce the cognitive effort required for independent idea generation, ultimately narrowing the range and novelty of students' writing (Zhai et al., 2024).

Despite growing scholarly attention to AI in education, the intersection of generative AI use and students' creative writing abilities remains largely unexplored. Most studies so far have focused on students' attitudes toward AI (Chan & Hu, 2023; Gasaymeh et al., 2024) or on general writing improvements (Chang et al., 2021; Mahapatra, 2024), rather than on creativity outcomes per se. This gap is especially pronounced for EFL learners, where the combined issues of language difficulty and creative writing have not been

jointly studied in the context of new AI tools. The present study addresses this gap by investigating the correlation between generative AI use patterns and creativity among English-major undergraduates at a Vietnamese university. Specifically, the study pursues two research questions:

1. How, and for what purpose(s), do English majors use generative AI tools in academic writing?

2. Are there any correlations between the use of generative AI tools and students' creativity performance in academic writing?

Review Of Literature

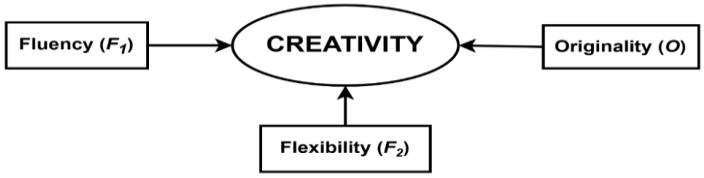
Creativity in EFL Academic Writing

Creativity is broadly defined as the cognitive process of producing responses that are both novel and useful within a given social context (Plucker et al., 2004; Runco & Jaeger, 2012). Despite its centrality to research on cognition and education, the construct has historically lacked definitional consensus. Plucker et al. (2004) found that only 38% of peer-reviewed articles featuring creativity in the title provided an explicit definition. For this study, the working definition follows Plucker et al. (2004): creativity is the interaction of ability, process, and environment to produce something new and useful within a social context. This definition accommodates the situated nature of academic writing, where creativity manifests through students' ability to generate, organize, and express ideas in response to disciplinary prompts. Kaufman and Beghetto (2009) proposed the Four C Model distinguishing mini-c (personal insight), little-c (everyday creativity), Pro-c (professional creativity), and Big-C (eminent creativity); most student writing falls within the mini-c to little-c range, where personal meaning-making intersects with task-appropriate novelty.

Operationally, creativity in divergent thinking research is assessed through three principal dimensions (see

Figure 1). Fluency captures the total quantity of ideas generated, reflecting ideational productivity (Anderson & Graham, 2021; Renzulli et al., 1974; Runco, 2010). Flexibility measures the number of distinct categories represented among those ideas, indicating cognitive breadth and the ability to draw on diverse semantic networks (Kenett et al., 2018; Weiss & Wilhelm, 2022). Originality assesses the statistical rarity of responses, distinguishing novel ideas from conventional ones (Plucker et al., 2011; Runco, 2010). The Alternative Uses Task (AUT; Guilford, 1968) remains the most widely used instrument for measuring these dimensions due to its open-ended format and strong validity across contexts (Pipes, 2023). Together, these three dimensions constitute the standard psychometric framework for divergent thinking assessment derived from Guilford's (1968) Structure of Intellect model and operationalized in the Torrance Tests of Creative Thinking (Torrance, 1966).

Figure 1
Creativity conceptualization



In EFL contexts, creativity in academic writing is constrained by several interrelated factors. Students frequently cite limited vocabulary, weak grammar control, and difficulty organizing ideas as barriers to creative expression (Dang, 2019; Moses & Mohamad, 2019). Dang (2019) reported that over half of Vietnamese EFL students identified a lack of ideas as a major difficulty when composing academic texts. At the institutional level, Chokwe (2013) and Ganapathy et al. (2022) argue that pedagogical systems that prioritize grammatical accuracy over ideational development further inhibit creative thinking. As Kang (2017) notes, EFL students often prioritize correctness over creativity due to fear of making errors. Kormos et al. (2024) demonstrated that L2 writing performance is significantly predicted by divergent thinking ability, suggesting that creativity and language proficiency interact in complex ways. These constraints make EFL writing a particularly important context for investigating whether AI tools can facilitate or hinder creative ideation.

Generative AI Use in EFL Academic Writing

Generative AI tools, including ChatGPT, Grammarly, QuillBot, and others, have been rapidly adopted by university students for academic writing support. A 2024 global survey found that 86% of students used AI in their studies, with ChatGPT (66%) and Grammarly (25%) the most common tools (Kelly, 2024). Freeman (2025) reported that 92% of UK undergraduates used AI in some form, with 88% applying it to assessed work. In Southeast Asian EFL settings, comparable patterns emerge: Thangthong et al. (2024) observed that Thai EFL undergraduates predominantly used familiar tools like Grammarly for vocabulary and grammar accuracy, while Syarifah and Fakhruddin (2024) documented widespread adoption among Indonesian EFL students for grammar checking, brainstorming, and draft refinement.

These tools support writing in several key ways: grammar and language correction, paraphrasing and vocabulary enhancement, idea generation and brainstorming, content expansion, and feedback on organization (Mahapatra, 2024; Chan & Hu, 2023). Mahapatra (2024) found that ChatGPT as a feedback tool had a significant positive impact on ESL students' writing skills. Chang et al. (2021) demonstrated that Chinese EFL students using Grammarly consistently over 16 weeks significantly outperformed controls in writing accuracy and cohesion. Aladini (2023) similarly found marked improvements in academic writing and logical thinking among students receiving AI-assisted instruction over 10 weeks. Hongxia and Razali (2025) likewise found, in a quasi-experiment with 70 Chinese English-major undergraduates, that ChatGPT use improved writing across content, structure, grammar, and vocabulary, and that most

students reported gains in affective and cognitive engagement. Meta-analytic evidence from Wang and Fan (2025) indicates a moderate positive effect of ChatGPT on students' learning performance overall. However, other findings complicate this picture. Han et al. (2023) found that while AI feedback improved surface-level features, critical thinking and argument development did not consistently improve. Alzubi et al. (2025) noted that both ESL students and teachers perceived AI as effective for grammar and vocabulary but less so for developing original arguments.

To systematically capture how students engage with AI tools, researchers have identified multiple dimensions of AI use. Tool selection reflects preferences for specific platforms (Aladini, 2023; Syarifah & Fakhruddin, 2024). Frequency of use indicates intensity of engagement (Chang et al., 2021; Yi & Kim, 2023). Context and purpose of use capture when and why students deploy AI, for grammar checking, feedback, idea expansion, translation, or proofreading (Bok & Cho, 2023; Han et al., 2023; Gasaymeh et al., 2024; Thangthong et al., 2024). Interaction patterns, adapted from Lin's (2025) five-level framework of LLM writing assistance, distinguish direct assistance (AI generates content), indirect assistance (AI provides suggestions), and mixed approaches. Content management addresses how students prepare prompts, modify AI output, and verify its accuracy (Aladini, 2023; Syarifah & Fakhruddin, 2024; Jin et al., 2025). Technical aspects capture challenges encountered during AI use (Utami et al., 2023).

The Relationship Between Generative AI Use and Creativity

From a theoretical standpoint, AI tools may serve as mediating artifacts that scaffold creative processes, particularly for students with limited linguistic resources. Vygotsky's theory of mediated action suggests that technology can extend cognitive capacity during task performance. However, Salomon's (1997) distinction between effects *with* technology (performance enhancement during use) and effects *of* technology (lasting cognitive change) raises the possibility that AI reliance may produce polished but derivative outputs. Kaufman and Beghetto's (2009) Four C Model further suggests that AI tools can support little-c or pro-c creativity if used reflectively, but may not contribute to deeper creative development if used passively. These mechanisms would not be expected to act uniformly across divergent-thinking dimensions. Exposure to AI-generated language and framings could plausibly shift the statistical rarity of the ideas a writer produces (originality), whereas the breadth of conceptual categories a writer can access (flexibility) draws on existing knowledge structures that a single AI-assisted session is unlikely to reshape, providing a theoretical reason to expect dimension-specific rather than global associations.

Empirical evidence on this relationship is mixed. Several studies support a facilitative role. De Vicente-Yagüe-Jara et al. (2023) found that ChatGPT assistance improved students' fluency, flexibility, and originality scores in a creative writing task. Han et al. (2023) reported that Korean university students used ChatGPT for idea clarification and content expansion, helping them overcome writer's block. Habib et al. (2024) reported that AI interaction had a positive effect on creative output, particularly when students used AI as a brainstorming partner rather than a content generator. Doshi and Hauser (2024) experimentally demonstrated that AI-generated story ideas made individual outputs more creative, better written, and more enjoyable, particularly for lower-ability

writers. Jin et al. (2025), in a structural equation modeling study of 1,073 postgraduates across 21 countries, found that advanced, transformative uses of AI positively predicted writing improvement and critical thinking, mediated by self-regulated learning strategies. Hongxia and Razali (2026) reinforced this point from a different angle: in a 16-week study of 60 Chinese EFL undergraduates, AI-assisted peer feedback improved self-regulated learning, self-efficacy, and motivation more than conventional teacher feedback, with medium-to-high effect sizes. Aligned with Jin et al. (2025), this suggests that the key is not AI exposure but the self-regulatory work students do around it. Moorhouse et al. (2025) found that postgraduate L2 writers positioned themselves as leaders of AI, maintaining control over content while using tools for linguistic polishing. Liu and Reinders (2025), in a preliminary longitudinal study, observed that learners' motivation rose significantly only after a generative-AI chatbot was introduced midway through a 15-week course, suggesting that some benefits of AI use emerge gradually and may be missed by a single cross-sectional snapshot.

Conversely, other research highlights risks. Bok and Cho (2023) observed over-reliance among Korean EFL students, where learners substituted critical thinking with AI-generated suggestions rather than engaging in reflective revision. Lin (2025) argued that students lacking metacognitive awareness may become passive recipients of AI-generated language, which dilutes individual voice. Doshi and Hauser (2024) also showed a critical trade-off: while AI boosted individual creativity, AI-aided stories were more similar to each other than stories by humans alone, indicating reduced collective diversity. Corpus-driven evidence gives this homogenization concern a concrete linguistic basis: Tudino and Qin (2024) found that ChatGPT-generated academic texts were syntactically and semantically more uniform than human-authored writing. Huang and Wu (2025), in a study of 337 Chinese university students, found that decreased independent and innovative thinking was the most commonly reported consequence of excessive AI reliance. Zawacki-Richter et al. (2019) and Zhai et al. (2024) further cautioned that uncritical acceptance of AI outputs risks eroding higher-order cognitive skills.

Notably, several studies suggest that the impact of AI on creativity is conditional on how students use these tools. Syarifah and Fakhruddin (2024) found that students who used AI selectively reported enhanced creativity, while those who used it as a one-step solution showed lower engagement. Chang et al. (2021) observed improved grammar and structure through Grammarly use but no significant gains in creativity-related indicators such as lexical diversity. Jin et al. (2025) demonstrated that self-regulated learning strategies mediate the relationship between AI-assisted writing and writing outcomes, implying that active, strategic engagement with AI, rather than passive acceptance, is the critical factor. This conditional relationship aligns with Moorhouse et al. (2025), who found that AI tools can empower L2 writers when used to support rather than replace their own writing processes.

Despite the growing body of work, significant gaps remain. Most studies focus on writing quality or student perceptions rather than creativity outcomes per se. Very few employ standardized creativity frameworks measuring fluency, flexibility, and originality, nor do they quantify how multiple dimensions of AI use relate to creative

performance. Research contextualized in Southeast Asian EFL university settings is particularly scarce. The present study is therefore distinctive in three respects: (1) it operationalizes creativity through a standardized divergent-thinking framework rather than holistic writing-quality ratings, (2) it relates creativity to seven separable dimensions of AI use rather than to AI use in general, and (3) it does so in an under-represented Southeast Asian EFL context.

METHOD

Research Design

This study employed a quantitative correlational design, incorporating qualitative coding procedures, to examine the relationship between creativity and the use of generative AI in academic writing. Quantitative methods were used to measure creativity performance and to analyze Likert-scale questionnaire data, while qualitative coding procedures were used to segment and score idea units in the AUT test.

Participants and Research Site

The study was conducted at a public university in Ho Chi Minh City, Vietnam. The target population was undergraduate students enrolled in the Bachelor of Arts in English Language program. As long as they were learning or had completed any writing courses, they were eligible to participate. Participants were recruited through purposive sampling, wherein all potential participants were contacted during break time at class. As this research requires participants to be equipped with writing skills or exposed to academic writing courses, purposive sampling enables the researcher to select participants who are especially knowledgeable about or experienced with the phenomenon of interest (Etikan et al., 2016). All participants were fully informed of the procedure and given a consent form to sign. The sample comprised 78 valid responses from 84 collected. Table 1 summarizes participant characteristics.

Table 1
Participant characteristics (N=78)

Characteristic	n	%
Gender		
Male	23	29.5
Female	52	66.7
Other/Prefer not to say	3	3.8
Academic Year		
Year 1	17	21.8
Year 2	18	23.1
Year 3	15	19.2
Year 4	28	35.9
Program Type		
Standard	27	34.6
High-quality	39	50.0
English Language only	12	15.4

Instruments

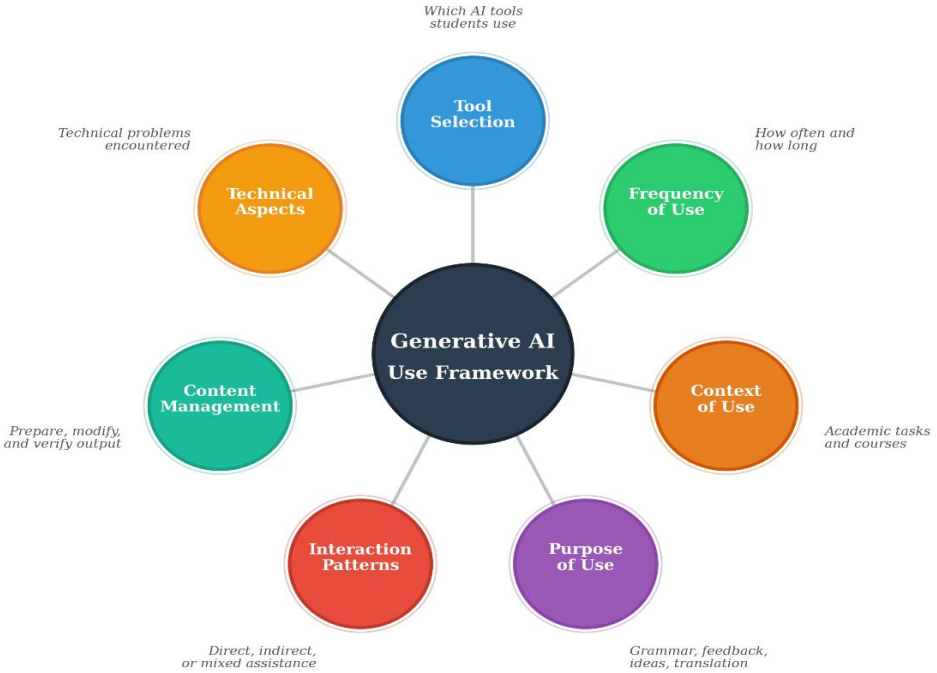
Creativity measurement: Modified AUT

The Alternative Uses Task (AUT) is a widely recognized instrument for assessing divergent thinking, which is considered a key component of creative potential. Originally developed by Guilford (1968), the AUT measures an individual's ability to generate multiple and varied responses to a given stimulus, typically involving a familiar object such as a brick. Participants are instructed to list as many alternative uses for the object as possible within a fixed time frame, usually between two to five minutes (Pipes, 2023). For the purposes of this study, the original format of the AUT was modified to better align with the context of academic writing. Rather than generating uses for a physical object, participants were asked to produce as many ideas as possible in response to a given academic writing prompt, allowing for the assessment of divergent thinking within a discipline-specific task. Participants were given a topic based on an IELTS Writing Task 2 sample question. They were required to list as many answers to the topic as possible within the span of 10 minutes. The test was administered through Google Forms, and participants were required to attend an online meeting via Microsoft Teams while they completed the task to receive instructions and to prevent cheating. Idea-unit segmentation, categorization, and originality scoring were carried out jointly by the researchers. The researchers worked through every response together and settled each segmentation and categorization decision by discussion until they agreed. This consensus procedure covered all 78 responses, so every score reflects agreement between coders, and segmentation followed the explicit eight-rule framework adapted from Gecchele et al. (2019).

AI usage questionnaire

To investigate the correlation between creativity and the use of generative AI in academic writing, this study employs a questionnaire designed around seven dimensions (see Figure 2) derived from prior research: tool selection (Aladini, 2023; Syarifah & Fakhruddin, 2024), frequency of use (Aladini, 2023; Chang et al., 2021; Yi & Kim, 2023), context of use (Han et al., 2023), purpose of use (Bok & Cho, 2023; Han et al., 2023; Syarifah & Fakhruddin, 2024), interaction patterns (Lin, 2025), content management (Aladini, 2023; Syarifah & Fakhruddin, 2024), and technical aspects (Utami et al., 2023). *Tool selection* examines the variety and frequency of AI tools used to understand individual preferences. Nine popular tools were examined: ChatGPT, Grammarly, QuillBot, Gemini, Copilot, Story AI, Jenni AI, Jasper, and Chibi. *Frequency* captures how often students use AI tools and the average duration of each session, reflecting overall intensity of use. The *context* dimension explores the types of academic tasks and courses where AI is typically applied. The *purpose of use* assesses how students use AI for different purposes, including correcting grammar, asking for feedback, generating content, and translating. *Interaction patterns* distinguish between direct use (AI editing the text), indirect use (AI suggesting edits), or a combination of both. *Content management* examines how students prepare content before using AI, how frequently they edit AI-generated output, and the extent to which they verify its accuracy. Finally, the *technical aspect* measures how often students face technical issues while using AI tools.

Figure 2
Generative AI use dimensions



In total the questionnaire comprised 35 items distributed across the seven dimensions. Most were rated on five-point Likert-type scales, ranging from 1 (Never) to 5 (Usually) for the frequency-based items and from 1 (Totally disagree) to 5 (Totally agree) for the interaction and content-management statements, alongside two ordinal items recording session and correction duration. Internal consistency was estimated on the present sample using Cronbach's alpha (Cronbach, 1951; Tavakol & Dennick, 2011; see Table 2). Five of the six indices reached or approached the conventional .70 threshold, and the questionnaire as a whole was reliable ($\alpha = .87$). Content management was the exception ($\alpha = .42$), and the reason is substantive rather than psychometric. Its three items describe distinct acts performed at different points in the writing process: preparing material before turning to AI, reshaping AI output, and checking that output for accuracy. There is no reason these should rise and fall together, so they were treated as separate indicators throughout, never summed into a scale, which is how the correlational analysis in fact uses them.

Table 2
Internal consistency reliability of the AI usage questionnaire dimensions

Dimension	No. of items	α
Tool selection	9	.71
Context of use	6	.79
Purpose of use	8	.78
Interaction patterns	4	.61
Content management	3	.42
Overall questionnaire	30	.87

Note. Cronbach's α computed on the present sample (N = 78). Frequency-and-duration and technical-aspect items are single-item indicators and are not entered into dimension-level α. The full modified AUT prompt and the complete AI usage questionnaire are available as supplementary online material (see Data Availability statement).

Pilot study

To ensure the reliability and validity of the research framework in an online environment, a pilot study was conducted with 7 participants, of whom 6 provided valid responses. This preliminary phase served multiple purposes: testing the stability of data collection methods, identifying potential issues with the questionnaire, refining data collection procedures, and assessing the feasibility of the 30-minute time allocation for the official test. Based on participant feedback, minor revisions were made to enhance clarity and better reflect actual AI usage behaviors. The successful completion within the allocated 30 minutes validated the timing for the official test.

Data Collection Procedure

Data collection occurred in a single session per participant. Each participant first joined a Microsoft Teams meeting for instructions, then completed the AUT via Google Forms (10 minutes), followed by the AI usage questionnaire via Fillout. Both instruments were confirmed as received before participants were dismissed. Online testing meetings required all participants to turn on their cameras to maintain assessment integrity; no AI use was allowed during the AUT test. After data collection, the results of the AUT test were coded using the idea unit framework and the questionnaire results were coded for analysis. Sessions ran between April and May 2025.

Data Analysis

Analyzing creativity

This study adapts the Alternative Uses Task to suit the specific context of assessing creativity in academic writing. Three out of four factors examined in the AUT were used: fluency (F_1), flexibility (F_2), and originality (O).

To count the ideas generated by participants, one idea must be defined and regarded as an *idea unit.* Gecchele et al. (2019) defined an idea unit as a textual fragment that logically expresses an idea or thought. An idea unit segmentation framework was proposed, analyzed based on semantic meaning, with the following rules adapted from Gecchele et al. (2019): (1) A single sentence has at least one main clause. (2) A compound sentence: if all clauses support one main idea, it is one idea unit; if a clause introduces a novel idea, it is a separate unit. (3) A complex sentence: if the dependent

clause carries additional information to elaborate the main clause, it is not counted separately, and vice versa. (4) In listing statements, each item is counted as one idea unit; repeated items are excluded. (5) Citations and examples are not counted if they align with the main clause; they are counted if they introduce new ideas. (6) Prepositional phrases, (7) infinitive clauses, and (8) bracketed content are not counted as separate idea units.

Fluency is defined as the total number of unique idea units generated in a single response. Each idea unit counted for one fluency point; duplicated units were disregarded. Fluency was calculated using the formula below.

$$F_1 = \sum i$$

where:
- F_1 (points) is the fluency point
- i is the idea unit(s) generated in a single response

Flexibility was calculated by categorizing the idea units from one participant, with each distinct category counting for one flexibility point.

$$F_2 = \sum cat$$

where:
- F_2 (points) is the flexibility point
- ***cat*** is the category

Originality was scored based on the statistical rarity of each idea category across all participants.

$$C_r = \frac{f_x}{n} \times 100$$

where:
C_r (%) is the originality rarity of the x category,
f_x (times) is the frequency of the category that idea unit x_n belongs to
n (ideas) is the number of all idea units collected.

We then scored each category based on its C_r according to a previous scoring system adapted from Alhashim et al. (2020) and Habib et al. (2024) in this manner:

- If $C_r \leq 3\%$ then $O_{cat} = 3\ pts$
- If $3\% < C_r \leq 5\%$ then $O_{cat} = 2\ pts$
- If $5\% < C_r \leq 10\%$ then $O_{cat} = 1\ pt$
- If $C_r > 10\%$ then $O_{cat} = 0\ pts$

Note. *O_{cat} (pts) is the originality point of C_r*

After we calculate the originality point for each category, we continued to calculate the originality point for each participant using the following formula:

$$O_x = \sum_{i=1}^{n} (N_i \times O_{cat_i})$$

where:
O_x (point) is the originality point of participant x
x is the participant index
i is the category index
n the index of the final category
N_i is the total IU of one participant in category i
O_{cat_i} is the originality point of category i

Statistical analysis.

All analyses were conducted using MS. Excel 365 and Jamovi version 2.6.44. Descriptive statistics were computed for all variables. The Shapiro-Wilk test assessed distributional normality. Because all variables departed significantly from normality (all $p<.001$), Spearman rank-order correlations were used for all bivariate analyses. Correlation magnitudes were interpreted following Cohen (1988): weak ($|\rho|<.30$), moderate (.30-.50), and strong (>.50). Because the design generated 105 correlations (35 AI use variables crossed with three creativity dimensions), the Benjamini-Hochberg false discovery rate procedure was applied so that the number of significant results would not be inflated by chance alone.

FINDINGS

Results are organized around the two research questions. Shapiro-Wilk tests confirmed that all variables departed significantly from normality (all $p<.001$), justifying the use of nonparametric methods throughout.

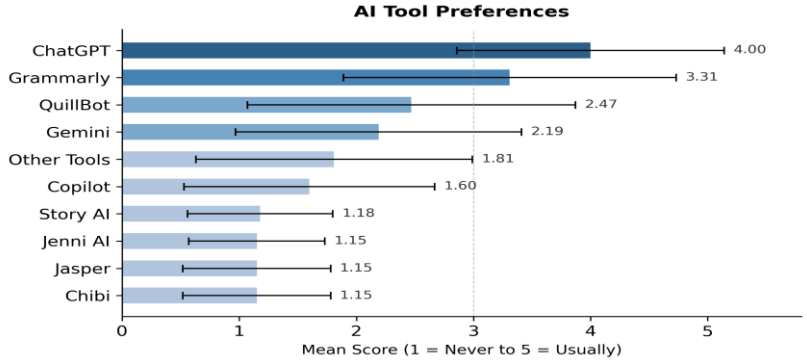
RQ1: How, and for what purpose(s), do English-major students use generative AI tools in academic writing?

Seven dimensions of generative AI use were examined: tool selection, frequency and duration, context of use, purpose of use, interaction patterns, content management, and technical aspects. Descriptive results for each dimension are reported below.

Tool Selection

Figure 3 visualizes average usage for each AI tool. ChatGPT emerged as the most frequently used tool ($M = 4.00$, $SD = 1.14$), corresponding to the "often" level on the response scale. Grammarly followed ($M = 3.31$, $SD = 1.42$), with notable variability suggesting divergent adoption patterns among participants. QuillBot was moderately used ($M = 2.47$, $SD = 1.40$), and Gemini showed limited adoption ($M = 2.19$, $SD = 1.22$). The remaining tools, Copilot ($M = 1.60$), Story AI ($M = 1.18$), Jenni AI ($M = 1.15$), Jasper ($M = 1.15$), and Chibi ($M = 1.15$), were rarely used, all with medians of 1.00 (i.e., "never"), indicating that the vast majority of participants did not engage with these platforms. This distribution reveals a strong concentration around two dominant tools, ChatGPT and Grammarly, with most other AI writing tools remaining peripheral in students' academic writing practices.

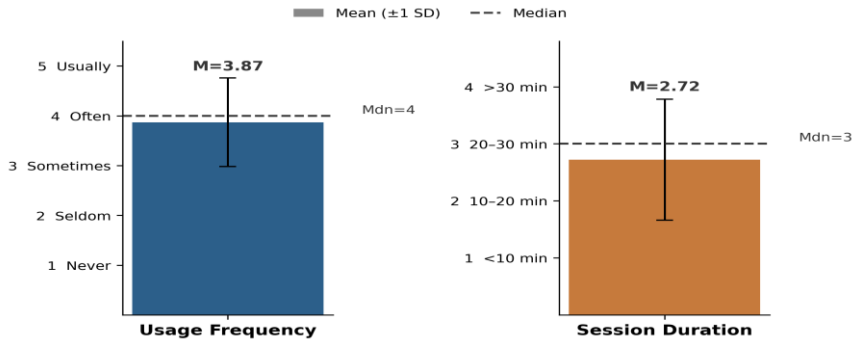
Figure 3
AI tool preferences (N = 78)



Frequency and Duration of Use

AI usage frequency was generally high (M = 3.87, SD = 0.89, Mdn = 4.00), corresponding to "often" on the response scale (see Figure 4). The majority of participants reported frequent use: 46.2% selected "often" (n = 36) and 24.4% selected "usually" (n = 19), while only 6.4% reported using AI seldom or never. Session duration averaged 2.72 (SD = 1.06, Mdn = 3.00), corresponding to approximately 20-30 minutes per session. Duration was fairly evenly distributed: 15.4% of students used AI for under 10 minutes, 26.9% for 10-20 minutes, 28.2% for 20-30 minutes, and 29.5% for over 30 minutes. These findings indicate that AI use is not merely frequent but also sustained, with most students engaging for nontrivial periods during each writing session.

Figure 4
Frequency and duration of AI tool use (N = 78)

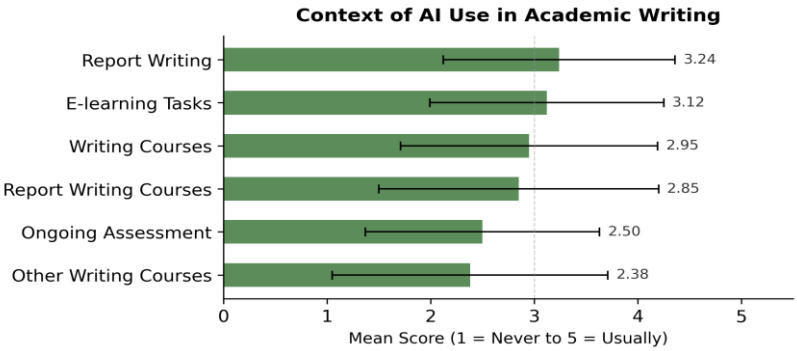


Context of Use

Figure 5 presents the frequency of AI use across six academic contexts. Report writing assignments received the highest mean AI use (M = 3.24, SD = 1.12), followed by e-

learning tasks (*M* = 3.12, *SD* = 1.13) and writing courses (*M* = 2.95, *SD* = 1.24). Report writing courses (*M* = 2.85, *SD* = 1.35), ongoing assessment (*M* = 2.50, *SD* = 1.13), and other writing courses (*M* = 2.38, *SD* = 1.33) scored lower. All context means fell at or below the "sometimes" level except for report writing and e-learning tasks, which slightly exceeded it. This pattern suggests that students deploy AI tools more readily for assignments requiring substantial textual output (reports, e-learning submissions) than for ongoing assessments or specialized writing courses, where task constraints or instructor monitoring may limit AI use.

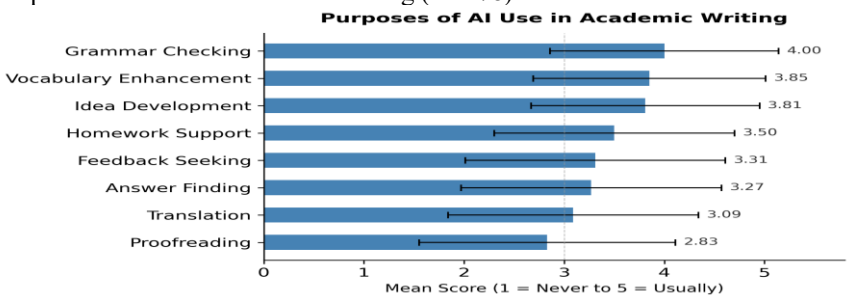
Figure 5
Context of AI use in academic writing (N = 78)



Purpose of Use

Figure 6 presents the frequency of AI use across eight purposes. Grammar checking was the most frequent purpose (*M* = 4.00, *SD* = 1.14), closely followed by vocabulary enhancement (*M* = 3.85, *SD* = 1.16) and idea development (*M* = 3.81, *SD* = 1.14). These three purposes all had medians of 4.00 ("often"), confirming their centrality in students' AI-assisted writing. Homework support (*M* = 3.50), feedback seeking (*M* = 3.31), answer finding (*M* = 3.27), and translation (*M* = 3.09) occupied middle positions, while proofreading was the least frequently reported purpose (*M* = 2.83, *SD* = 1.28). This hierarchy indicates that students primarily use AI for language accuracy and idea generation rather than for final-stage quality assurance, suggesting an orientation toward production-phase support over revision-phase support.

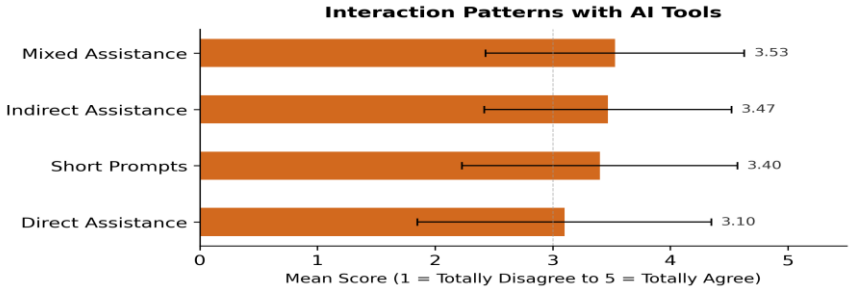
Figure 6
Purposes of AI use in academic writing (N = 78)



Interaction Patterns

Figure 7 summarizes interaction patterns. Mixed assistance, in which AI provides feedback and then edits the student's writing, received the highest mean score (M = 3.53, SD = 1.10), followed by indirect assistance (M = 3.47, SD = 1.05) and short prompts (M = 3.40, SD = 1.17). Direct assistance, in which AI intervenes and edits the student's writing, was the least preferred interaction mode (M = 3.10, SD = 1.25). These results indicate a preference for collaborative rather than delegative AI interaction: students favor receiving AI-generated suggestions and combined feedback-editing cycles over allowing AI to directly modify their texts, reflecting a degree of authorial agency in the writing process.

Figure 7
Interaction patterns with AI tools (N = 78)

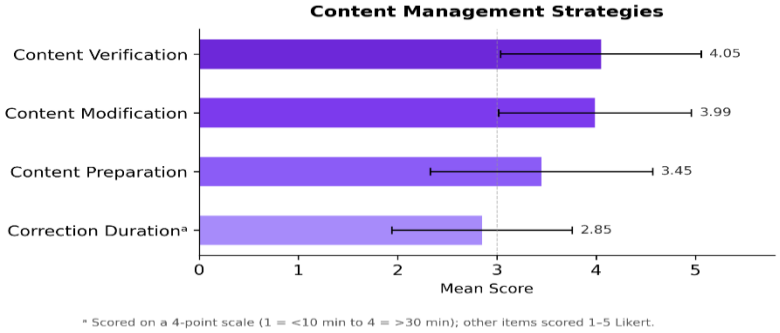


Content Management

Figure 8 presents content management strategies. Content verification received the highest mean score (M = 4.05, SD = 1.01), indicating that most participants regularly verified the accuracy of AI-generated content. Content modification scored similarly high (M = 3.99, SD = 0.97), suggesting that students frequently adapted AI output to fit their purposes. Content preparation was somewhat lower (M = 3.45, SD = 1.12), indicating that while students prepared content before using AI, they did so less consistently than they modified or verified AI output afterward. The mean correction

duration was 2.85 (*SD* = 0.91), corresponding to approximately 20-30 minutes on a 4-point scale, with 60.3% of participants spending 20 minutes or more correcting AI-generated writing. This profile of high post-generation engagement—actively modifying, verifying, and spending substantial time correcting AI output—depicts students who treat AI contributions as preliminary drafts requiring critical evaluation rather than finished products.

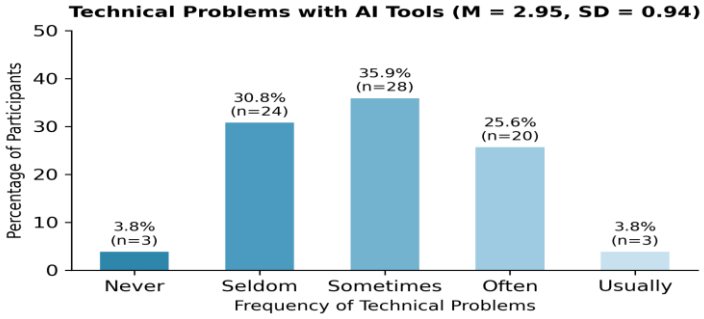
Figure 8
Content management strategies (N = 78)



Technical Aspects

Technical problem frequency averaged 2.95 (*SD* = 0.94, *Mdn* = 3.00), corresponding to "sometimes" on the response scale (see Figure 9). The distribution was roughly symmetric: 34.6% of participants reported encountering technical problems "seldom" and 35.9% "sometimes," while 25.6% reported "often" and 3.8% "usually." Only 3.8% never encountered any issues. These findings indicate that technical difficulties with AI tools are a common but not dominant concern, representing a non-negligible friction point in AI-assisted writing workflows.

Figure 9
Frequency of technical problems with AI tools (N = 78)



RQ2: Are there any correlations between the use of generative AI tools and students' creativity performance in academic writing?

Table 3 presents the results of creativity analyses. Fluency ranged from 1 to 19 (*M*=6.36, *SD*=3.41). Flexibility ranged from 1 to 7 (*M*=3.86, *SD*=1.24). Originality was notably low and right-skewed (*M*=0.92, *SD*=1.40, *Mdn*=0.00), indicating that most participants generated common rather than statistically rare idea categories. As Figure 10 demonstrates, the pronounced gap between the mean and median for originality reflects the strong positive skew in this dimension, with the majority of participants scoring zero.

Table 3
Descriptive statistics for creativity metrics (N=78)

Dimension	M	Mdn	SD	Min	Max	W	p
Fluency	6.36	6.00	3.41	1	19	.832	< .001
Flexibility	3.86	4.00	1.24	1	7	.929	< .001
Originality	0.92	0.00	1.40	0	6	.693	< .001

Note. Fluency=count of idea units; Flexibility=count of distinct categories; Originality=sum of weighted rarity scores.

Figure 10
Creativity metrics

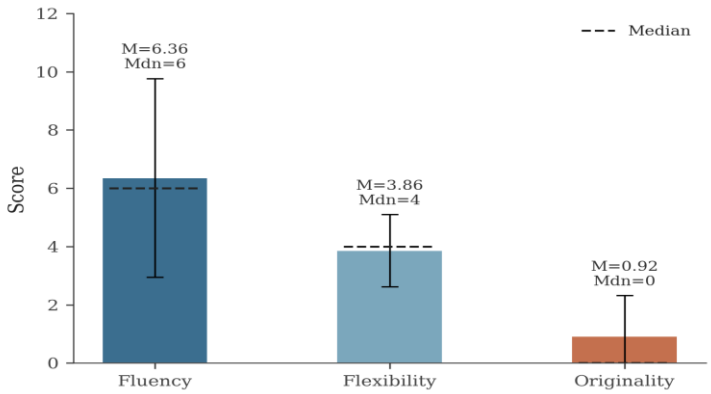


Table 4 (and see also Figure 11) summarizes all statistically significant Spearman correlations. Of the 105 correlations tested (35 AI variables × 3 creativity dimensions), nine reached significance at α=.05. Eight of these coefficients fell in the weak range; the strongest, the negative correlation between Copilot use and fluency (ρ = -.312), sat just over the conventional cut-off and is best read as weak-to-moderate (|ρ| = .225-.312).

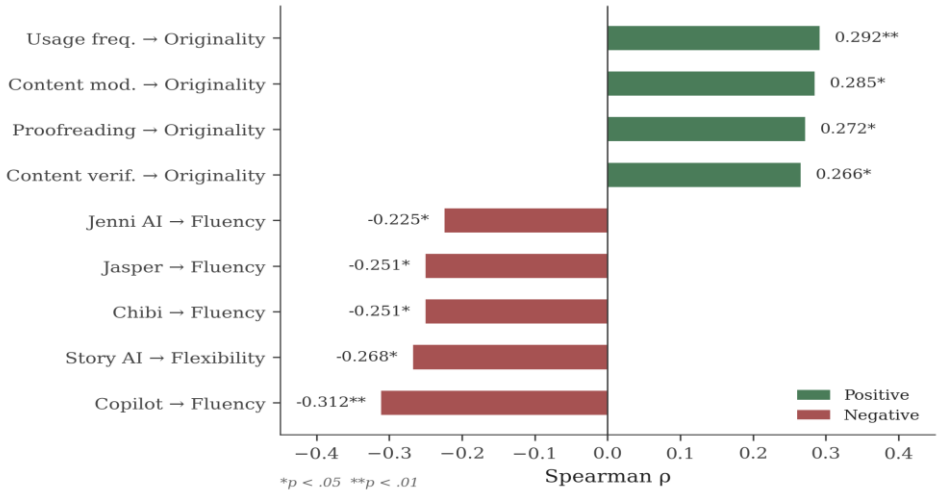
One caution attaches to this count. Testing 105 correlations inflates the chance that some will reach significance by accident, so the Benjamini-Hochberg procedure was applied to control the false discovery rate. At a false discovery rate of .05, none of the nine correlations survived correction. The associations reported below should therefore be read as exploratory and hypothesis-generating rather than as confirmed effects, a point the discussion returns to.

Table 4
Statistically significant Spearman correlations between AI use and creativity (N=78)

AI Use Variable	Creativity	Spearman ρ	p-value	Strength
Copilot use	Fluency	-.312	.005	Weak-moderate
Jenni AI use	Fluency	-.225	.048	Weak
Jasper use	Fluency	-.251	.026	Weak
Chibi use	Fluency	-.251	.026	Weak
Story AI use	Flexibility	-.268	.018	Weak
Usage frequency	Originality	.292	.009	Weak
Proofreading	Originality	.272	.016	Weak
Content modification	Originality	.285	.011	Weak
Content verification	Originality	.266	.019	Weak

Note. Only correlations significant at α=.05 are presented. Strength: weak=|ρ|<.30; moderate=.30-.50 (Cohen, 1988). The Copilot-fluency coefficient (ρ=-.312) lies just above the .30 cut-off and is interpreted as weak-to-moderate. None of the nine correlations remained significant after Benjamini-Hochberg correction for the 105 tests.

Figure 11
Significant Spearman correlations between AI use variables and creativity dimensions



Tool selection and creativity. Among the ten AI tools, four showed significant negative correlations with fluency: Copilot (ρ=-.312, *p*=.005), Jasper (ρ=-.251, *p*=.026), Chibi (ρ=-.251, *p*=.026), and Jenni AI (ρ=-.225, *p*=.048). Story AI correlated negatively with flexibility (ρ=-.268, *p*=.018). No tool showed a significant association with originality. ChatGPT and Grammarly, the two most widely used tools, showed no significant correlations with any creativity dimension. These negative correlations involve tools used by very few participants (medians at 1.00), which limits interpretive weight. Because these tools were used by almost no one, their coefficients are likely artefacts of restricted range and near-zero variance, and should not be read as evidence that the tools harm fluency or flexibility.

Usage frequency and creativity. AI usage frequency was positively associated with originality (ρ=.292, *p*=.009) but showed no significant relationship with fluency (ρ=.059, *p*=.605) or flexibility (ρ=.114, *p*=.320). Session duration was not correlated with any creativity dimension (all *p*>.29).

Context of use and creativity. No significant correlations were found between the six academic contexts and any creativity dimension (all *p*>.05). The association between e-learning frequency and originality approached significance (ρ=.220, *p*=.053).

Purpose of use and creativity. Among eight purposes, only proofreading showed a significant correlation, specifically with originality (ρ=.272, *p*=.016).

Interaction patterns and creativity. None of the four interaction variables reached significance with any creativity dimension (all *p*>.10).

Content management and creativity. Content modification (ρ=.285, *p*=.011) and content verification (ρ=.266, *p*=.019) were significantly associated with originality. Content preparation approached significance with flexibility (ρ=-.202, *p*=.075). Time spent correcting AI output was unrelated to any creativity dimension.

Technical aspects and creativity. Technical problem frequency was not significantly correlated with any creativity measure, though its association with fluency approached significance (ρ=.215, *p*=.059).

DISCUSSION

AI use in academic writing among Vietnamese EFL students

The first research question asked how and for what purposes English majors use generative AI tools in academic writing. The results show that AI use is a routine part of these students' writing practices. ChatGPT (*M* = 4.00) and Grammarly (*M* = 3.31) were the most commonly used tools, while the remaining tools were rarely or never used by most participants. This pattern of concentrated usage around a small number of well-known platforms is consistent with global survey data (Kelly, 2024) and with findings in other Southeast Asian EFL settings. Thangthong et al. (2024) observed that Thai EFL undergraduates relied primarily on Grammarly for vocabulary and grammar assistance, and Syarifah and Fakhruddin (2024) documented similar patterns among Indonesian EFL students. The strong preference for ChatGPT and Grammarly in this sample is likely explained by their wide availability, free or low-cost access tiers, and the breadth of writing support they offer, from grammar checking to idea generation. Specialized tools such as Jasper, Jenni AI, and Chibi may be less well known among Vietnamese undergraduates, or may be perceived as unnecessary given the broad functionality of ChatGPT.

The three most frequent purposes of AI use were grammar checking (*M* = 4.00), vocabulary enhancement (*M* = 3.85), and idea development (*M* = 3.81). These primary use purposes align well with the linguistic difficulties that Vietnamese EFL students experience. Dang (2019) reported that over half of Vietnamese EFL students cited a lack of ideas as a major difficulty in academic writing, and Moses and Mohamad (2019) identified limited vocabulary and weak grammar as the most common barriers to

effective writing in ESL/EFL contexts. The students in this sample appear to be using AI tools to address exactly these difficulties. It is worth noting that idea development was reported almost as frequently as grammar checking. This means that the students were not only using AI for surface-level language correction but also for the kind of higher-order cognitive work that Kormos et al. (2024) identified as a predictor of L2 writing quality. The relatively low frequency of proofreading (M = 2.83) compared to grammar checking (M = 4.00) suggests that students distinguish between targeted error correction during drafting and holistic review of a finished text. They appear to turn to AI more during the production phase of writing than during the final revision phase.

AI was used most frequently for report writing and e-learning tasks, and least for ongoing assessments and other writing courses. This context-dependent pattern could be explained by the nature of these tasks. Report writing and e-learning assignments typically involve extended, autonomous writing where students have time and opportunity to consult AI tools. Ongoing assessments, by contrast, are often conducted under supervised or timed conditions that limit opportunities for AI use. The lower AI use in specialized writing courses may also reflect instructor expectations or explicit restrictions on AI in those courses, consistent with Alshreef and Saaty's (2025) observation that generative AI has changed writing classroom dynamics and prompted some instructors to set boundaries on its use.

Two findings about how students interact with AI are particularly informative. First, students preferred mixed assistance (M = 3.53) and indirect assistance (M = 3.47) over direct assistance (M = 3.10). In other words, they preferred AI to give suggestions and feedback rather than to directly edit their texts. This preference indicates that the students in this sample maintained a degree of authorial control over their writing. Moorhouse et al. (2025) reported a similar pattern among postgraduate L2 writers who positioned themselves as "leaders" of AI, keeping authority over content decisions while using tools for linguistic support. Second, content verification (M = 4.05) and content modification (M = 3.99) scored high, and over 60% of participants spent 20 minutes or more correcting AI output per session. This suggests that the students did not accept AI-generated content passively but actively evaluated and revised it. This behavior contrasts with the uncritical acceptance that Bok and Cho (2023) observed among Korean EFL students. However, since these data are self-reported, the possibility of social desirability bias should be considered. Students may have overstated the extent of their verification and modification behaviors.

The relationship between AI use and creativity

The study also investigated whether there is a relationship between generative AI use and students' creativity in academic writing. The overall answer is that the relationship is weak and largely absent. Of the 105 correlations tested across 35 AI use variables and three creativity dimensions, only nine reached statistical significance at $\alpha = .05$, and all of them fell in the weak range. Moreover, once the 105 tests were corrected for multiple comparisons using the Benjamini-Hochberg false discovery rate, none of these nine correlations remained significant, which reinforces rather than weakens this conclusion. This means that most AI use dimensions, including tool selection, academic context,

interaction patterns, and most purposes of use, were not significantly associated with fluency, flexibility, or originality. The link between generative AI use and creativity in EFL academic writing, at least as captured by the measures in this study, is not clear.

The few significant correlations that did emerge clustered around originality. Content modification (ρ=.285, *p*=.011), content verification (ρ=.266, *p*=.019), proofreading (ρ=.272, *p*=.016), and usage frequency (ρ=.292, *p*=.009) were each weakly and positively associated with originality. No AI use variable was significantly associated with flexibility. These correlations, while statistically significant, explain only a small portion of the variance in originality scores (approximately 7-9%), which limits the strength of any conclusion that can be drawn from them.

One possible explanation for the association between content management behaviors and originality is that the cognitive processes involved in modifying and verifying AI output overlap with the evaluative thinking that contributes to original idea generation. In Salomon's (1997) framework, technology use can produce effects *with* the tool (performance improvement during use) or effects *of* the tool (lasting cognitive change through internalized strategies). Students who routinely evaluate and transform AI output may be developing, or may already possess, the kind of critical evaluation skills that support original thinking. This interpretation is consistent with Jin et al.'s (2025) finding that self-regulated learning strategies mediated the relationship between AI-assisted writing and critical thinking. However, the correlational design of the present study does not allow us to determine the direction of this relationship. It is equally possible that students who are already more original in their thinking are also the ones who are more inclined to critically evaluate and modify AI output, rather than the reverse.

The positive association between usage frequency and originality is worth considering. Concerns have been raised that frequent AI use may reduce independent and innovative thinking (Huang & Wu, 2025; Zhai et al., 2024). The present finding does not support this concern, but it does not refute it either, because the effect is weak and specific to originality. One possibility is that students who use AI frequently are exposed to a wider range of vocabulary, phrasing, and conceptual framings than they would encounter in their own writing, and this exposure may contribute to the generation of less common ideas. Doshi and Hauser (2024) found experimentally that individuals who received AI-generated ideas produced more novel individual outputs, although collective diversity across participants decreased. A similar mechanism could operate here: frequent AI users may be incorporating less conventional framings into their thinking, which then show up as statistically rarer ideas in the originality scoring. However, it is also possible that the association simply reflects a third variable. For example, students who are more engaged with their studies in general may both use AI more frequently and think more creatively.

The large number of insignificant correlations is informative in itself. The fact that using AI tools, no matter what tool and how frequently it is used, was not correlated with any creativity dimension suggests that AI use does not in itself predict creative performance. Chang et al. (2021) similarly found that sustained Grammarly use

improved grammatical accuracy and cohesion but did not produce gains in creativity-related measures. The absence of significant correlations for interaction patterns indicates that whether students ask AI to edit their text, to suggest changes, or to combine both approaches makes no detectable difference to creativity. This is important because it suggests that pedagogical efforts to prescribe specific modes of AI interaction are unlikely to affect creative outcomes.

Flexibility was the creativity dimension least related to any AI use variable. This may reflect the nature of flexibility as a construct. Flexibility captures the breadth of conceptual categories a student can access during idea generation, which depends on the organization of knowledge in long-term memory (Kenett et al., 2018; Weiss & Wilhelm, 2022). Unlike originality, which can be influenced by exposure to novel stimuli, flexibility reflects the structural properties of a student's existing knowledge base. Within Kaufman and Beghetto's (2009) Four C Model, student-level creativity draws on personal knowledge structures that develop through sustained learning. AI tools may introduce novel vocabulary or ideas that affect originality, but they are unlikely to restructure a student's conceptual categories within a single writing session. If this reasoning is correct, then interventions aimed at improving flexibility in EFL writing would need to focus on building disciplinary knowledge and broadening students' intellectual engagement rather than on AI tool use.

It should also be noted that originality was markedly low in this sample ($M = 0.92$, $Mdn = 0.00$), with the majority of participants scoring zero. This means most students generated only idea categories shared by more than 10% of the sample. This floor effect is consistent with the well-documented difficulty Vietnamese EFL students have with generating ideas in academic writing (Dang, 2019), and it also limits the statistical power of the correlational analyses. The significant associations with originality are driven by the variation among the smaller group of students who did produce some original ideas. For the majority of participants, whose originality was zero regardless of AI use, the challenge is more basic: creating conditions under which original thinking in academic writing becomes possible in the first place. AI tools, however useful for grammar and vocabulary, do not appear to solve this more fundamental problem.

A further consideration concerns what the creativity task actually measures. The modified AUT indexes divergent, ideational creativity, that is, the capacity to generate many, varied, and statistically rare ideas in response to a prompt. This is related to, but not the same as, creativity in finished academic writing, which also involves argument development, organization, elaboration, voice, and rhetorical control. The associations reported here therefore speak to the ideational component of creativity rather than to creative quality in completed academic texts, and they should be interpreted within that scope.

Taken together, these findings do not support a strong facilitative relationship between AI use and creativity. De Vicente-Yagüe-Jara et al. (2023) found that ChatGPT assistance improved fluency, flexibility, and originality in a controlled experiment, but their study measured creativity during AI-assisted task performance, which captures what Salomon (1997) called effects *with* technology. The present study measured

creativity through a separate AUT administered independently of AI use, capturing the students' creative abilities outside of the AI-mediated writing environment. This design difference likely explains the discrepancy. The findings also do not support the strongly cautionary position of Bok and Cho (2023) or Huang and Wu (2025), who emphasized that AI use diminishes independent thinking. The students in this sample reported active engagement with AI rather than passive reliance. What the findings do suggest, consistent with Jin et al. (2025), Syarifah and Fakhruddin (2024), and Moorhouse et al. (2025), is that the relationship between AI and creativity is conditional: it depends on how students engage with AI output. However, the present data indicate that even this conditional relationship exists, and the practical significance of the observed correlations is limited. The overall picture that emerges is that generative AI use, as practiced by these Vietnamese EFL undergraduates, operates largely independently of their creative thinking in academic writing.

CONCLUSION

This study investigated how Vietnamese English-major undergraduates use generative AI in academic writing and whether their AI use patterns relate to creativity as measured by divergent thinking. The findings show that the students are active and purposeful AI users and that they frequently use common and free AI tools for grammar correction, vocabulary enhancement, and idea development. They also reported modifying and verifying AI-generated content when using them. However, the correlational analysis reveals that the relationship between AI use and creativity is weak and largely absent. Only a small number of weak positive correlations emerged between originality and specific post-generation behaviors (i.e. content modification, content verification, and proofreading), while most AI use dimensions were not associated with any creativity measure. The link between generative AI use and creativity in EFL academic writing is therefore not established by the present data.

For EFL writing pedagogy, these findings suggest that AI tools can serve as useful aids for addressing linguistic difficulties in L2 academic writing, but they should not be expected to enhance or diminish students' creative abilities. Pedagogical efforts would be better directed toward promoting students' evaluative habits that this study, though weakly, associates with originality, that is to teach students to critically assess, modify, and verify AI-generated content rather than to accept it at face value. The low originality scores in this sample point to a more fundamental challenge that AI tools cannot address. Developing the ideational breadth and disciplinary knowledge that support creative writing requires sustained investment in content-rich instruction, not merely better tool use.

Several limitations should be noted. The correlational design does not allow causal inferences. The self-reported AI use measures are subject to recall bias and social desirability. The creativity assessment relied on a single modified AUT administration, which excluded the elaboration dimension, and produced floor effects in originality that constrained statistical power. Aditionally, the modified AUT captures divergent, ideational creativity rather than the full construct of creativity in finished academic writing, which also encompasses argument development, organization, elaboration,

voice, and rhetorical control; the findings should be read with that scope in mind. The sample of 78 students from a single Vietnamese university limits generalizability, so the results should be read as specific to this population rather than as representative of EFL students in general. Future research should employ experimental or longitudinal designs to test whether specific AI engagement behaviors genuinely cultivate creative thinking over time. Additionally, log-based measures of AI interaction would provide more reliable data than self-reports. Finally, expanding the creativity framework to include elaboration and replicating the study across different EFL populations and proficiency levels would help clarify whether the patterns observed here hold in other contexts.

DECLARATIONS

The consent form included information about the study, the duration of the test and questionnaire, and projected benefits and drawbacks for participants. This form was approved by the institutional review board. Each participant was handed one consent form, given time to read it thoroughly before signing. The researchers signed to confirm mutual agreement. Participation was voluntary, and participants could withdraw at any time.

COMPETING INTERESTS

The authors declare no competing interests.

FUNDING

This research received no external funding.

AI USE DISCLOSURE

AI tools (Gemini and Grammarly) were used for language editing and grammar corrections only. All content was reviewed and verified by the authors, who take full responsibility for the final manuscript.

DATA AVAILABILITY

Data is available upon reasonable request to the corresponding author. Supplementary material is available at: https://osf.io/bz47p

REFERENCES

Aladini, A. (2023). AI applications impact on improving EFL university academic writing skills and their logical thinking. *Journal of Educational Sciences*, *31*(2), 25-44. https://doi.org/10.21608/ssj.2023.320166

Alhashim, G., Marshall, M., Hartog, T., Jonczyk, R., Dickson, D., van Hell, J., Okudan Kremer, G. E., & Siddique, Z. (2020, June). *WIP: Assessing Creativity of Alternative Uses Task Responses: A Detailed Procedure*. https://doi.org/10.18260/1-2--35612

Alshreef, N., & Saaty, N. (2025). The influence of students' independent use of generative AI tools on the EFL writing classroom. *International Journal of English Language Education, 13*(1), 152. https://doi.org/10.5296/ijele.v13i1.22832

Alzubi, A. A. F., Nazim, M., Alkhrisheh, H., Yahia, A. E., & Aldawsari, A. (2025). Perception of ESL students and teachers on the use of AI tools in English language learning. *International Journal of English Language Education, 13*(1), 79. https://doi.org/10.5296/ijele.v13i1.22660

Anderson, R. C., & Graham, M. (2021). Creative potential in flux. *Thinking Skills and Creativity, 40,* 100816. https://doi.org/10.1016/j.tsc.2021.100816

Bok, E., & Cho, Y. (2023). Examining Korean EFL college students' experiences and perceptions of using ChatGPT as a writing revision tool. *STEM Journal, 24*(4), 15-27. https://doi.org/10.16875/stem.2023.24.4.15

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI. *International Journal of Educational Technology in Higher Education, 20*(1), 43. https://doi.org/10.1186/s41239-023-00411-8

Chang, T.-S., Li, Y., Huang, H.-W., & Whitfield, B. (2021). *Exploring EFL students' writing performance and their acceptance of AI-based automated writing feedback.* Proceedings of the 2021 2nd International Conference on Education Development and Studies, 31-35. https://doi.org/10.1145/3459043.3459065

Chokwe, J. M. (2013). Factors impacting academic writing skills of English second language students. *Mediterranean Journal of Social Sciences, 4*(14), 377-383.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/BF02310555

Dang, T. N. A. (2019). EFL student's writing skills: Challenges and remedies. IOSR *Journal of Research & Method in Education, 9*(6), 74-84. https://doi.org/10.9790/7388-0906017484

de Vicente-Yagüe-Jara, M. I., López-Martínez, O., Navarro-Navarro, V., & Cuéllar-Santiago, F. (2023). Writing, creativity, and artificial intelligence. *Comunicar, 77*, 47-57. https://doi.org/10.3916/C77-2023-04

Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances, 10*(28). https://doi.org/10.1126/sciadv.adn5290

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1-4. https://doi.org/10.11648/j.ajtas.20160501.11

Freeman, J. (2025). *Student generative AI survey 2025.* HEPI. https://www.hepi.ac.uk/2025/02/26/student-generative-ai-survey-2025/

Ganapathy, M., Tan, D. A. L., & Phan, J. (2022). Narrative inquiry of ESL and EFL teachers' conceptions of creativity. *Journal of Research, Policy & Practice of Teachers and Teacher Education, 12*(1), 78-92.

Gasaymeh, A.-M. M., Beirat, M. A., & Abu Qbeita, A. A. (2024). University students' insights of generative artificial intelligence writing tools. *Education Sciences, 14*(10), 1062. https://doi.org/10.3390/educsci14101062

Gecchele, L., Rossi, R., & Spada, M. (2019). Idea unit analysis. *Reading in a Foreign Language, 31*(2), 210-229.

Guilford, J. P. (1968). *Intelligence, creativity and their educational implications* (1st ed.). Robert R. Knapp.

Habib, S., Vogel, T., & Anli, E. (2024). How does generative artificial intelligence impact student creativity? *Computers & Education, 222,* 105006.

Han, J., Yoo, H., Kim, Y., Myung, J., Kim, M., Lim, H., Kim, J., Lee, T. Y., Hong, H., Ahn, S.-Y., & Oh, A. (2023). RECIPE: How to integrate ChatGPT into EFL writing education. *Proceedings of the Tenth ACM Conference on Learning @ Scale,* 416-420. https://doi.org/10.1145/3573051.3596200

Hongxia, H., & Razali, A. B. (2025). Impact of ChatGPT on English academic writing ability and engagement of Chinese EFL undergraduates. *International Journal of Instruction, 18*(2), 323–346. https://doi.org/10.29333/iji.2025.18218a

Hongxia, H., & Razali, A. B. (2026). Impact of AI-assisted peer feedback on EFL students' self-regulated learning, self-efficacy, and motivation in English academic writing. *International Journal of Instruction, 19*(2), 163–182. https://doi.org/10.29333/iji.2026.1929a

Huang, T., & Wu, C. (2025). The chain mediating effect of academic anxiety and performance expectations between academic self-efficacy and generative AI reliance. *Computers and Education Open, 9,* 100234. https://doi.org/10.1016/j.caeo.2025.100234

Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level. *Contemporary Educational Technology, 15*(4). https://doi.org/10.30935/cedtech/13605

Jin, F., Lin, C.-H., & Lai, C. (2025). Modeling AI-assisted writing. *Computers in Human Behavior, 165,* 108538. https://doi.org/10.1016/j.chb.2024.108538

Jurecic, A. (2006). Is teaching writing still possible? *New Jersey Writing Alliance,* 1-13.

Kang, D. (2017). EFL learner and teacher beliefs about grammar learning in Korea. *English Teaching, 72*(2), 51-69. https://doi.org/10.15858/engtea.72.2.201706.51

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The Four C Model of creativity. *Review of General Psychology, 13*(1), 1-12. https://doi.org/10.1037/a0013688

Kelly, R. (2024, August 28). *Survey: 86% of students already use AI in their studies*. Campus Technology.

Kenett, Y. N., Levy, O., Kenett, D. Y., Stanley, H. E., Faust, M., & Havlin, S. (2018). Flexibility of thought in high creative individuals. *Proceedings of the National Academy of Sciences, 115*(5), 867-872. https://doi.org/10.1073/pnas.1717362115

Kormos, J., Suzuki, S., & Rossi, O. (2024). The role of creativity in second language writing performance. *Learning and Individual Differences, 114,* 102500. https://doi.org/10.1016/j.lindif.2024.102500

Lin, Z. (2025). Techniques for supercharging academic writing with generative AI. *Nature Biomedical Engineering, 9*(4), 426-431. https://doi.org/10.1038/s41551-024-01185-8

Liu, M., & Reinders, H. (2025). Do AI chatbots impact motivation? Insights from a preliminary longitudinal study. *System, 128,* 103544. https://doi.org/10.1016/j.system.2024.103544

Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills. Smart Learning Environments, 11(1), 9. https://doi.org/10.1186/s40561-024-00295-9

Moorhouse, B. L., Wan, Y., Wu, C., Wu, M., & Ho, T. Y. (2025). Generative AI tools and empowerment in L2 academic writing. *System, 133,* 103779. https://doi.org/10.1016/j.system.2025.103779

Moses, R. N., & Mohamad, M. (2019). Challenges faced by students and teachers on writing skills in ESL contexts. *Creative Education, 10*(13), 3385-3391. https://doi.org/10.4236/ce.2019.1013260

Pipes, R. (2023). Measuring creativity: A comprehensive review. Psychology of Aesthetics, Creativity, and the Arts.

Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? *Educational Psychologist, 39*(2), 83-96. https://doi.org/10.1207/s15326985ep3902_1

Plucker, J. A., Qian, M., & Wang, S. (2011). Is originality in the eye of the beholder? *The Journal of Creative Behavior, 45*(1), 1-22. https://doi.org/10.1002/j.2162-6057.2011.tb01081.x

Renzulli, J. S., Owen, S. V., & Callahan, C. M. (1974). Fluency, flexibility, and originality as a function of group size. *The Journal of Creative Behavior, 8*(2), 107-113.

Runco, M. A. (2010). Divergent thinking, creativity, and ideation. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity.* Cambridge University Press.

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal, 24*(1), 92-96. https://doi.org/10.1080/10400419.2012.650092

Salomon, G. (1997). *Distributed cognitions: Psychological and educational considerations.* Cambridge University Press.

Syarifah, E. F., & Fakhruddin, A. (2024). Exploring students' experience in using AI to assist their writing. *Journal of English Language Learning, 8*(1), 558-564. https://doi.org/10.31949/jell.v8i1.10028

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2,* 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Thangthong, P., Phiromsombut, J., & Imsa-Ard, P. (2024). Navigating AI writing assistance tools. *THAITESOL Journal, 37*(1), 111-131.

Torrance, E. P. (1966). *Torrance Tests of Creative Thinking: Norms-technical manual.* Personnel Press.

Utami, S. P. T., Andayani, Winarni, R., & Sumarwati. (2023). Utilization of artificial intelligence technology in an academic writing class. *Contemporary Educational Technology, 15*(4). https://doi.org/10.30935/cedtech/13419

Tudino, G., & Qin, Y. (2024). A corpus-driven comparative analysis of AI in academic discourse: Investigating ChatGPT-generated academic texts in social sciences. *Lingua, 312,* 103838. https://doi.org/10.1016/j.lingua.2024.103838

Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance. *Humanities and Social Sciences Communications, 12*(1), 621. https://doi.org/10.1057/s41599-025-04787-y

Weiss, S., & Wilhelm, O. (2022). Is flexibility more than fluency and originality? *Journal of Intelligence, 10*(4), 96. https://doi.org/10.3390/jintelligence10040096

Yi, J., & Kim, Y. (2023). Use of AI-based platforms in teaching and learning of general English writing for college EFL learners. *The Journal of Learner-Centered Curriculum and Instruction, 23*(7), 663-683. https://doi.org/10.22251/jlcci.2023.23.7.663

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

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