



Generative AI Use, Self-Regulated Learning, and Critical Thinking in EFL Writing: A Correlational Survey of Indonesian University Students

Siti Nur Diana^{1*}, Syaiful Islam²

^{1,2}Universitas Nurul Jadid, Probolinggo, Indonesia

ARTICLE INFO

Article history:

Submitted: Juny 26, 2026
Final Revised: July 22, 2026
Accepted: August 05, 2026
Published: September 31, 2026

Keywords:

Generative AI; Self- Regulated Learning; Critical Thinking; EFL Writing.



ABSTRACT

Purpose: This study aims to examine Indonesian university students' patterns of generative artificial intelligence (GenAI) use, critical thinking practices, self-regulated learning (SRL) behaviours, perceived benefits and challenges, and the relationships between GenAI use, SRL, and critical thinking in EFL writing. **Methods:** This study employed a quantitative descriptive-correlational survey design with a cross-sectional approach. Data were collected through an online questionnaire completed by 208 students enrolled in EFL-related programs at several Indonesian universities. The data were analysed using frequency distributions, percentages, means, standard deviations, the Shapiro-Wilk normality test, and Spearman's correlation to describe patterns of GenAI use and examine relationships among the variables. **Findings:** Nearly all respondents 99.0% of respondents had used GenAI for EFL writing assignments. ChatGPT was the most frequently used tool (75.0%), and planning was the writing stage in which GenAI was most commonly used (46.6%). Students reported high levels of critical thinking practices ($M = 4.30$, $SD = 0.34$) and SRL behaviours ($M = 4.22$, $SD = 0.34$). They also reported a high perceived risk of over-reliance on GenAI ($M = 3.90$, $SD = 0.73$). The frequency of GenAI use was not significantly associated with critical thinking practices ($p = 0.033$; $p = 0.638$). By contrast, SRL behaviours in GenAI use were positively and significantly associated with critical thinking practices ($p = 0.449$; $p < 0.001$). **Research Implications:** These findings indicate that EFL writing instruction should not focus solely on whether to allow or restrict the use of GenAI. Lecturers need to guide students in planning their use of AI, formulating purposeful prompts, verifying AI-generated information, revising outputs through their own reasoning, and maintaining academic integrity. Process-based writing instruction and reflective assessment are therefore needed to support responsible, meaningful, and self-regulated use of GenAI. **Originality:** This study integrates patterns of GenAI use, critical thinking, SRL, and perceived benefits and hazards into a single analytical model in the context of Indonesian EFL higher education. Also, the results demonstrate that the quality of self-regulated use is more important than the intensity of use with respect to critical thinking.



Doi: <https://doi.org/10.61255/jupiter.v4i4.1479>

INTRODUCTION

Generative artificial intelligence (GenAI) has quickly entered academic practice in higher education, including language learning and writing instruction. Its expanding presence is particularly noticeable in the context of EFL, where students use AI-based tools to generate ideas, structure arguments, revise drafts, improve grammar and refine language choices. This growth is crucial in writing pedagogy, since EFL writing necessitates not only grammatical accuracy but also coherence, argumentation and rhetorical control. Recent research suggests that GenAI can be harnessed to promote writing development when it is utilised as a scaffold for planning, revision, and feedback, rather than as a replacement for students' own producing processes.

The pedagogical value of GenAI, however, cannot be evaluated solely from cleaner text or faster task completion. A growing body of research suggests that the use of AI in academic work may also encourage cognitive offloading and over-reliance when learners accept generated responses without sufficient evaluation. A systematic review has shown that over-reliance on AI dialogue systems may affect students' cognitive abilities, while empirical evidence also indicates that stronger AI dependence tends to be associated with lower critical thinking. Within EFL education, recent reviews and empirical studies further suggest that GenAI can either support or weaken critical thinking depending on whether learners use it reflectively, critically, and with adequate verification of AI-generated content.

Thus, critical thinking and self-regulated learning (SRL) become important to the conversation of GenAI-assisted writing. Critical thinking is important because students must assess the recommendations offered by AI for relevance, logic, and accuracy before utilising them in their texts. The same importance applies to SRL, as efficient use of GenAI depends on goal-setting, planning, monitoring, rewriting, and managing dependencies throughout the writing process. In AI-assisted academic writing research, the 6-P pedagogy of plan, prompt, preview, produce, peer review, and portfolio tracking has been proposed, explicitly framing AI use as a regulated and reflexive learning process. Other research on EFL writing also indicates that integrating large language models can boost writing performance, motivation, and SRL when students are directed to utilise AI as a support rather than a substitute for thinking and writing.

The study to date has yielded valuable insights, but the data is still fragmentary. The international studies reviewed in this paper have generally examined distinct aspects individually, such as the pedagogical affordances and challenges of ChatGPT in second-language writing, the impact of ChatGPT-generated formative feedback on writing performance, and students' critical interaction with AI-supported learning (Ibrahim & Kirkpatrick, n.d.; Mahapatra, 2024; Nasr et al., 2025).

Indonesia has followed a similar study course. Previous studies have explored teachers' perceptions of ChatGPT use to improve EFL writing instruction (Rahma & Fithriani, 2024), the beliefs of Indonesian EFL students about AI-supported writing and the perceived benefits and risks for English writing development (Riwayatiningih et al., 2026), and the acceptance of pre-service teachers about the utilisation of ChatGPT for writing-related tasks (Syamsi et al., 2025). Other studies have focused on Indonesian students' perceptions of AI-assisted writing and its influence on their writing proficiency, autonomy and creativity (Marhaban et al., 2025), and students' experiences of managing AI assistance for textual authenticity and their strategic use of ChatGPT in the EFL writing process (Werdiningsih, Marzuki, & Rusdin, 2024; Werdiningsih, Marzuki, Indrawati, et al., 2024). Another fascinating area has been the function of AI-supported writing tools in aiding planning, monitoring, the use of feedback and evaluation in self-regulated writing assignments (Amani & Bisriyah, 2025a). In general, these findings reflect an emerging research interest in AI-assisted writing in Indonesia. However, evidence is divided across teachers' and students' opinions, acceptance of technology, perceived benefits and hazards, strategic use, self-regulation, authenticity and specific writing-related outcomes.

Therefore, a more specific gap in the literature is the extremely limited information that jointly maps patterns of GenAI use, critical-thinking approaches, SRL behaviours, and perceived benefits and risks in Indonesian EFL higher education. This difference is important because the issue in education is not whether students are using GenAI, but how they use it, how critically they analyse it, and how well they regulate their dependence on it for writing. Therefore, this study aims to explore the application of GenAI in EFL writing of Indonesian university students from descriptive and correlational perspectives. This study seeks to solve the following six research questions: (1) What patterns of GenAI use do Indonesian students use on EFL writing assignments? (2) To what extent do students apply critical thinking strategies when they use GenAI? (3) The extent to which students engage in SRL behaviours in their interactions with GenAI. (4) How do the students see the benefits and challenges of using GenAI in EFL writing? (5) The association between students' critical thinking in EFL writing and their use of GenAI; and (6) The relationship between students' self-regulated use of GenAI and their critical thinking in EFL writing.

METHOD

The study used a quantitative descriptive correlational survey design with a cross-sectional design. A descriptive approach was employed to examine students' self-reported patterns of generative artificial intelligence (GenAI) use, critical thinking practices, self-regulated learning (SRL) behaviours, and perceived dangers of overreliance on GenAI in English as a Foreign Language (EFL) writing. In the correlational part, the statistical correlations between the frequency of GenAI use and students' reported critical thinking activities and between SRL behaviours in GenAI use and critical thinking practices were examined. All characteristics were evaluated using a single administration of an

online questionnaire, without experimental modification. Therefore, the correlational results were considered as statistical connections rather than predictive, influential or causal (Pérez-Guerrero et al., 2024).

The target population of the study was all diploma and undergraduate students in EFL-related study programs or courses at universities in Indonesia, including those in English Education curricula. The criteria for participation were: being an active student, having done EFL writing tasks, willingness to engage, and having completed the questionnaire. Recruitment of participants took place from January to May 2026 through university class communication groups and students' social media networks. The study employed a non-probability convenience sampling technique, whereby respondents were selected based on their accessibility, willingness, and ability to complete the online questionnaire (Andrade, 2021). Thus, 208 complete responses were retained for analysis, and there were no missing data in the demographic, category, and Likert-scale items. No information is provided about the distribution of the sample by institution, and respondents were not requested to state their universities. The sample obtained cannot be regarded as typical of all Indonesian university students because the probability of selection was unknown, and the respondents were mostly female students in the English Education curriculum and in middle- to later-semester courses. No a priori power calculation was reported. The final sample included all complete and eligible responses received during the data collection period. Limitations regarding the representativeness and external validity of the results were noted, including the lack of distribution at the institutional level and the sampling decision (Andrade, 2021).

The data were collected via a standardised questionnaire constructed by the researcher and tailored to the context of GenAI-assisted EFL writing. The instrument was not a direct modification of an already verified scale. The items were constructed by linking the research questions with the conceptual descriptions of critical evaluation, verification, reflective writing, goal setting, planning, monitoring, revision, self-control, academic integrity, and overreliance on AI-assisted writing. The process of questionnaire construction by the researcher includes systematic definition of constructs, item development, pretesting, and subsequent evaluation of validity and reliability (Ranganathan et al., 2024a; Ranganathan & Caduff, 2023). The questionnaire was separated into three sections. The first segment consisted of seven demographic questions on gender, age, academic level, current semester, course of study, self-rated English competency and primary language of teaching. The second segment comprised four categorical questions regarding past use of GenAI in EFL writing, frequency of GenAI use, the most commonly used GenAI tool and the stage of writing at which GenAI was most frequently used. The last section consisted of 21 closed-ended items measured on a five-point Likert scale, with 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree. The typical five-point Likert-style scale is used to express graded levels of agreement, but individual items and composite scales need to be carefully constructed and interpreted (Jebb et al., 2021). Nine items assessed self-reported engagement in critical thinking practices (e.g., assessing the logic and correctness of AI-produced content; comparing AI suggestions with personal knowledge; challenging ambiguous content; verifying information with credible sources; owning arguments). Nine items were used to measure SRL behaviours, such as goal-setting, planning AI use, monitoring reliance, reviewing AI-produced work, following lecturers' requirements, maintaining academic integrity, and recognising when to continue writing independently. Three items evaluated perceived overreliance, including dependence on AI to construct arguments, diminished independence in reasoning, and unease about writing without AI assistance. Perceived benefits were not included as a separate psychometric construct because the initial benefit indicators were derived from items previously assigned to the critical thinking and SRL subscales. The elements would have been utilised to produce a single composite score. This would have led to construct overlap and potentially misleading indicators of construct validity. As such, any facilitative properties of GenAI represented by those items were known only at the individual-item level and not measured as a separate benefits subscale (Ranganathan et al., 2024).

Initial content relevance was examined by mapping each questionnaire item to the research questions and operational definitions of critical thinking practices, SRL behaviours, and perceived overreliance. This procedure provided preliminary evidence of conceptual alignment but did not constitute formal content or construct validation. Formal questionnaire validation commonly involves procedures such as expert assessment of item relevance, pretesting or pilot testing, and empirical assessment of the proposed measurement structure (Ranganathan et al., 2024). The available research documentation did not include an expert panel review, the number and qualifications of validators, a content validity index, Aiken's V coefficients, pilot testing results, exploratory factor analysis, or confirmatory factor analysis. Claims regarding formal content validity and construct validity were therefore avoided, and the absence of these procedures was recognised as a limitation of the instrument. Internal consistency was examined using Cronbach's alpha based on the 208 complete responses. The reliability coefficients were $\alpha = .494$ for critical thinking practices, $\alpha = .515$ for self-regulated learning behaviours, and $\alpha = .582$ for perceived overreliance. These coefficients indicate limited internal consistency at the subscale level. The relatively low coefficients may be associated with the conceptual

breadth of the behaviours covered by each subscale, the small number of items in the perceived-overreliance subscale, and the concentration of responses in the agree and strongly agree categories, which restricted score variability. The pooled 21-item instrument produced a higher overall coefficient; however, the aggregate coefficient was not interpreted as evidence that the instrument represented a single unidimensional scale because the items measured three conceptually distinct domains. Cronbach's alpha should not be treated as independent evidence of unidimensionality, particularly when items represent different constructs or specific content domains (Doval et al., 2023). The critical thinking, SRL, and perceived-overreliance composite scores were consequently treated as exploratory measures. Descriptive and correlational findings involving these scores were interpreted cautiously, and the limited subscale reliability was considered when evaluating the strength and stability of the results.

The survey was sent electronically through Google Forms and to student groups on social media and in their university classes. Prior to completing the questionnaire, potential respondents were informed about the study's purpose, voluntary participation, and the use of the collected data. Informed consent is necessary in online surveys, which means that explicit information about the study and the voluntary nature of participation must be provided (Mondal et al., 2023). We did not collect any names, student ID numbers, email addresses or any other personal information. Responses were automatically time-stamped and downloaded into one data set for screening and analysis. Participation was voluntary, and anyone could fill out the questionnaire after reading the study literature. Only complete responses were included. We obtained 208 valid responses without missing data for screening. The questionnaire was anonymous; however, this did not rule out self-presentation or social desirability bias, and the likelihood of direct identification was limited. Even anonymous self-reports and Likert-type questions are subject to acquiescence, central tendency, and social desirability biases (Westland, 2022). This adjustment does not affect the number of ethical approvals in the manuscript. We therefore do not claim to have obtained unverifiable ethical institutional clearance in this manuscript. Where formal approval has been granted, the name of the ethical committee that approved the research should be provided, along with the approval number and date. Strategies to collect data in a voluntary, anonymous, and non-identifying manner cannot substitute for ethical approval when the institution explicitly requires it (Mondal et al., 2023).

Data were analysed using the pandas and scipy packages in Python. Demographic characteristics and patterns of GenAI use were summarised using frequencies (f) and percentages (%). Descriptive statistics were used to assess the Likert-scale items, including means (M) and standard deviations (SD). Composite scores were created for the critical thinking approaches, SRL behaviours and perceived overreliance by summing the items assigned to each applicable subscale. Composite Likert scores can be constructed by aggregating related items, but the interpretation of the resulting scores should take into account conceptual and psychometric coherence (Jebb et al., 2021). Mean scores were interpreted as low (1.00-2.33), moderate (2.34-3.67) and high (3.68-5.00). The categories were used only as aids in descriptive terms of interpretation, not as evidence of behaviour objectively witnessed. The high mean scores revealed that the respondents tended to agree with the actions or attitudes described but did not necessarily mean that the behaviours were consistently carried out in real EFL writing activities. The Shapiro-Wilk tests indicated that the composite scores of critical thinking practices and SRL behaviours were statistically significantly different from a normal distribution ($p < .001$). Consequently, the Spearman's rank-order correlation coefficient (ρ) was utilised to examine the association between the frequency of GenAI use and critical thinking skills and between SRL behaviours and critical thinking skills. The Spearman's coefficient is an adequate nonparametric method when numeric data do not meet the normality assumptions of the parametric correlation analysis (Pérez-Guerrero et al., 2024). Statistical significance was assessed using two-tailed tests at $\alpha = .05$. The correlation studies were interpreted in light of the cross-sectional design, self-report measures, limited score variability, and low subscale reliability. The results demonstrated the direction and strength of statistical relationships, but not the temporal order, predictive validity, causal effect, or dominance of one type of GenAI use over another (Pérez-Guerrero et al., 2024).

RESULTS

This section presents the results of descriptive and correlational analyses based on the six research questions. Interpretation of the results is limited to the patterns indicated by the questionnaire data and is not used to infer actual behavior, influence, prediction, or causal relationships.

Respondent Profile

Table 1 presents the demographic characteristics of the 208 student-respondents who participated in the study.

Table 1. Demographic Profile of Respondents (N = 208)

Characteristic / Category	F	%
Gender		
Female	173	83.2
Male	35	16.8
Age		
Below 18	2	1.0
18–20	48	23.1
21–23	147	70.7
24–26	9	4.3
Above 26	2	1.0
Academic Level		
Bachelor	170	81.7
Diploma	38	18.3
Current Semester		
1–2	6	2.9
3–4	51	24.5
5–6	78	37.5
7–8	73	35.1
Program of Study		
English Education	190	91.3
Other	18	8.7
Self-Rated English Proficiency		
Beginner	10	4.8
Intermediate	128	61.5
Upper-Intermediate	51	24.5
Advanced	19	9.1
Primary Language of Instruction		
Indonesian	22	10.6
English	36	17.3
Both	150	72.1

Note. Percentages are based on the total sample (N = 208).

Table 1 shows that 83.2% of the respondents were female, and 70.7% were in the 21–23 age range. Most of the respondents were undergraduate students (81.7%) in their 5th–8th semester (72.6%) and enrolled in the English Language Education Study Program (91.3%). 95% of respondents reported intermediate or higher English proficiency, while 72.1% preferred a mixture of Indonesian and English as their preferred language of instruction. The composition shows that the sample was concentrated among female students, English Language Education students, and mid- to final-semester students. There is an imbalance in the convenience sampling technique and the respondents' characteristics, which limits the generalisability of the results. Therefore, the research findings solely represent the participating respondents and cannot be generalised to all EFL students in Indonesia.

RQ1: Patterns of Generative AI Use

The first research question examined how students used generative AI tools during EFL writing tasks, drawing on four instrument items: prior use of generative AI, frequency of use, the tool used most frequently, and the writing stage at which generative AI was most often used. Results are presented in **Table 2**.

Table 2. Patterns of Generative AI Use Among Student-Respondents (N = 208)

Item / Category	f	%
Prior Use of Generative AI for EFL Writing		
Yes	206	99.0
No	2	1.0
Frequency of Generative AI Use		
Rarely	4	1.9
Sometimes	12	5.8

Item / Category	f	%
Often	157	75.5
Very often	35	16.8
Generative AI Tool Used Most Frequently		
ChatGPT	156	75.0
Gemini	28	13.5
Claude	10	4.8
Other AI tools	10	4.8
Copilot	4	1.9
Writing Stage Where Generative AI Is Most Often Used		
Planning	97	46.6
Drafting	50	24.0
Revising	61	29.3

Note. Frequency response options ranged from rarely to very often; no respondent selected “never.”

Table 2 indicates that generative AI use was both widespread and frequent among the respondents: virtually all students (99.0%) reported having used generative AI for EFL writing, and three-quarters (75.5%) described their use as “often,” with an additional 16.8% using it “very often.” ChatGPT was by far the dominant tool (75.0%), followed distantly by Gemini (13.5%), Claude (4.8%), other tools (4.8%), and Copilot (1.9%). In terms of where AI was integrated into the writing process, students most commonly turned to generative AI during the planning stage (46.6%), followed by revising (29.3%) and drafting (24.0%). This pattern suggests that students tend to use generative AI as an idea-generation and organisational aid early in the writing process, and again as a polishing tool at the revision stage, rather than primarily as a tool for producing first-draft text.

RQ2: Critical Thinking Practices in Generative AI Use

The second research question examined the extent to which students engaged in critical thinking practices when using generative AI for EFL writing. Nine instrument items capturing evaluative, comparative, verification, and reflective behaviours toward AI-generated content were analysed; results are reported in Table 3.

Table 3. Students’ Critical Thinking Practices in Generative AI Use (N = 208)

Item	M	SD	Level
Checking the accuracy of AI-provided facts/examples	4.37	0.78	High
Keeping my own opinions/arguments clear when using AI	4.38	0.66	High
Comparing AI suggestions with my own understanding	4.36	0.82	High
Evaluating the logic of AI-generated suggestions	4.36	0.75	High
Verifying AI information against reliable sources	4.30	0.70	High
Exploring alternative ideas before deciding on my own approach	4.28	0.77	High
Questioning vague or overly general AI ideas	4.28	0.77	High
Reflecting on whether AI supports or weakens my thinking	4.23	0.81	High
Avoiding unreliable AI content	4.16	0.86	High
Overall Critical Thinking Practices	4.30	0.34	High

Note. Scale: 1.00–2.33 = Low, 2.34–3.67 = Moderate, 3.68–5.00 = High. The overall mean is the composite of the nine items. The reliability of the critical thinking practices subscale was $\alpha = .494$, indicating limited internal consistency.

The overall score of critical thinking techniques indicated by the respondents was in the high range (M = 4.30, SD = 0.34). The statement with the greatest mean score was about retaining clarity of personal viewpoints or arguments (M = 4.38); the statement with the lowest mean score was about avoiding unreliable AI content (M = 4.16). All products stayed on the high side. These results indicate that respondents were more likely to agree with comments about reviewing, comparing, confirming and reflecting on GenAI output. This result does not indicate that these behaviours are always observed in real writing practice, as the data were collected through self-assessment. A clustering of scores in the high range and minimal variability in the composite score could indicate a potential ceiling effect and respondents’ inclination to answer in line with academic expectations. The low dependability of the subscales suggests that this composite score should be viewed as an exploratory tool.

RQ3: Self-Regulated Learning Behaviours in Generative AI Use

The third research question identified respondents' level of agreement with nine statements regarding goal setting, planning, monitoring, revision, dependency control, and academic integrity. as presented in Table 4.

Table 4. Students' Self-Regulated Learning Behaviours in Generative AI Use (N = 208)

Item	M	SD	Level
Planning which part of the task needs AI help	4.40	0.70	High
Revising AI text to match my personal writing style	4.39	0.75	High
Understanding lecturers' rules on AI use	4.28	0.64	High
Ensuring submitted writing reflects my own reasoning	4.25	0.71	High
Avoiding direct copying of AI-generated sentences	4.23	0.80	High
Avoiding AI use that violates academic integrity	4.15	0.81	High
Deciding when to stop using AI and write independently	4.12	0.85	High
Using AI with clear goals (not automatic content generation)	4.10	0.56	High
Limiting reliance on AI during writing	4.02	0.93	High
Overall Self-Regulated Learning Behaviours	4.22	0.34	High

Note. Interpretation range: 1.00–2.33 = low; 2.34–3.67 = moderate; 3.68–5.00 = high. The overall score is the average of the nine items. The reliability of the SRL subscale was $\alpha = .515$, indicating limited internal consistency.

The overall score of SRL behaviour was in the high category, $M=4.22$, $SD=0.34$. The highest mean score, $M = 4.40$, was obtained for planning portions of tasks that require AI assistance, followed by rewriting AI content to suit one's particular writing style ($M = 4.39$). Limiting dependency throughout the writing process received the lowest mean score ($M = 4.02$) but was still in the high category. Scores suggest that the respondents generally agreed with the assertions on planning, editing, regulating use, and academic integrity. It is not clear from the data how frequently these behaviours were used in actual writing projects. The concentration of responses in the high category and the relatively small SD of the composite suggest little variation in responses. Given the low internal consistency of the subscales, the SRL scores should be regarded with caution as self-reported views and as exploratory measures.

RQ4: Perceived Benefits and Challenges of Generative AI Use

The fourth research question addressed respondents' perceptions of the facilitative and risk aspects of GenAI use. Three items related to facilitative use were presented descriptively at the item level. These items were not treated as stand-alone benefit constructs or combined into a composite score because they were derived from the critical thinking and SRL subscales.

Table 5. Descriptive Responses to GenAI Facilitative Use Items (N = 208)

Item	M	SD	Level
AI output can be adapted to my personal writing style	4.39	0.75	High
AI support helps me keep my arguments clear	4.38	0.66	High
AI helps me explore a wider range of ideas	4.28	0.77	High

Note. The three items are derived from the critical thinking and SRL subscales. They are presented only as item-level descriptions and are not used to form a construct or composite benefit score.

Table 5 shows that respondents indicated relatively high levels of agreement with all three statements. These results indicate that adjusting writing style, maintaining clarity of argument, and exploring alternative ideas emerged as aspects of GenAI use that respondents positively assessed. This interpretation is limited to each item and does not constitute evidence that the study measured the benefit constructs separately.

Table 6. Students' Perceived Challenges/Risks of Generative AI Use (N = 208)

Item	M	SD	Level
Relying on AI to construct arguments I cannot build myself	4.15	0.80	High
Depending less on my own reasoning skills when using AI	3.88	1.12	High
Feeling uncomfortable writing without AI assistance	3.68	1.03	High
Overall Perceived Challenge/Over-Reliance	3.90	0.73	High

Note. Interpretation range: 1.00–2.33 = low, 2.34–3.67 = moderate, 3.68–5.00 = high; higher scores reflect a greater perceived danger of reliance. The perceived over-reliance subscale had a reliability of $\alpha = .582$, indicating low internal consistency.

Overall, the perceived over-reliance score was in the high group ($M = 3.90$, $SD = .73$). Concerns about using AI to develop arguments that are not possible on their own received the highest mean score, $M = 4.15$. Concerns about reduced use of personal reasoning, $M = 3.88$. Concerns about uncomfortable writing without AI support were $M = 3.68$. These findings confirm respondents' concerns about relying on GenAI. There is no evidence in the numbers to suggest that GenAI has reduced cognitive ability or autonomous writing. Because the number of questions is small (3) and the internal consistency of the subscale is low, the scores should be interpreted in an explorative way.

RQ5: Relationship Between Generative AI Use and Critical Thinking Practices

The fifth research question examined the relationship between students' frequency of generative AI use (rarely = 1, sometimes = 2, often = 3, very often = 4) and their composite critical thinking practices score. A Spearman rank-order correlation was computed.

Table 7. Spearman Correlation Between Frequency of Generative AI Use and Critical Thinking Practices ($N = 208$)

Variable	1	2
Frequency of Generative AI Use	-	.033
Critical Thinking Practices		-

Note. $\rho = .033$, $p = .638$, $N = 208$. The correlation is positive but negligible and not statistically significant at $\alpha = .05$.

As shown in Table 7, There was no significant association between the frequency of GenAI use and the respondents' reported critical-thinking practices ($\rho = .033$, $p = .638$). The mean for critical thinking activities was $M = 4.25$ (rarely), $M = 4.44$ (occasionally), $M = 4.29$ (often), and $M = 4.31$ (very often). These descriptive differences were neither linear nor statistically significant. These results indicate that the usage of GenAI was not associated with differences in critical thinking scores in the research sample. This finding does not mean that intensity of use has no repercussions for critical thinking in other circumstances. Again, the averages for the rarely and occasionally categories should not be overinterpreted, as they, too, had relatively small numbers of respondents.

RQ6: Relationship Between Self-Regulated Use of Generative AI and Critical Thinking

The sixth research question examined the relationship between students' self-regulated learning (SRL) behaviours in generative AI use and their critical thinking practices. A Spearman rank-order correlation was computed between the composite SRL and critical thinking scores.

Table 8. Spearman Correlation Between Self-Regulated Generative AI Use and Critical Thinking Practices ($N = 208$)

Variable	1	2
Self-Regulated Learning Behaviours	-	.449**
Critical Thinking Practices		-

Note. ** $p < .001$. $\rho = .449$, $N = 208$, indicating a moderate, positive, and statistically significant relationship.

As shown in Table 8, There was a moderately strong, statistically significant positive association between respondent-reported SRL behaviours and critical thinking practices, $\rho = .449$, $p < .001$. Respondents with higher SRL ratings had higher scores on critical thinking techniques. This is a correlational association and does not suggest that SRL promotes advances in critical thinking. Both ratings were obtained from the same respondents, using the same questionnaire, at the same time. Hence, the effect of common procedure bias cannot be ruled out. Both subscales demonstrated low internal consistency. The association coefficients should be viewed as exploratory and need to be confirmed using more robust instruments, behavioural observations, written products, prompt historical data, or longitudinal designs.

DISCUSSION

The high level of GenAI use suggests that this technology has become a common feature of respondents' EFL writing experiences. The most used tool was ChatGPT, and GenAI was most involved in the planning stage. This coincides with research showing students use AI to plan their writing, generate ideas, improve language, and get feedback. The findings from this study merely show respondents' preferred phases of use; the specific function of GenAI in the writing process cannot be identified without investigating the history of prompts, interactions, and draft evolution (Amani & Bisriyah, 2025; Polakova & Ivenz, 2024).

Respondents who scored highly in critical thinking were more likely to agree with statements about reviewing facts, logically evaluating, comparing ideas, and verifying sources. These results are still self-reported practices and not evidence that these behaviours are consistently used in actual writing projects. The influence of GenAI on critical

thinking depends on how users interact with it, appraise its output, and remain cognitively engaged during the activity (Nasr et al., 2025).

High SRL ratings correlate with planning, monitoring, revising and limiting GenAI use among responders. This trend is consistent with the 6-P pedagogy, which characterises planning, rapid development, review, production, peer review, and portfolio reflection as the stages of purposeful AI use. Research on EFL writing instruction further indicates that language model support should be linked to student learning goals, monitoring, and evaluation (Kong et al., 2024).

As the three items utilised are drawn from the subscales of critical thinking and SRL, rather than from a benefits construct built separately, the interpretation of benefits in GenAI must be limited. These are only items that reflect respondents' approval of using AI to alter writing style, maintain argument clarity, and explore ideas. High dependency risk scores also reflected conflicted attitudes, with respondents acknowledging the facilitative role of GenAI while expressing concern about reduced autonomous reasoning. These concerns correlate with evidence that excessive dependence on AI dialogue systems can affect students' analytical processes, decision-making, and critical thinking (Zhai et al., 2024).

The frequency of GenAI use was not substantially associated with critical-thinking approaches ($\rho = .033$, $p = .638$). That finding does not mean that the frequency of consumption has no cognitive consequences. Nearly all respondents had used GenAI, and most used it often or very frequently; thus, there was little variation in frequency. SRL was moderately favourably linked with critical thinking strategies, $\rho = .449$, $p < .001$. This correlation implies that higher SRL scores are associated with higher critical thinking scores; the data do not imply that SRL predicts, causes, or is of greater causal importance than the intensity of GenAI use.

Caution in interpretation is also warranted, as both constructs were measured using the same respondents, methods, response formats, and collection durations. This context is prone to common-method bias, acquiescence bias, and responses that align with academic expectations (Podsakoff et al., 2026). The reliability coefficients for critical thinking practices, self-regulated learning behaviours, and perceived over-reliance were $\alpha = .494$, $\alpha = .515$, and $\alpha = .582$, respectively. The statistics indicate limited internal consistency; therefore, the composite score is considered exploratory. Further, the distribution of responses in the high category implies a probable ceiling effect and minimal score variability. Reliability evaluations should consider the construct structure, number of items, and response distribution rather than depending on rigid numerical boundaries (Hussey et al., 2025).

Characteristics of the sample also limit interpretation. The sample was mostly female, consisted of students in English Education in the middle-to-final semesters, and had intermediate-level or above English proficiency. Although language competency, semester, writing experience, experience with AI, and lecturer policies may be associated with students' perception of GenAI output, these variables were not considered in the analysis. Hence, the findings of this study cannot be generalised to all the EFL students in Indonesia.

The educational implications of this study suggest the necessity to integrate GenAI into the writing process transparently and with purpose. Assignments could require students to state the intent of the request, verify the AI output against reputable sources, record modifications to drafts, and justify whether to accept or reject AI ideas. Such assessments can include the quality of the final writing, the revision history, the verification of sources, and contemplation on who owns the argument. These recommendations were not evaluated in interventions in this study. Further study is needed using more reliable instruments, more diverse samples, and behavioural data (such as prompt histories, writing drafts, observations, interviews, and writing assessments).

CONCLUSION

The study included 208 participants with experience in EFL writing tasks who used generative artificial intelligence (GenAI). 99% of respondents said they had used GenAI; ChatGPT was the most-used tool, and the planning stage was the writing stage in which GenAI assistance was most often involved. Respondents indicated high levels of critical thinking practices ($M = 4.30$; $SD = 0.34$), high levels of self-regulated learning (SRL) behaviours ($M = 4.22$; $SD = 0.34$), and high levels of perceived danger of overreliance ($M = 3.90$; $SD = 0.73$). The results are based on self-reported beliefs and practices rather than observable proof of the actual behaviour of all EFL students in Indonesia.

GenAI use frequency was not significantly associated with critical thinking skills in this sample of study participants ($\rho = .033$; $p = .638$). SRL behaviour in GenAI use was positively, modestly, and significantly associated with critical thinking techniques ($\rho = .449$; $p < .001$). This is a correlational link and does not establish that SRL causes, determines or is causally more relevant than frequency of GenAI use. This data merely suggests that respondents who report higher SRL scores are also likely to report higher scores on critical-thinking activities.

This study contributes to the scientific community by mapping, for the first time, the usage patterns of GenAI, the self-reported practices of critical thinking, the behaviours of SRL, the dangers of over-reliance, and the statistical differences between the frequency of use and the self-regulated use of GenAI in the context of EFL writing. Items related to the facilitative function of GenAI were not examined as independent sub-scales, as they were not defined and validated as separate benefit constructs and were only analysed at the item level.

The reliability coefficients were $\alpha = .494$ for critical thinking practices, $\alpha = .515$ for self-regulated learning behaviours and $\alpha = .582$ for perceived over-reliance. Results show limited internal consistency; therefore, subscale scores should be viewed as exploratory measures. Furthermore, the findings may be limited by the generalisability of the convenience sampling, the predominantly female students studying English Education, and the cross-sectional design and self-report questionnaires at one measurement point.

The research offers pedagogical implications for process-based writing education, including planning for AI usage, formulating guided prompts, confirming sources, assessing outputs, documenting modifications, assuring transparency in AI use, and reflecting on argument ownership. In this study, these instructions were not viewed as interventions. Future research needs to improve sample heterogeneity, construct validity, and instrument reliability, and to add behavioural data such as prompt history, draft development, observations of the writing process, interviews, and measures of writing quality.

REFERENCE

- Amani, N., & Bisriyah, M. (2025). University students' perceptions of AI-assisted writing tools in supporting self-regulated writing practices. *IJELTAL: Indonesian Journal of English Language Teaching and Applied Linguistics*, 10(1), 91–107.
- Andrade, C. (2021). The inconvenient truth about convenience and purposive samples. *Indian Journal of Psychological Medicine*, 43(1), 86–88. <https://doi.org/10.1177/0253717620977000>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- Creswell, J. W. (2002). Educational research: Planning, conducting, and evaluating quantitative.
- Dahlan, S. P. (2026). Generative artificial intelligence and critical digital literacy in EFL academic writing: An integrative review and pedagogical framework. *Journal of Language and Literature Inquiry*, 5(2), 182–203. <https://journal.privietlab.org/index.php/JLLI/article/view/1850>
- Deng, L. (2026). Roles of generative artificial intelligence (GenAI) in English as a foreign language instruction: A systematic literature review. *SAGE Open*. <https://doi.org/10.1177/21582440261418315>
- Diasti, K. S. (2026). EFL students' self-regulation in the writing process using ChatGPT. *Journal of English Teaching, Literature, and Applied Linguistics*, 10(1), 137–142. <https://doi.org/10.30587/jetlal.v10i1.10837>
- Doval, E., Viladrich, C., & Angulo-Brunet, A. (2023). Coefficient alpha: The resistance of a classic. *Psicothema*, 35(1), 5–20. <https://doi.org/10.7334/psicothema2022.321>
- Dzahabiyah, L. (2025). Exploring the role of ChatGPT in enhancing EFL students' writing skills: A qualitative study at Universitas Negeri Semarang. *Journal of English Language Education*, 10(6). <https://jele.or.id/index.php/jele/article/view/1560>
- Fadillah, R. (2025). EFL Students' Perceptions of AI-Powered Feedback in Academic Writing: A Qualitative NVivo Study. *ELITE JOURNAL*, 7(2), 125–136.
- Fujiono, F., Hartono, R., & Wahyuni, S. (2025). Students' perceptions and experiences of ChatGPT-assisted writing in EFL contexts: A literature review. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 13(1). <https://doi.org/10.24256/ideas.v13i1.6943>
- Gagnier, J. J., Lai, J., Mokkink, L. B., & Terwee, C. B. (2021). COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Quality of life research*, 30(8), 2197–2218.
- Gonsalves, C. (2024). Generative AI's impact on critical thinking: Revisiting Bloom's taxonomy. *Journal of Marketing Education*, 48(1), 4–19. <https://doi.org/10.1177/02734753241305980>
- Gunawan, F., Rahmah, S. M., Rusdyansyah, R., & Tarihoran, N. (2026). Students' perceptions and attitudes toward the use of ChatGPT in English language writing. *Jurnal Jendela Pendidikan*, 6(2), 379–394. <https://doi.org/10.57008/jjp.v6i02.2604>
- Hadi, A., Aribowo, K., & Prasetyo, D. E. (2026). Students' self-regulation in learning academic writing with AI mediation tools. *Journal of English Language Education*, 11(3), 355–365. <https://doi.org/10.31004/jele.v11i3.2437>

-
- Hartono, R. (2025). Students' perceptions and experiences of ChatGPT-assisted writing in EFL contexts: A literature review. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*. <https://ejournal.iainpalopo.ac.id/index.php/ideas/article/view/6943>
- Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, 5(1), 37–45. <https://doi.org/10.61414/jeti.v5i1.103>
- Hussey, I., Alsalti, T., Bosco, F., Elson, M., & Arslan, R. (2025). An aberrant abundance of Cronbach's alpha values at .70. *Advances in Methods and Practices in Psychological Science*, 8(1). <https://doi.org/10.1177/25152459241287123>
- Ibrahim, K., & Kirkpatrick, R. (2024). Potentials and Implications of ChatGPT for ESL Writing Instruction. *The International Review of Research in Open and Distributed Learning*, 25(3), 394–409. <https://doi.org/10.19173/irrodl.v25i3.7820>
- Janah, R., Tusino, T., & Dewi, P. (2026). Exploring the role of ChatGPT in developing writing skills and critical thinking among vocational high school students. *Al-Ishlah: Jurnal Pendidikan*, 18(1). <https://doi.org/10.35445/alishlah.v18i1.8580>
- Jebb, A. T., Ng, V., & Tay, L. (2021). A review of key Likert scale development advances: 1995–2019. *Frontiers in Psychology*, 12, Article 637547. <https://doi.org/10.3389/fpsyg.2021.637547>
- Kaltsonoudi, K., Tsigilis, N., & Karteroliotis, K. (2022). Common method bias in research designs using self-report tools: Literature overview and recommended remedies. *Psychology: The Journal of the Hellenic Psychological Society*, 27(1), 210–231. <https://doi.org/10.12681/psyhps.30843>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., . . . Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kong, S.-C., Lee, J. C.-K., & Tsang, O. (2024). A pedagogical design for self-regulated learning in academic writing using text-based generative artificial intelligence tools: 6-P pedagogy of plan, prompt, preview, produce, peer-review, portfolio-tracking. *Research and Practice in Technology Enhanced Learning*, 19, Article 030. <https://doi.org/10.58459/rptel.2024.19030>
- Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, 8(1), Article 33267. <https://doi.org/10.1525/collabra.33267>
- Lambert, L. S., & Newman, D. A. (2023). Construct development and validation in three practical steps: Recommendations for reviewers, editors, and authors. *Organizational Research Methods*, 26(4), 574–607.
- Larson, B. Z., Moser, C., Caza, A., Muehlfeld, K., & Colombo, L. A. (2024). Critical thinking in the age of generative AI. *Academy of Management Learning & Education*. <https://doi.org/10.5465/amle.2024.0338>
- Lee, Y.-J., Davis, R. O., & Choi, J.-I. (2025). Integrating generative AI into EFL writing: University students' strategies and perceptions. *Online Journal of Communication and Media Technologies*, 15(4), e202541. <https://doi.org/10.30935/ojcm/17545>
- Liu, Z.-M., Hwang, G.-J., Chen, C.-Q., Chen, X.-D., & Ye, X.-D. (2024). Integrating large language models into EFL writing instruction: Effects on performance, self-regulated learning strategies, and motivation. *Computer Assisted Language Learning*, 1–25. <https://doi.org/10.1080/09588221.2024.2389923>
- Mahapatra, S. Impact of ChatGPT on ESL students' academic writing skills: a mixed methods intervention study. *Smart Learn. Environ.* 11, 9 (2024). <https://doi.org/10.1186/s40561-024-00295-9>
- Marhaban, S., Sulisty, T., & Widiastuti, O. (2025). Artificial intelligence in writing: Students' writing competencies and voices. *JEELS: Journal of English Education and Linguistics Studies*, 12(1), 403–425. <https://doi.org/10.30762/jeels.v12i1.3927>
- Maslan, M., & Ubaidah, N. (2026). Artificial intelligence in education and its conditional impact on critical thinking: A systematic literature review. *Jurnal Didactical Mathematics*, 8(1), 188–204. <https://doi.org/10.31949/dm.v8i1.18001>
- Mekheimer, M. (2025). Generative AI-assisted feedback and EFL writing: A study on proficiency, revision frequency and writing quality. *Discover Education*, 4, 170. <https://doi.org/10.1007/s44217-025-00602-7>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
-

-
- Mohammed, S. J., & Khalid, M. W. (2025). Under the world of AI-generated feedback on writing: Mirroring motivation, foreign language peace of mind, trait emotional intelligence, and writing development. *Language Testing in Asia*, 15, 7. <https://doi.org/10.1186/s40468-025-00343-2>
- Mondal, H., Mondal, S., & Soni, S. (2023). A cross-sectional audit of informed consent of online survey: Characteristics and adherence to prevalent guidelines. *Perspectives in Clinical Research*, 14(3), 123–129. https://doi.org/10.4103/picr.picr_175_22
- Moon, H., Chung, Y., & Randolph, A. W. (2025). Teaching and learning languages with ChatGPT: Challenges and opportunities in multilingual classrooms in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics*, 10(1), 207–223. <https://doi.org/10.21093/ijeltal.v10i1.1991>
- Mubarak, M. K. (2026). EFL learners' preferences and needs for AI writing tools: A systematic literature review. *Proceedings of the International Conference on English Education, Linguistics, and Literature*. <https://ejournal.unuja.ac.id/index.php/iceet/article/view/14564>
- Naqvi, W. M., Ganjoo, R., Rowe, M., Pashine, A. A., & Mishra, G. V. (2025). Critical thinking in the age of generative AI: Implications for health sciences education. *Frontiers in Artificial Intelligence*, 8. <https://doi.org/10.3389/frai.2025.1571527>
- Nasr, N. R. (2025). Exploring the impact of generative AI ChatGPT on critical thinking in higher education: Passive AI-directed use or active AI-facilitated engagement? *Education Sciences*, 15(9), 1198. <https://doi.org/10.3390/educsci15091198>
- Ngo, C.-L., Nguyen, T. T., & Nguyen, K. N. H. (2025). Critical thinking in the age of generative AI: Effects of a short-term experiential learning intervention on EFL learners. *International Journal of TESOL Studies*. <https://doi.org/10.58304/ijts.250522>
- OECD. (2026). *OECD digital education outlook 2026: Exploring effective uses of generative AI in education*. OECD Publishing. <https://doi.org/10.1787/062a7394-en>
- Pérez-Guerrero, E. E., Guillén-Medina, M. R., Márquez-Sandoval, F., Vera-Cruz, J. M., Gallegos-Arreola, M. P., Rico-Méndez, M. A., Aguilar-Velázquez, J. A., & Gutiérrez-Hurtado, I. A. (2024). Methodological and statistical considerations for cross-sectional, case-control, and cohort studies. *Journal of Clinical Medicine*, 13(14), Article 4005. <https://doi.org/10.3390/jcm13144005>
- Podsakoff, P. M., Podsakoff, N. P., Williams, L. J., Huang, C., & Yang, J. (2024). Common method bias: It's bad, it's complex, it's widespread, and it's not easy to fix. *Annual Review of Organizational Psychology and Organizational Behavior*, 11, 17–61. <https://doi.org/10.1146/annurev-orgpsych-110721-040030>
- Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2410101>
- Putri, I. C., Imaniah, I., & Nargis, N. (2025). Using Chatgpt To Support Efl Writing: Student Insights And Experiences. *Eltin Journal: Journal of English Language Teaching in Indonesia*, 13(2), 423-432.
- Qawqzeh, Y. (2024). Exploring the influence of student interaction with ChatGPT on critical thinking, problem solving, and creativity. *International Journal of Information and Education Technology*, 14(4), 596–601. <https://doi.org/10.18178/ijiet.2024.14.4.2082>
- Rahma, A., & Fithriani, R. (2024). The potential impact of using Chat GPT on EFL students' writing: EFL teachers' perspective. *Indonesian EFL Journal*, 10(1), 11–20. <https://doi.org/10.25134/ieflj.v10i1.9222>
- Ramanda, D., Kardena, A., Eliza, & Syafitri, W. (2026). The Impact Of Artificial Intelligence On The Writing Practices Of Sixth-Semester English Students At Uin Sjech M. Djamil Djambek Bukittinggi. *Elp (Journal of English Language Pedagogy)*, 11(1), 101–112. <https://doi.org/10.36665/elp.v11i1.1172>
- Ranganathan, P., & Caduff, C. (2023). Designing and validating a research questionnaire—Part 1. *Perspectives in Clinical Research*, 14(3), 152–155. https://doi.org/10.4103/picr.picr_140_23
- Ranganathan, P., Caduff, C., & Frampton, C. M. A. (2024). Designing and validating a research questionnaire—Part 2. *Perspectives in Clinical Research*, 15(1), 42–45. https://doi.org/10.4103/picr.picr_318_23
- Rawas, S. (2024). ChatGPT: Empowering lifelong learning in the digital age of higher education. *Education and Information Technologies*, 29(6), 6895–6908. <https://doi.org/10.1007/s10639-023-12114-8>
- Riwayatiningsih, R., Vivekanantharasa, R., & Prastikawati, E. F. (2026). EFL students' beliefs about AI-assisted writing and its perceived impact on their English writing development. *JEELS: Journal of English Education and Linguistics Studies*, 13(1), 363–388. <https://doi.org/10.30762/jeels.v13i1.7343>
- Roy, B. K., & Jony, M. S. (2026). Generative AI and critical thinking in higher education: Students' narratives of cognitive offloading. *The Journal of Interdisciplinary Studies in Education*, 15(3), 263–282. <https://doi.org/10.32674/jeys9q25>
-

- Salwa, A. (2024). Exploring students' perception of EFL on the use of ChatGPT in writing classes. *Berumpun: International Journal of Social, Politics, and Humanities*, 7(1). <https://berumpun.ubb.ac.id/index.php/BRP/article/view/186>
- Settiawan, D. (2025). Leveraging ChatGPT in EFL writing assessment. *Journal of English and Arabic Language Teaching*, 16(1), 80–92. <https://ejournal.uin-suska.ac.id/index.php/jealt/article/view/37850>
- Shabir, A. (2025). Generative AI writing tools and academic writing in higher education: A systematic review. *INTEC: Information Technology Education Journal*, 4(3), 543–557. <https://doi.org/10.59562/intec.v4i3.10494>
- Shen, B., & Bai, B. (2024). Enhancing Chinese university students' writing performance and self-regulated learning (SRL) writing strategy use through a strategy-based intervention. *System*, 122, 103249.
- Syamsi, K., Nugroho, A., Cahyani, I., & Retnaningsih, W. (2025). EFL pre-service teachers' acceptance of ChatGPT for writing: A sequential explanatory mixed-method study. *Indonesian Journal of Applied Linguistics*, 15(1).
- Tahir, M., Jahrir, A. S., Asrifan, A., & Ariatna, A. (2025). Examining the role of AI-powered writing assistants in enhancing critical thinking in EFL academic writing. *JOLLT Journal of Languages and Language Teaching*, 13(4), 1817–1831. <https://doi.org/10.33394/joltt.v13i4.15077>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10, 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Tusino, T., Dewi, P., Hati, M. L. P., & Iglesia, M. J. C. (2025). ChatGPT in academic writing classes: Exploring EFL students' perspectives. *JOLLT Journal of Languages and Language Teaching*, 13(4), 2162–2172. <https://doi.org/10.33394/joltt.v13i4.15293>
- UNESCO International Bureau of Education. (2024). *Critical thinking and generative artificial intelligence*. UNESCO. <https://www.ibe.unesco.org/en/articles/critical-thinking-and-generative-artificial-intelligence>
- Vangsness, L., Voss, N. M., Maddox, N., Devereaux, V., & Martin, E. (2022). Self-report measures of procrastination exhibit inconsistent concurrent validity, predictive validity, and psychometric properties. *Frontiers in Psychology*, 13, 784471. <https://doi.org/10.3389/fpsyg.2022.784471>
- Werdiningsih, I., Marzuki, & Rusdin, D. (2024). Balancing AI and authenticity: EFL students' experiences with ChatGPT in academic writing. *Cogent Arts & Humanities*, 11(1), Article 2392388. <https://doi.org/10.1080/23311983.2024.2392388>
- Werdiningsih, I., Marzuki, Indrawati, I., Rusdin, D., Ivone, F. M., Basthomi, Y., & Zulfahreza. (2024). Revolutionizing EFL writing: unveiling the strategic use of ChatGPT by Indonesian master's students. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2399431>
- Westland, J. C. (2022). Information loss and bias in Likert survey responses. *PLOS ONE*, 17(7), Article e0271949. <https://doi.org/10.1371/journal.pone.0271949>
- Xu, J. (2026). How generative AI enhances self-regulated learning in EFL writing. *Frontiers in Psychology*, 17, Article 1808183. <https://doi.org/10.3389/fpsyg.2026.1808183>
- Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments*, 11, Article 28. <https://doi.org/10.1186/s40561-024-00316-7>
- Zhou, M. (2025). Artificial intelligence in English writing feedback: A systematic review, 2020–2025. *Vietnam Journal of Education*. <https://vjes.vnies.edu.vn/en/artificial-intelligence-english-writing-feedback-systematic-review-2020-2025>

ACKNOWLEDGEMENT

The author would like to express appreciation and gratitude to all student respondents who participated in this study, as well as to the students who volunteered to assist with the distribution of the survey instruments and data collection. The author also extends appreciation to the academic staff who supported this study. Thanks are also extended to the faculty members and lecture collaborating for their academic support, substantive input, and scientific collaboration, which contributed to the drafting and refinement of this manuscript.

AUTHOR CONTRIBUTION STATEMENT

Siti Nur Diana contributed to the formulation of the research concept, the development of the research design, data analysis, and the writing of the initial draft of the manuscript and contributed to the development of the research instruments, the data collection process, and the critical review of the manuscript's content. Syaiful Islam provided

academic input, helped interpret the research results, and refined the manuscript's substance. All authors have read, reviewed, and approved the final version of the manuscript.

AI DISCLOSURE STATEMENT

The authors used ChatGPT (OpenAI) during the preparation of this work for limited language editing to improve clarity of expression and refine academic style. After using this tool, the authors thoroughly reviewed and edited the text as needed and take full responsibility for the publication's content. The AI tool had no role in the study design, data collection, data analysis, interpretation of the findings, decision to publish, or the development of the scientific content of the manuscript. All scientific content, academic judgment, and final responsibility for this paper remain exclusively with the authors.

***Siti Nur Diana (Corresponding Author)**

English Education Department Faculty of Social and Humaniora,
Universitas Nurul Jadid,
Karanganyar, Paiton, Probolinggo, East Java, 67291, Indonesia
Email: dianadns053@gmail.com

Syaiful Islam

English Education Department Faculty of Social and Humaniora,
Universitas Nurul Jadid,
Karanganyar, Paiton, Probolinggo, 67291, East Java, 60237, Indonesia
Email: syaifulislam@unuja.ac.id
