

AI-Driven Writing Assistance in a Flipped Classroom Model: A Pedagogical Approach to Enhancing Academic Writing and Critical Thinking in EFL Learners

Pirman Ginting*

Universitas Muhammadiyah Sumatera Utara, Indonesia

Yenni Hasnah

Universitas Muhammadiyah Sumatera Utara, Indonesia

Adib Jasni Kharisma

Universitas Muhammadiyah Sumatera Utara, Indonesia

Dodi Mulyadi

Universitas Muhammadiyah Semarang, Indonesia

Correspondence

Email: pirmanginting@umsu.ac.id

Abstract

This study explores the AI-based flipped model to enhance academic writing proficiency and critical thinking among EFL learners. Employing a mixed method through a one-group pretest-posttest design, the study combined quantitative assessment using the IELTS Academic Writing Task 2 rubric and Facione's Critical Thinking framework with qualitative data drawn from open-ended post-intervention questionnaires. Descriptive statistics showed that the students' writing skills advanced from a pre-test mean of 4.75 to a post-test mean of 6.17, which stand for an average increase of 1.42 points (29.8%). Meanwhile, critical thinking scores increased from a mean of 42.87 to 68.14 with an average increase of 25.27 points (58.9%). The Wilcoxon signed-rank test further confirmed substantial rises for both outcomes (writing: $Z = -5.163$, $p < .001$; critical thinking: $Z = -5.107$), with the effect sizes (writing: $r = .873$, $rrb = 1.000$; critical thinking: $r = .851$, $rrb = .976$). Qualitative findings corroborated this trend, denoting significant improvements in writing organization, argument construction, evidence-based argumentation, and source evaluation. However, challenges such as over-reliance on AI output and difficulties related to adapting to self-paced learning were also observed. Implications for pedagogical design were highlighted, implying that AI-integrated flipped learning supported the teaching of writing and critical thinking skills.

ARTICLE HISTORY

Received: 08 June 2026

Revised: 24 September 2026

Accepted: 29 September 2026

KEYWORDS

Academic Writing, Critical Thinking, AI-Driven Writing Assistance, Flipped Classroom Model

How to cite this article (APA 7th Edition):

Ginting, P., Hasnah, Y., Kharisma, A. J., & Mulyadi, D. (2026). AI-driven writing assistance in a flipped classroom model: A pedagogical approach to enhancing academic writing and critical thinking in EFL learners. *Language Teaching Research Quarterly*, 56, 132–154. <https://doi.org/10.32038/ltrq.2026.56.07>

Introduction

Proficiency in writing, coupled with strong critical skills, enables students to engage in meaningful discussions and tackle various subjects across academic and real-world settings. Valuing these competencies, writing prowess and critical thinking (CT) constitutes a crucial foundation for fostering personal and educational development (Kafri, 2022). In the EFL context, in particular, although these two competencies are separately evaluated, they are interconnected and revealed as two pivotal elements that contribute to language learning and communicative competence (Stephen & Muriel, 2013). In written communication, for instance, critical thinking involves analyzing varied aspects of reasoning, such as purposes, issues at hand, assumptions, perspectives, information, and implications (Celuch & Slama, 2000), which requires diverse modes of thought through careful scrutiny (Tanaka & Gilliland, 2016) and shifts the focus from mere content to an understanding and appreciation of the cognitive processes in examining that content (Celuch & Slama, 2000). Building on this perspective, developing students' critical thinking competencies allow them to enhance their writing skills, encompassing stylistic, textual, and grammatical proficiency, while concurrently augmenting their potential for critical self-reflection on their own learning processes (Schicker, 2018). Analogously, looking through the prism of writing dimensions, writing endeavors allow learners to dive into specific topics through which the process of writing constructs insights that elevate higher-order thinking (Güven et al., 2020; Yildirim et al., 2011). Notably, mastery of writing augments critical thinking faculties; conversely, critical thinking, as a higher-order cognitive process, is critical in developing writing skills (Ebadi & Rahimi, 2018), and together, these cognitive functions and activities are evident in a writer's rhetorical output (Ebadi & Rahimi, 2018; Yin et al., 2024).

Given the significance of writing and critical thinking for students' academic development, notably for EFL learners (Yin et al., 2023), a plethora of endeavors, such as reflective writing, have been employed to cultivate these skills (Carter et al., 2017); however, empirical evidence regarding their efficacies to enhance both skills remains limited (Naber & Wyatt, 2014). Studies report that a substantial cohort of students depart without having thoroughly acquired these competencies (Piergiovanni, 2014; Bandyopadhyay & Szostek, 2018), as evidenced by difficulties in elucidating the meanings of words, comparing and contrasting ideas, assessing assumptions, and correlating with relevant facts in their writing (Umar & Rathakrishnan, 2012). Therefore, addressing this emerging issue requires appropriate approaches, and endorsing AI-driven writing tools, particularly when integrated with a flipped classroom model offers a transformative strategy for enhancing EFL learners' academic writing and critical thinking skills.

The advent of generative AI tools in the writing process has been a breakthrough for enhancing students' writing skills and overall writing quality. In line with the growing potential of AI-assisted writing, a range of studies have reported considerable progress in writing performance facilitated by such tools. Inquiries through mixed-method designs reveal remarkable support for the use of AI writing assistants in academic writing tasks, including generating research content, summarizing literature, refining written work and proofreading (Al-Sofi, 2024; Xu & Jumaat, 2024). Echoing these findings, Kim et al. (2024) and Jia and Lu (2026) emphasized the value of automatic word processing tools in improving academic arguments by strengthening textual coherence, logical flow, and clarity. In addition to supporting basic writing mechanics, AI-driven writing tools are increasingly recognized for their role in providing personalized and peer-like feedback (Seo, 2024), which accounts for edits and corrections that potentially exacerbate the accuracy (Crusan & Mohebbi, 2026; Wang et al., 2024; Zhai & Ma, 2022), thus, it enables students to focus on the content of their writing rather than on mechanics, saving time and leading to greater writing efficiency (Golan et al., 2023). More importantly, as a digital peer, AI feedback can preserve learners' autonomy, provided that they engage with it meaningfully (Dong, 2026).

Taking into account that academic writing is a cognitively laborious and timely feedback endeavor from instructors, infusing AI-driven personalized learning tools into flexible instructional models, such as the flipped classroom, can support both autonomous and collaborative learning processes. Such a hybrid feedback model that blends the responsiveness of generative AI with the nuance of human insight by following through comprehensive, balanced feedback across multiple levels, including process-level, self-regulation, and motivational support, promotes writing proficiency and enhances various writing skills (Zhang et al., 2025). Complementing this technological support, the flipped learning model offers personalized and flexible learning by providing instructional content, typically in video format, prior to class, allowing students to engage with the material at their own pace and adapt it to their preferred learning styles (Bergmann & Sams, 2012). With the transmission of content shifting to out-of-class activities (e.g., via pre-recorded videos or readings), class time is repurposed for active problem-solving, either individually or collaboratively, that enables students to engage in deeper cognitive processing supported by the teacher's scaffolded guidance and the social dynamics of peer learning (Abeysekera & Dawson, 2014; Bezzazi, 2019). Therefore, it provides more time for EFL writing and input during face-to-face meetings (Ghufron & Nurdianingsih, 2021), which significantly improves their English writing skills, primarily in organisation and linking of ideas (Angelini & García-Carbonell, 2019), sentence structure, and mechanics (Putri et al., 2024).

Prior studies indicate that these learning strategies improve the quality of writing (Fathi & Rahimi, 2024; Marzuki et al., 2023) and significantly enhance students' engagement due to their effective features, such as personalized learning, real-time feedback, and

interactive content (Gjermeni & Prodani, 2024), and flexibility (Alghasab, 2020). Encouraged with technology-enhanced resources (e.g., video recordings, narrated presentations, podcasts, and course notes) before attending classes, students can use classroom time for active engagement in learning through varied activities, collaboration with peers and instructors, and immediate personalized feedback (Adnan, 2017), thereby promoting effective learning (Gondal et al., 2024). Owing to the potential of AI-assisted writing tools and the flipped classroom model, researchers have delved into their pedagogical contributions to writing instruction, primarily examined separately; however, empirical evidence on their collaborative effects on improving academic writing and critical thinking skills, in particular, EFL contexts remain scarce. This inadequacy calls for a more in-depth inquiry into how these dual models influence language learning, especially in academic writing, and their interconnection with critical thinking aspects. Considering the paucity of inquiries exploring the benefits of integrating AI-enhanced tools with the flipped classroom approach, this study contributes valuable insights to the literature on advanced pedagogical models for EFL academic writing, along with critical thinking abilities. Of considerable importance, this research fills a crucial void in educational and technological practices by investigating how AI-enhanced writing tools, when merged into a flipped classroom model, can enhance the academic writing and critical thinking competencies of EFL learners.

Method

Research Design

The research applied a field-based one-group pretest-posttest pre-experimental design as the participants comprised a single student cohort (Knapp, 2016), which was embedded in a mixed method framework (Creswell & Creswell, 2018). The quantitative strand measured the changes in students' academic writing and critical thinking performance before and after instructional intervention through the integration of AI-driven writing tools within a flipped classroom model. On the other hand, the qualitative strand explores students' perceptions and lived experiences of the intervention (Rad et al., 2022). This concurrent mixed-methods design was adopted because the two strands addressed complementary research concerns, in which the quantitative component assessed the extent of changes in students' academic writing and critical thinking performance, whereas the qualitative component illuminates the experience and perception through which students engaged with the pedagogical design. The one-group design was adopted as the intervention was implemented in a single intact writing class. Furthermore, withholding access to the AI tools from a subset of students was considered ethically inappropriate because it would create unequal access to instructional support among students enrolled in the same course; therefore, adopting one-group design would allow all the participants to access the AI tools than limiting their use to a control group in an educational context

Participant

As the study adopted the quasi-experimental design with non-randomized participants, an intact cohort of EFL students was selected. A total of 36 second-year students (6 males and 30 females), aged 19-22, enrolled in an academic essay writing course at a university in Medan, North Sumatra, Indonesia. The participants were purposively recruited based on their enrolment in the target course during the study semester. The class was assigned through the university's standard course registration procedure, and no additional selection requirements beyond course enrolment criteria was applied. This intact-class approach facilitated the assessment of the intervention in natural instructional context and preserved the existing course structure. In the pre-intervention phase, all participants have already completed required courses: basic and advanced structure, paragraph writing, and creative writing, to provide foundational writing skills and prepared them for the greater demands of the course. In addition, selecting the second-year students marks a shift from basic skills to advanced cognitive processes, in accordance with the revised Bloom's taxonomy (analysis, evaluation, and creation). Furthermore, all participants possessed basic knowledge of essay composition (introduction, body, and conclusion), yet struggled to develop arguments and construct coherent paragraphs. Notwithstanding their previous experiences in writing classes, none of them had received formal instruction with the assistance of AI writing tools (e.g., ChatGPT, QuillBot, Grammarly) within writing pedagogy. Such a deficient exposure designated them as ideal subjects for assessing the potential of AI-driven writing tools in bolstering students' writing and critical thinking capabilities. Before commencing the study, all the participants were scrupulously briefed about the research objectives, procedures, and phases to ensure that they understood the skeleton of the study. Written informed consent was also obtained from all the participants to confirm voluntary participation and the right to quit or withdraw from the study at any time without academic or personal consequences. This set of protocols was called to uphold ethical standards and ensure transparency at each stage of the research process. For the participants' confidentiality, pseudonyms were used instead of real names throughout the research process, and all the data were used solely for research purposes (Ebadi & Rahimi, 2018). These ethical measures were essential to ensure that students participated openly and confidently in the study.

Instruments and Procedures of Data Collection

The study employed both quantitative and qualitative instruments. Quantitatively, a writing competency test (pretest-posttest) was administered to assess students' writing and critical thinking skills pre-and post-intervention. The test comprised essay writing tasks aligned with the standard EFL academic writing rubrics (Hyland, 2019), assessing content and ideas, organization and coherence, and linguistic accuracy. To design the writing competency test, an essay-writing format modelled on the IELTS Academic Writing Task 2 was adopted, which required students to compose an essay presenting arguments in response to a given topic. This task type was selected for its emphasis on

argumentative writing that aligns with the study's objective of enhancing students' critical thinking through written discourse. The assessment of students' writing performance was referred to the IELTS Academic Writing Task 2 band descriptors, comprising four dimensions: task response (e.g., accurate/relevant and word count), coherence and cohesion (e.g., information organization, effective paragraphing, and the use of linking devices), lexical resources (e.g., suitable collocation and the frequency of errors), and grammatical range and accuracy (e.g., the range of grammatical structures and tenses, correct punctuation, and the incidence of mistakes), with a scoring system ranging from 1 to 9 in steps of 0.5. To uphold the reliability and objectivity of the assessment, the student essays were also reviewed by experienced IELTS instructors as independent raters to mitigate the potential for scoring bias. Moreover, a standardized critical thinking assessment was administered before and after the intervention to assess students' critical thinking development. How well students demonstrate critical thinking skills (as per the California Critical Thinking Skills Test or CCTST) within their essay writing was evaluated through six core cognitive skills: interpretation, analysis, evaluation, inference, explanation, and self-regulation (Facione, 1990). Based on the six critical thinking categories, a modified rubric was developed to assess students' critical thinking skills in their essays using a five-point scale (1 = poor, 5 = excellent). As with the writing assessment, critical thinking was assessed from the same essay texts to ensure that both instruments drew on identical written products rather than separate tasks. Scoring of writing and critical thinking dimensions was conducted separately to prevent halo effects across the two instruments. With respect to the qualitative instrument, the data were collected through an open-ended questionnaire administered to all 36 participants at the end of the intervention. The questionnaire comprised items that addressed students' experiences with the flipped classroom model, their use of AI writing tools, perceived changes in their writing and critical thinking abilities, and the challenges they encountered during the intervention. The items were written in Bahasa Indonesia to ensure the comprehension and provision of the fuller responses. The responses were collected in written form during a dedicated class session to ensure complete data capture.

Before data collection commenced, informed consent was obtained from participants, and the purposes and procedures of the study were explained to ensure ethical compliance. Subsequently, the instruments, including AI writing assistants, instructional videos for the flipped classroom, and scoring rubrics for writing and critical thinking assessments, were prepared. At the beginning of the study, a pretest of the essay writing was administered to assign students to complete the test without AI assistance to measure their baseline writing and critical thinking skills. The writing was, then, assessed by two raters using the IELTS rubric, which evaluates task response, coherence and cohesion, lexical resources, and grammatical range and accuracy. In addition, a critical thinking rubric, adapted from the seminal works of Facione (1990) and Ennis (1993), was utilized to safeguard a rigorous appraisal of critical thinking skills. Following the pretest,

a flipped classroom intervention was implemented. Before each instruction was carried out, the students were provided with resources: videos or readings related to writing skills and the application of AI-powered writing tools. They were encouraged to use AI tools at home for generating ideas, organizing thoughts, or checking grammar, while critically evaluating and refining the AI-generated suggestions. During in-class sessions, the students engaged in discussions, collaborative writing tasks, peer reviews, and revision activities informed by both AI-generated and human feedback. After completing the intervention, they undertook a posttest by composing a new essay under conditions comparable to the pretest without using AI assistance. To uphold consistency, the posttest was assessed using the same rubrics and scoring procedures as the pretest. To gain a deeper understanding of the instructional impact, they were compelled to complete open-ended questionnaires about their experiences with AI tools, flipped learning activities, and perceived improvements in both writing and critical thinking competencies.

Data Analysis

Drawing on both quantitative and qualitative inquiries, this study adopted systematic procedures to strengthen both methodological strands. The initial stage entailed statistical analysis of pretest and posttest data, with descriptive statistics (mean, median, and standard deviation), to provide a detailed account of students' writing and critical thinking skills before and after the intervention. Before subsequent analysis, data normality was assessed to determine whether the dataset conformed to the distribution of normality. As the p-value was below 0.05, the data were deemed non-normally distributed (Ghasemi & Zahediasl, 2012; Razali & Wah, 2011); therefore, the Wilcoxon signed-rank test, a non-parametric alternative, was employed. A p-value < 0.05 was interpreted as indicating a statistically significant difference (Gravetter & Wallnau, 2016; Coolican, 2018). The effect size (Cohen's d) was also calculated to appraise the significance of observed differences alongside the statistical significance. The statistical measurements were performed using IBM SPSS Statistics to establish analytical accuracy. Subsequently, qualitative data from open-ended questionnaire responses were analyzed using thematic analysis, which adhered to Braun and Clarke's (2006) six-phase framework: familiarization, initial coding, identifying themes, reviewing themes, defining and naming the themes, and drawing conclusions. In order to maintain the credibility and trustworthiness of the findings, member checking was employed to verify the accuracy of the interpretations. Peer debriefing and triangulation was further held to strengthen analytical validity by cross-verifying insights across different data sources. In addition, the use of triangulation that synthesized the quantitative and qualitative findings provided a comprehensive evaluation of the effects of AI writing assistance on EFL learners' academic writing and critical thinking in flipped classroom settings.

Results

Quantitative Outcomes: Measurable Learning Gains in Writing and Critical Thinking

This section illustrates a statistical analysis of students' performance in academic writing and critical thinking performance before and after the instructional intervention.

Normality assessment

Prior to the inferential testing, a Shapiro-Wilk test was performed to determine whether the dataset for each variable adhered to normal distribution.

Table 1

Shapiro-Wilk Normality Test

Variable	W Statistic	Df	Sig. (p-value)
Writing pre-test	.915	36	.009
Writing post-test	.958	36	.186
CT pre-test	.930	36	.025
CT post-test	.830	36	.000

As depicted in Table 1 above, three out of the four variables did not meet the assumption of normality ($p < .05$), including the writing post-test, CT pre-test and CT post-test scores. Although the Writing Post-test was normally distributed ($W = .958$, $p = .186$), the Wilcoxon Signed-Rank Test, a non-parametric approach, was applied to both variables for analytical consistency, as the normality was violated in the corresponding pre-test distributions. Therefore, the use of the Wilcoxon Signed-Rank Test was considered appropriate and methodologically justified for analyzing the data.

Descriptive Statistics

Table 2 summarizes the descriptive statistics for both variables, which include means, medians, and standard deviations for both variables at pre-test and post-test.

Table 2

Descriptive Statistics of Pre-test and Post-test Scores (Median Values)

Variable	Mean	Median	SD	Mean Gain	% Improvement
Writing Pre-Test	4.75	5.00	0.88	-	-
Writing Post-Test	6.17	5.95	0.80	+1.42	29.8%
CT Pre-Test	42.87	38.34	13.55	-	-
CT Post-Test	68.14	63.30	7.69	+25.27	58.9%

Students' writing scores increased from a mean of 4.75 ($SD = 0.88$) to 6.17 ($SD = 0.80$), representing a mean gain of 1.42 points, equivalent to a 29.8% improvement. Similarly, critical thinking scores rose more substantially, from a mean of 42.87 ($SD=13.55$) in the pre-test to 68.14 ($SD=7.69$) in the post-test, with a mean gain of 25.27 percentage points, or a 58.9% improvement. In addition to the increase in mean scores, the standard deviation of CT post-test scores (7.69) was considerably smaller than at pre-test (13.55). This suggests that student performance became more homogeneous following the

intervention. These descriptive trends indicate meaningful pre-to-post gains across both skill domains.

Wilcoxon Signed-Rank test

To further examine whether the changes observed between the pre-test and post-test scores were statistically significant, the Wilcoxon Signed-Rank Test was applied to both variables. The results are presented in Table 3.

Table 3

Wilcoxon Signed-Rank Test Results

Variable	Z-Statistic	P-Value	Positive Ranks	Negative Ranks	Ties	Interpretation
Writing Skill	-5.163	< .001	35	0	1	Significant Improvement
Critical Thinking	-5.107	< .001	35	1	0	Significant Improvement

Results from the Wilcoxon signed-rank test showed statistically notable growth in both outcome measures. For writing, the analysis yielded a Z value of -5.163 ($p < .001$), based on 35 non-tied pairs, in which 35 students demonstrated improvement and none declined, with one tied pair was excluded from the analysis that followed standard Wilcoxon procedures. For critical thinking, the test produced a Z value of -5.107 ($p < .001$), based on all 36 paired observations. In this measure, 35 students showed improvement, whereas one student exhibited a marginal decline. The negative Z values in both evaluations confirm the post-test scores which consistently exceeded the pre-test scores, and the p values below .001 provide strong evidence against the null hypothesis. This suggests that the observed gains in both variables were statistically significant. It is important to note, however, that although the Wilcoxon signed-rank test statistically significant increase in both academic writing and critical thinking, these findings should be interpreted with caution. Given the one-group pretest-posttest design, which does not permit strong causal inferences, the observed improvements were associated with the intervention period but cannot be attributed exclusively to the intervention. Alternative explanations, including maturation, history, testing effects, and regression to the mean, remain plausible alternative explanations that a controlled design with a comparison group was required to rule out

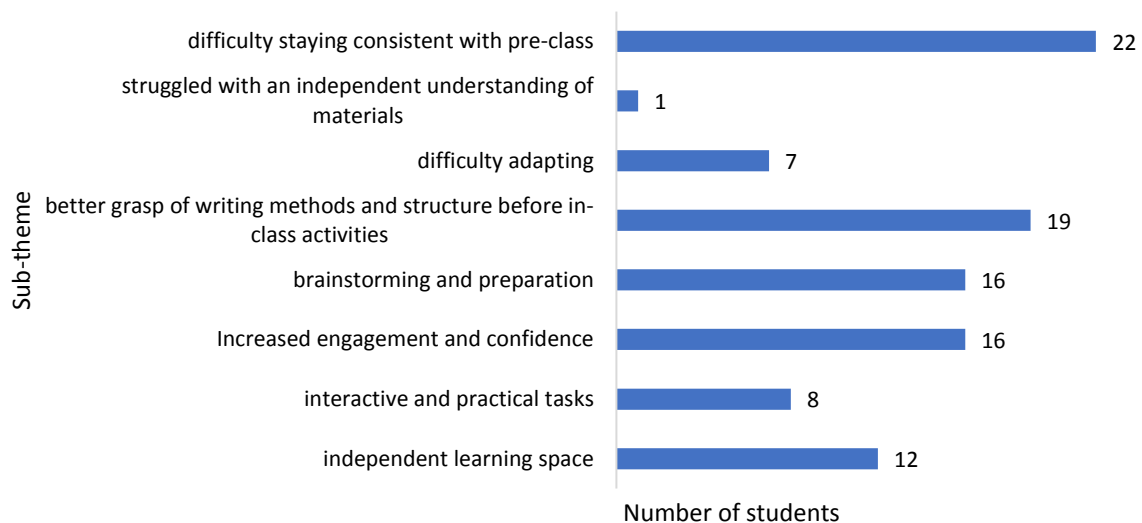
Effect sizes

To assess the practical magnitude of the observed differences, two complementary effect size measures were calculated and are reported in Table 4, namely Rosenthal's r ($r = Z/\sqrt{n}$, where n denotes the number of non-tied pairs) and the matched-pairs rank-biserial correlation (r_{rb}).

Table 4*Effect Size Estimates Based on Wilcoxon Z-Scores*

Variable	Z-Value	N (Non-Tied)	$r=Z/\sqrt{n}$	r_{rb} (Rank-Biserial)	Interpretation
Writing Skill	-5.163	35	.873	1.000	Very Large Effect
Critical Thinking	-5.107	36	.851	.976	Very Large Effect

Using Rosenthal's r ($r = Z/\sqrt{n}$), effect sizes were .873 for writing and .851 for critical thinking. Using the matched-pairs rank-biserial correlation (r_{rb}), the corresponding values were 1.000 and .976, respectively, indicating that nearly all participants improved on both measures. According to conventional benchmarks, these values represent very large effects. Nevertheless, these findings warrant cautious interpretation. First, the effect sizes of this magnitude are relatively uncommon in educational intervention research and may be may partially reflect characteristics of the one-group design. In the absence of a comparison group, the observed changes may reflect not only the intervention but also other factors, such as maturation, testing effects, or naturally occurring development over time. Second, although the $r=Z/\sqrt{n}$ values remain technically valid, the matched-pairs rank-biserial correlation is generally considered more appropriate for Wilcoxon tests, as it reflects the balance between favorable and unfavorable paired differences. Its inclusion in the analysis provides a more methodologically transparent account of the effect. As a whole, the results indicate a strong and consistent positive change over the intervention period, while the limitations of the study design should be considered when interpreting the magnitude of these effects.

*Qualitative Insights Drawn from Post-Intervention Questionnaires**Experiential insight into flipped learning***Figure 1***Student Feedback on the Flipped Classroom Experience*

Evidence from the sub-theme suggests that participants expressed varied viewpoints of the AI-assisted flipped classroom model, acknowledging its potential and practical limitations. From a pedagogical standpoint, a considerable proportion of students (12) admitted that pre-class access to learning materials, such as videos and readings, afforded an independent learning space, which endorsed flexibility to engage with the content at their own pace and learning preferences. A closely related value concerned interactive and practical tasks reported by eight students, explicating that hands-on, experiential activities held during in-class sessions made the learning more engaging and meaningful. Additionally, 16 students attested that early engagement with pre-class materials strengthened their brainstorming and preparation, as reflected in their enhanced capability of planning and organizing ideas more effectively. Such a pedagogical approach also contributed to their increased engagement and confidence, as echoed by other 16 students, impacting more active and meaningful class participation. In parallel, 19 students voiced a better grasp of writing methods and structure before in-class activities, which enabled them to approach the writing tasks with improved guidance and clarity. Benefits notwithstanding, several concerns were also raised. According to students (7) one of the main barriers was the difficulty to adapt to the flipped learning format, including time management and adapting to the self-regulated learning. Student (1) also had problems adapting to the learning model and found it difficult to learn the materials by themselves. Such a concern reflected limited experience of self-regulated learning and difficulty to understand the materials without guided instruction. More common, however, was the challenge of staying consistent with pre-class, with 22 students reporting this. This suggested the significance of ongoing encouragement and follow up to support active participation in pre-class learning.

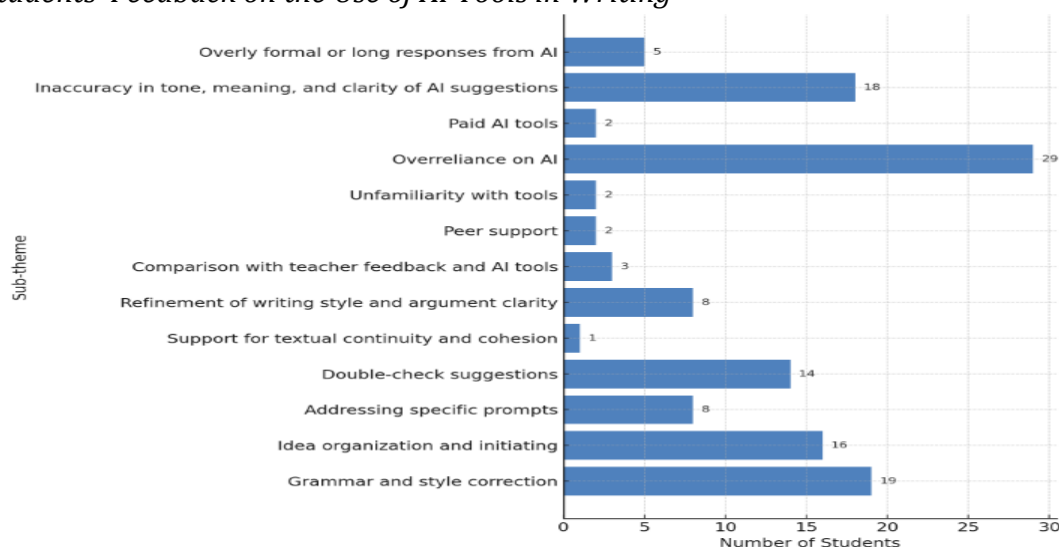
AI writing tools use

The integration of AI-assisted writing tools into instructional settings coincided with meaningful changes in how students performed the writing. As denoted in the chart below, the use of AI writing assistants indicated nuanced views, which presented both potential gains and hurdles that hindered their effective utilization. The most frequently mentioned advantage, cited by 19 students, was the capacity of the tools to deliver grammar and feedback, thus improving linguistic accuracy across grammar, spelling, and expression. Additionally, 16 students perceived the tools supported idea generation and organization, and functioned as a scaffold during the early writing stage. Further advantages were evident in the usefulness of the tools for addressing specific prompts (8) and double-checking suggestions (14), both of which were instrumental in producing and verifying the accuracy and appropriateness of outputs before adopting them into their writing. A smaller number of students also acknowledged the use of AI tools for more targeted purposes, such as supporting textual continuity and cohesion (1) and refining writing style and argument clarity. These affordances enabled students to improve the overall flow of their writing at both sentence and paragraph levels, thus ensuring the logical organization and reader-oriented coherence. Moreover, the other

students (3) used the AI tools as reference points to compare teacher feedback with AI suggestions, which fostered students' greater awareness of the credibility of source and evaluative thinking when revising the writing. Initial operational difficulties were mitigated through peer collaboration, which emerged as a mechanism for developing effective and critical use of AI in writing.

Figure 2

Students' Feedback on the Use of AI Tools in Writing

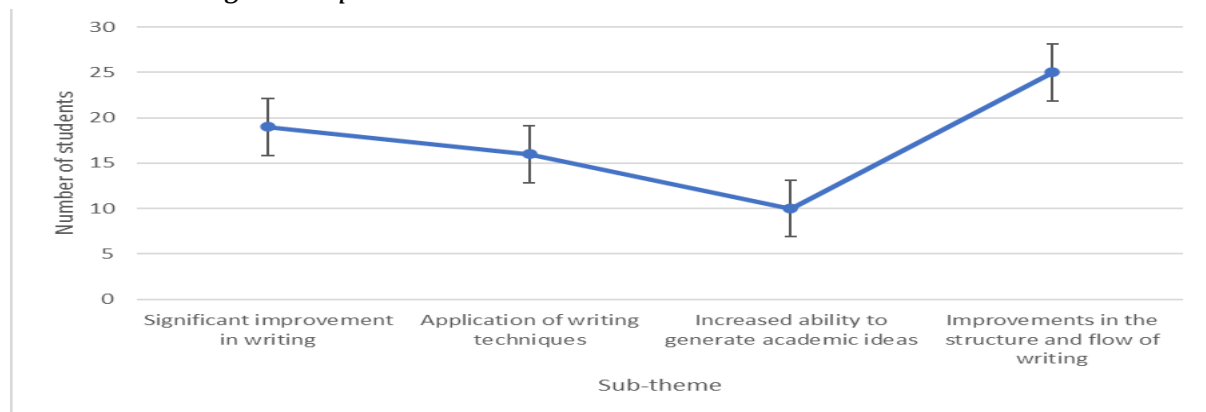


That said, engagement with AI writing tools also raised practical discomforts, most notably overreliance, reported by 29 participants. Students perceived that excessive dependence could undermine confidence in autonomous writing and hinder the development of critical thinking decisive for academic writing. Impediments were also raised by students (18) regarding the *inaccuracy in tone*, meaning, and clarity of AI suggestions, demanding further critical evaluation for appropriate outcomes before embedding them into their written works. Likewise, some students (5) mentioned the unease about the AI's feedback, for it was too formal and long, thus interfering with the well-structured flow and diverging from the intended topic or topic relevance. While, furthermore, a smaller subset of participants (2) reported difficulties stemming from tool unfamiliarity, others (2) accentuated hindrances on limited access to more advanced features due to paid tools. These varied perspectives signal the need for not just using the AI but also discerning how to use it reasonably; thus, its integrations into learning settings should be accompanied by ongoing support to help students benefit from its convenience without becoming overly reliant or overwhelmed by its flaws.

Development of writing performance

When exploring the students' writing performance, the post-questionnaire data revealed four main areas of improvement, coming up significantly across participants' reflections, as visualized in the chart below.

Figure 3
Students' Writing Development

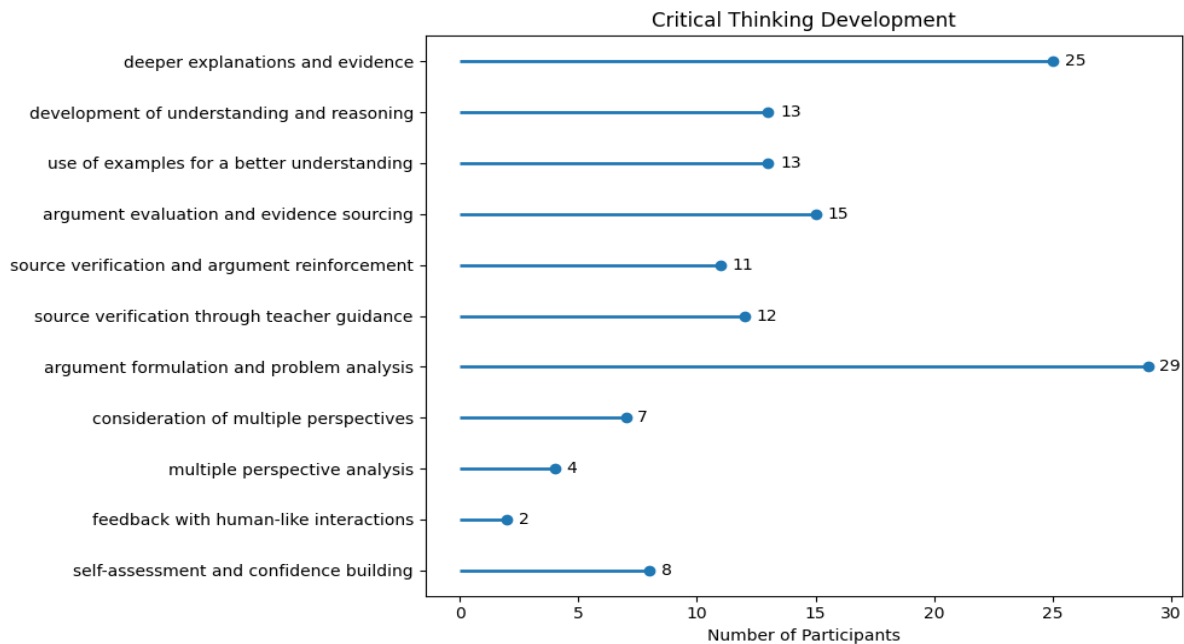


It showcases how the integrated teaching approach fostered measurable strides in their writing skills. Spanning the four sub-themes, a group of 19 students claimed that the flipped classroom model, when combined with AI, drove significant progress in their overall writing skills and quality. This growth was evident in their improved ability to organize ideas and select relevant vocabulary, alongside more formal academic language, factors that impact their writing proficiency and process. This illuminated the capacity of instructional design to cultivate students' technical aptitude, considerably required for effective writing. In light of the second subcategory, application of writing techniques, such as outlining, planning, and revising, was another benefit of the designed pedagogical strategies, according to 16 students. This consequently helped students not only deepen their understanding of the writing principles but also apply them in the composition practices, culminating in structured and coherent written output. It is also evident that the designed model benefited from an increase in the ability to generate academic ideas, as shared by 10 students. Within this strand, it allowed learners to formulate and express their thoughts more fluently, particularly in developing a thesis statement and strong supporting arguments. Coupled with this, across the sub-themes, the benefits most frequently reported by students were improvements in structure and flow. Through the AI-assisted writing in the flipped model, students were capable of developing their thoughts more effectively.

Critical thinking development

Of the four categories of benefits revealed by students during in-class learning practices, the development of critical thinking stood out as the most prominent, as displayed in the following chart.

Figure 4
Critical Thinking Skill Improvement



Within this parameter, 25 students admitted that the AI-supported flipped learning model enhanced their ability to provide deeper explanations and evidence. Learning within such a guided framework strengthened students' awareness of evidence-based argumentation, resulting in more academically sound writing. Supporting this trend, 13 students reported the development of understanding and reasoning, which reflected their ability to interpret concepts and, in turn, craft them into more logical and coherent ideas. Similarly, 13 students noted greater proficiency in explaining abstract concepts through the use of examples for a better understanding. Another, though related, 15 students clarified growth in argument evaluation and evidence sourcing, underpinned by 11 students who professed a favourable effect of blended instructional scaffolds in source verification and argument reinforcement, strengthening their confidence in scrutinizing argument validity supported by credible sources. This pattern resonates with higher-order thinking, in which students moved beyond constructing arguments to rigorously verifying their credibility. Closely related, though less frequently mentioned, was increased familiarity with evaluating sources, noted by two participants. Moreover, the progress was also ascribed to source verification through teacher guidance, as manifested by 12 students, reflecting their heightened awareness of the importance of evaluating information credibility and the instructors' role in assessing the validity and credibility of information and sources. A related cognitive shift involved advances in argument formulation and problem analysis, expressed by 29 students. This increase represents a key outcome of the course's pedagogical design, as students became more adept at addressing complex issues and constructing well-reasoned arguments, such as pinpointing core issues, proposing logical responses, and articulating ideas coherently.

Completing this, consideration of multiple perspectives also emerged as one of the constructs under this theme, noted by seven participants, recognizing the importance of adopting various viewpoints to build compelling arguments. This view was further reinforced by four students who undertook multiple perspective analysis, which they perceived as another strength of the combined pedagogical practices. Within this benefit, students demonstrated the ability to critically compare divergent views and merge them into their reasoning. Additionally, for the other small number of students (2), the AI-supported model created feedback with human-like interactions that encouraged them to question their own work, reflect more deeply, and refine their ideas, making them more polished. Ultimately, eight students reported that the integrated pedagogical scaffolds supported self-assessment and confidence building, which advanced their reflective thinking, particularly in identifying flaws in reasoning and restructuring their arguments, alongside self-directed commitment to continual improvement. Drawing on this strand, the analysis revealed no reported drawbacks in the combined learning model's support for critical thinking; instead, participants highlighted enhanced opportunities for deeper exploration, more systematic reasoning, and increased confidence in their thinking.

Discussion

This study points to the strong potential of integrating AI-driven writing feedback with the flipped classroom model to enhance EFL learners' academic writing and critical thinking skills. The substantial development of both skills was evidenced by a mean of 1.42 (29.8%) for writing performance, and critical thinking scores increased by a mean of 25.27 (58.9%). These gains were statistically significant at $p < .001$ for both variables and were accompanied by very large effect sizes, specifically the rank-biserial correlations of $r_{rb} = 1.000$ for writing, which suggests that the non-tied student improved, whereas the $r_{rb} = .976$ for critical thinking indicates improvement among 35 of the 36 participants. Taken as a whole, the data provide a strong and directionally consistent development of both outcomes over the course of the intervention. It is important to clarify, however, that, as the study adopted a one-group pretest–posttest design, the observed gains in both variables cannot be interpreted solely as evidence of the intervention's causal efficacy, as they may also be associated with other factors, such as maturation and testing effects. Accordingly, a controlled design incorporating a delayed-treatment comparison group would be needed to isolate the treatment effect more precisely. One additional finding worth noting is the marked reduction in the standard deviation of critical thinking scores from pre-test ($SD = 13.55$) to post-test ($SD = 7.69$). This convergence may suggest that students who entered the intervention with lower critical thinking scores experienced relatively greater gains, thereby reducing the variability in scores. Although this interpretation should be treated with caution, this pattern is consistent with the potential scaffolding function of both AI feedback tools and the flipped model's pre-class preparation structure, which provided support for students with weaker baseline levels.

The statistically significant gain in writing performance, substantiated by qualitative themes associated with language accuracy, structural organization, and the application of writing techniques, points to the combined intervention contributed to the development of students' linguistic and compositional abilities. In particular, students demonstrated measurable improvements in organizing ideas, deploying appropriate vocabulary, maintaining coherence and cohesion, and managing grammatical accuracy. This finding converges with prior research which demonstrated that AI writing tools improved overall writing quality (Fathi & Rahimi, 2024; Namaziandost, 2025), and that flipped learning models strengthened students' writing production by enabling more deliberate engagement with content prior to in-class application (Putri et al., 2024; Angelini & García-Carbonell, 2019). A notable qualitative dimension of the improvement in writing concerned the role of pre-class exposure to learning materials. As reflected in students' reported experience, the flipped learning model was perceived to contribute to improvements, in particular, in their brainstorming and idea-organization abilities through engagement with preparatory materials before class sessions. This implies that the time afforded by the flipped model for independent pre-reading and viewing supported the planning and structuring stages of the writing process. This interpretation accords with Bergmann and Sams' (2012) foundational argument that realigning content delivery outside the classroom creates greater cognitive space for deeper processing and application during face-to-face instruction. Complementing this, the AI tools' capacity to provide immediate and personalized feedback on grammar, style, and textual coherence facilitate students to focus greater cognitive effort on higher-order thinking, such as argument development, critical evaluation and clarity of idea, as echoed by Kim et al. (2024) and Polakova and Ivenz (2024).

With respect to cognitive enrichment, the marked increase in mean critical thinking scores comprises the most striking quantitative finding of this study and warrants careful interpretation. Proportionally, the intervention was associated with a greater improvement in critical thinking than in language-based outcomes, a pattern that reflects the relatively lower baseline position of many students on the critical thinking scale, as well as the potentially generative nature of argument-focused AI feedback in stimulating analytical engagement. The observed improvement corresponded with the findings of Yin and Dou (2025) who found that AI-assisted critical thinking interventions enhanced EFL learners' writing proficiency by fostering analytical reasoning. This provides further evidence that integrating critical thinking into writing instruction can help students move beyond formulaic writing toward more analytically sophisticated forms of expression. The qualitative interpretation further highlighted the improved argument formulation and problem analysis as the most frequently cited dimensions of cognitive growth. As indicated in the students' responses, they became more adept at identifying core issues, constructing well-reasoned responses, and articulating ideas in coherent argumental structures. Such development in analytical reasoning echoes Abeysekera and Dawson's (2014) theoretical account of the flipped learning model, which proposes that

reallocating in-class time to active problem-solving and collaborative engagement creates the conditions conducive to the development of higher-order thinking, a skill essential for 21st-century learning (Ayçiçek, 2021). Within this designed pedagogical practice, AI tools emerged to further extended these opportunities by prompting students to evaluate, verify, and revise their own arguments in response to automated feedback, a process that resonates with the evaluative and inferential dimensions of Facione's (1990) critical thinking framework. These indicate that the intervention effectively moved students beyond surface-level comprehension toward a more critically engaged approach to writing. Further qualitative evidence points to the development in evidence-based reasoning, source verification, and the consideration of multiple perspectives. As reported by students, the gains in argument evaluation and evidence sourcing, and the critical appraisal of information credibility. Taken together, these accounts indicate that the most substantial cognitive gains observed in the intervention may not have resulted from AI use alone, but rather from the interaction among AI scaffolding, peer collaboration, and instructor-mediated guidance. This pattern supports the conceptualization of the flipped classroom as a socially and cognitively dynamic learning environment rather than merely a reorganization of instructional delivery.

Furthermore, a recurring theme that emerged from the quantitative findings and the qualitative data was the interdependent relationship between the development of writing and critical thinking. Those who reported greater involvement in argumentation also tended to report gains in the structure and coherence of their writing. Those who reported greater gains in writing skills often linked those gains to a clearer understanding of how to develop and support an argument. This reciprocal patterning, whereby the development of one skill triggered the other, is well documented in the literature (Shen & Teng, 2024; Wang & Newell, 2025; Woldt & Nenad, 2021) and is corroborated by the present data. The integrated pedagogical design of this study, where AI writing tools and flipped learning structures were applied simultaneously and continuously, not as separate interventions, were particularly successful in nurturing this reciprocal relationship, as students were constantly asked to write and think critically about their writing in the same task contexts, which possibly offered opportunities for the development of these skills in tandem. While the pedagogical promise of the paired learning approach is evident, there are challenges that have emerged including dependence on AI, unfamiliarity with the tools, errors in the tone, meaning, and clarity of the feedback generated by AI, responses that are overly formal or verbose, inconsistent engagement with pre-class preparation, and challenges with independent understanding of the materials. These challenges require pedagogical approaches that encourage students to critically engage with AI-generated outputs instead of relying on them uncritically (Tang et al., 2024) and informed policies and ethical guidelines (Cong et al., 2024). Moreover, appropriate training schemes should be designed to progressively introduce AI functionalities and allow students to become acquainted with the tools.

Implications for Pedagogy and Future Research

The findings offer several implications for technology-integrated EFL writing instruction. First, integrating AI tools into the flipped learning model provide a productive pedagogical approach for developing linguistic and cognitive competencies in parallel, in particular, among learners who enrol academic writing courses with limited prior exposure to writing conventions. Second, the concerns raised by the students about overreliance on AI and the inaccuracy of some AI-generated suggestions highlight the need for explicit instruction in critical AI literacy, particularly in how to evaluate, interrogate, and apply automated feedback. Providing access to AI tools without accompanying effective guidance on their limitations may therefore be insufficient to foster the deeper analytical engagement evident among the students who used the technology more reflectively. Third, given the limitations of the one-group design, future research was recommended to employ comparative designs, ideally incorporating delayed-treatment comparison groups, to more rigorously examine the extent to which observed gains can be attributed to the intervention rather than to developmental or other confounding factors. Longitudinal studies would also be valuable for determining whether the observed improvements are sustained beyond the immediate post-test period. Finally, the marked proportional gains in critical thinking relative to writing suggest that AI-enhanced flipped learning may have particular potential as a model for developing higher-order cognitive skills in EFL contexts, an area that has received comparatively less empirical attention than surface-level language accuracy. Further research should therefore investigate the mechanisms through which AI feedback may support analytical reasoning, rather than focusing solely on whether improvement occurs.

Conclusion

The findings emphasize that integrating AI-driven writing assistants into a flipped classroom exerts measurable improvements in writing and critical thinking performance among EFL college students. Anchored in a mixed-methods design, the study revealed that access to pre-class writing materials afforded students greater opportunities to comprehend key concepts in advance of classroom engagement. Meanwhile, the qualitative evidence indicated generally positive feedback toward the AI-assisted flipped classroom model with some constraints. The AI-facilitated flipped learning intervention based on participants' insights offered advantages such as learner autonomy, interactive and practice-oriented learning, and effective pre-class brainstorming and preparation. Also, early access and review of instructional materials allowed learners to comprehend better during in-class activities, while the integration of AI tools gave flexibility in improving their writing performance. AI specifically helps to improve style and clarity of argument, improve cohesion and coherence, confirm suggestions through double checking, and improve organization and grammatical accuracy. These affordances support students' development of writing especially in core writing strategies such as planning, outlining, and revising and support the generation of academic content and

structural flow, making a significant contribution to students' writing development (Marzuki et al., 2023). Most importantly, the integrated pedagogical design was related to the significant gains of higher-order cognitive competencies, especially in forming arguments, evaluating evidence, verifying information sources, and considering diverse perspectives, the capacity that relates to the core dimensions of Facione's (1990) critical thinking framework. The findings suggest that writing and critical thinking in EFL contexts are highly interdependent, suggesting mutually reinforcing benefits of integrated instructional designs. The parallel development of both skills is probably due to similar cognitive processes. Moreover, the study redefines the role of AI in language learning from a pedagogy, not only a technological aid (Werdiningsih et al., 2024) but also a cognitive scaffold to promote reflective thinking and learner autonomy. This study extends the flipped classroom pedagogy by showing that AI-supported pre-class engagement can enhance in-class collaboration and peer interaction, especially in EFL writing contexts.

ORCID

 <https://orcid.org/0000-0002-1697-8906>

 <https://orcid.org/0000-0002-3639-3948>

 <https://orcid.org/0009-0009-4334-6216>

 <https://orcid.org/0000-0002-1443-7567>

Publisher's Note

The claims, arguments, and counter-arguments made in this article are exclusively those of the contributing authors. Hence, they do not necessarily represent the viewpoints of the authors' affiliated institutions, or EUROKD as the publisher, the editors and the reviewers of the article.

Acknowledgements

The authors gratefully acknowledge the institutional support provided by Universitas Muhammadiyah Sumatera Utara and sincerely thank all the students who participated in this study for their valuable time and contributions.

Funding

This work was financially supported by Universitas Muhammadiyah Sumatera Utara, Indonesia under Grant No. 151/II.3-AU/UMSU-LP2M/C/2023.

CRedit Authorship Contribution Statement

Pirman Ginting: Conceptualization, Investigation, Methodology, Project Administration, Supervision, Writing–Original Draft

Yenni Hasnah: Data Curation, Formal Analysis, Investigation, Writing–Review & Editing

Adib Jasni Kharisma: Formal Analysis, Investigation, Visualization, Writing–Review & Editing

Dodi Mulyadi: Validation, Writing–Review & Editing

Generative AI Use Disclosure Statement

The authors used Grammarly and GPT to improve the manuscript's proofreading and language polish. The accuracy and integrity of the final manuscript were examined and revised thoroughly.

Ethics Declarations

Prior to data collection, ethical approval for this study was granted by the appropriate institutional authority. Written informed consent was secured from all participants, ensuring that participation was entirely voluntary and that participants could withdraw from the study at any time without academic or personal repercussions. Participants were fully informed about the objectives of the study, procedures, and their rights before giving consent. To ensure the confidentiality, all data were used solely for research purposes and remained strictly confidential.

Consent for Publication

Informed consent was obtained from the participants for the use and publication of their research data. No personally identifiable information is included in this article.

Competing Interests

The authors declare that they have no competing interests.

Data Availability

The data supporting the findings of this study comprise students' writing and interview data. Owing to ethical consideration and participant confidentiality, these data are not publicly accessible. Data may be obtained from the author upon reasonable request.

References

- Abeysekera, L., & Dawson, P. (2014). Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. *Higher Education Research & Development*, 34(1), 1–14. <https://doi.org/10.1080/07294360.2014.934336>
- Adnan, M. (2017). Perceptions of senior-year ELT students for flipped classroom: A materials development course. *Computer Assisted Language Learning*, 30(3–4), 204–222. <https://doi.org/10.1080/09588221.2017.1301958>
- Alghasab, M. B. (2020). Flipping the writing classroom: Focusing on the pedagogical benefits and EFL learners' perceptions. *English Language Teaching*, 13(4), 28–40. <https://doi.org/10.5539/elt.v13n4p28>
- Al-Sofi, B. B. M. A. (2024). Artificial intelligence-powered tools and academic writing: To use or not to use ChatGPT. *Saudi Journal of Language Studies*, 4(3), 145–161. <https://doi.org/10.1108/SJLS-06-2024-0029>
- Angelini, M. L., & García-Carbonell, A. (2019). Enhancing students' written production in English through flipped lessons and simulations. *International Journal of Educational Technology in Higher Education*, 16(1). <https://doi.org/10.1186/s41239-019-0131-8>
- Ayçiçek, B. (2021). Integration of critical thinking into curriculum: Perspectives of prospective teachers. *Thinking Skills and Creativity*, 41, Article 100895. <https://doi.org/10.1016/j.tsc.2021.100895>

- Bandyopadhyay, S., & Szostek, J. (2018). Thinking critically about critical thinking: Assessing critical thinking of business students using multiple measures. *Journal of Education for Business*, 94(4), 259–270. <https://doi.org/10.1080/08832323.2018.1524355>
- Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day*. International Society for Technology in Education.
- Bezzazi, R. (2019). Learning English grammar through flipped learning. *The Asian Journal of Applied Linguistics*, 6(2), 170–184. <https://caes.hku.hk/ajal/index.php/ajal/article/view/597>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Carter, A. G., Creedy, D. K., & Sidebotham, M. (2017). Critical thinking evaluation in reflective writing: Development and testing of carter assessment of critical thinking in midwifery (reflection). *Midwifery*, 54, 73–80. <https://doi.org/10.1016/j.midw.2017.08.003>
- Celuch, K., & Slama, M. (2000). Student perceptions of a marketing course taught with the critical thinking approach. *Marketing Education Review*, 10(1), 57–64. <https://doi.org/10.1080/10528008.2000.11488695>
- Cong, Y., Yang, L., & Ergün, A. L. P. (2024). Exploring the relationship between burnout, learning engagement and academic self-efficacy among EFL learners: A structural equation modelling analysis. *Acta Psychologica*, 248, Article 104394. <https://doi.org/10.1016/j.actpsy.2024.104394>
- Coolican, H. (2018). *Research methods and statistics in psychology* (7th ed.). Routledge.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approach* (5th ed.). Sage Publications.
- Crusan, D., & Mohebbi, H. (2026). Editorial: The role of AI and ChatGPT in giving corrective feedback: Panacea or unfulfilled promise. *Feedback Research in Second Language*, 4, 1–5. <https://doi.org/10.32038/frsl.2026.04.01>
- Dong, L. (2026). When machines teach writers: Reflecting on GenAI's role in giving corrective feedback on linguistic CAF facets. *Feedback Research in Second Language*, 4, 28–39. <https://doi.org/10.32038/frsl.2026.04.04>
- Ebadi, S., & Rahimi, M. (2018). An exploration into the impact of WebQuest-based classroom on EFL learners' critical thinking and academic writing skills: a mixed-methods study. *Computer Assisted Language Learning*, 31(5–6), 617–651. <https://doi.org/10.1080/09588221.2018.1449757>
- Ennis, R. H. (1993) Critical thinking assessment. *Theory into Practice*, 32(3), 179–186. <https://doi.org/10.1080/00405849309543594>
- Facione, P. A. (1990). *The California critical thinking skills test--college level: Technical REPORT #1, experimental validation and content validity*. California Academic Press.
- Fathi, J., & Rahimi, M. (2024). Utilising artificial intelligence-enhanced writing mediation to develop academic writing skills in EFL learners: a qualitative study. *Computer Assisted Language Learning*, 39(1–2), 263–302. <https://doi.org/10.1080/09588221.2024.2374772>
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486–489. <https://doi.org/10.5812/ijem.3505>
- Ghufron, M. A., & Nurdianingsih, F. (2021). Flipped classroom method with computer-assisted language learning (CALL) in EFL writing class. *International Journal of Learning Teaching and Educational Research*, 20(1), 120–141. <https://doi.org/10.26803/ijlter.20.1.7>
- Gjermeni, F., & Prodani, F. (2024). AI and student engagement: A comparative analysis. *Interdisciplinary Journal of Research and Development*, 11(3), 195. <https://doi.org/10.56345/ijrdv11n326>
- Golan, R., Reddy, R., Muthigi, A., & Ramasamy, R. (2023). Artificial intelligence in academic writing: A paradigm-shifting technological advance. *Nature Reviews Urology*, 20, 327–328. <https://doi.org/10.1038/s41585-023-00746-x>
- Gondal, S. A., Khan, A. Q., Cheema, E. U., & Dehele, I. S. (2024). Impact of the flipped classroom on students' academic performance and satisfaction in pharmacy education: A quasi-experimental study. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2378246>
- Gravetter, F. J., & Wallnau, L. B. (2016). *Statistics for the behavioral sciences* (10th ed.). Cengage Learning.
- Güven, Ş. D., Calpbinici, P., Kuzgun, H., & Çelik, G. K. (2020). Does writing patient care daily effects on critical thinking? A pilot study with 1st year nursing students. *Thinking Skills and Creativity*, 35, Article 100638. <https://doi.org/10.1016/j.tsc.2020.100638>
- Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press.
- Jia, C., & Lu, Y. (2026). Exploring generative AI assisted-EFL academic writing: An activity theory perspective. *Journal of English for Academic Purposes*, 81, Article 101672. <https://doi.org/10.1016/j.jeap.2026.101672>

- Kafri, B. A. (2022). Critical thinking (CT) in sustainable higher education: Ensuring consistent CT perception-practice and identifying gaps between college instructors' and students' perceptions in advanced academic writing courses in the UAE. *Thinking Skills and Creativity*, 46, Article 101182. <https://doi.org/10.1016/j.tsc.2022.101182>
- Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies*, 30(1), 1265–1300. <https://doi.org/10.1007/s10639-024-12878-7>
- Knapp, T. R. (2016). Why is the one-group pretest-posttest design still used? *Clinical Nursing Research*, 25(5), 467–472. <https://doi.org/10.1177/105477381666662>
- Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, 10(2). <https://doi.org/10.1080/2331186X.2023.2236469>
- Naber, J., & Wyatt, T. H. (2014). The effect of reflective writing interventions on the critical thinking skills and dispositions of baccalaureate nursing students. *Nurse Education Today*, 34(1), 67–72. <https://doi.org/10.1016/j.nedt.2013.04.002>
- Namaziandost, E. (2025). Integrating flipped learning in AI-enhanced language learning: mapping the effects on metacognitive awareness, writing development, and foreign language learning boredom. *Computers and Education Artificial Intelligence*, 9, Article 100446. <https://doi.org/10.1016/j.caeai.2025.100446>
- Piergiorganni, P. R. (2014). Creating a critical thinker. *College Teaching*, 62(3), 86–93. <https://doi.org/10.1080/87567555.2014.896775>
- 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, E. D., Cahyono, B. Y., & Zubaidi, N. (2024). Flipped learning in EFL classrooms effects on tertiary students' writing skills and perceptions. *Teaching English as a Second or Foreign*, 28(2). <https://doi.org/10.55593/ej.28110a7>
- Rad, H. S., Namaziandost, E., & Razmi, M. H. (2022). Integrating STAD and flipped learning in expository writing skills: Impacts on students' achievement and perceptions. *Journal of Research on Technology in Education*, 55(4), 710–726. <https://doi.org/10.1080/15391523.2022.2030265>
- Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2(1) 21–33. <https://www.nrc.gov/docs/ML1714/ML17143A100.pdf>
- Schicker, J. (2018). Teaching writing and grammar through critical thinking in a genre-based upper-level undergraduate German class. *Die Unterrichtspraxis/Teaching German*, 51(1), 63–75. <https://doi.org/10.1111/tger.12055>
- Seo, J. Y. (2024). Exploring the educational potential of ChatGPT: AI-assisted narrative writing for EFL college students. *Language Teaching Research Quarterly*, 43, 1–21. <https://doi.org/10.32038/ltrq.2024.43.01>
- Shen, X., & Teng, M. F. (2024). Three-wave cross-lagged model on the correlations between critical thinking skills, self-directed learning competency and AI-assisted writing. *Thinking Skills and Creativity*, 52, Article 101524. <https://doi.org/10.1016/j.tsc.2024.101524>
- Stephen, C. C., & Muriel, L. Y. (2013). Narrative and performative acts in cultural education: the teaching of writing as critical communication. *Inter-Asia Cultural Studies*, 14(2), 213–227. <https://doi.org/10.1080/14649373.2013.769747>
- Tanaka, J., & Gilliland, B. (2016). Critical thinking instruction in English for academic purposes writing courses: A dialectical thinking approach. *TESOL Journal*, 8(3), 657–674. <https://doi.org/10.1002/tesj.291>
- Tang, K-S., Cooper, G., Rappa, N., Cooper, M., Sims, C., & Nonis, K. (2024). A dialogic approach to transform teaching, learning & assessment with generative AI in secondary education: a proof of concept. *Pedagogies: An International Journal*, 19(3), 493–503. <https://doi.org/10.1080/1554480X.2024.2379774>
- Umar, I. N., & Rathakrishnan, M. (2012). The effects of online teachers' social role and learning style on students' essay writing performance and critical thinking in a wiki environment. *Procedia-Social and Behavioral Sciences*, 46, 5730–5735. <https://doi.org/10.1016/j.sbspro.2012.06.506>
- Wang, J., Liao, Y., Liu, S., Zhang, D., Wang, N., Shu, J., & Wang, R. (2024). The impact of using ChatGPT on academic writing among medical undergraduates. *Annals of Medicine*, 56(1). <https://doi.org/10.1080/07853890.2024.2426760>

- Wang, Q., & Newell, G. E. (2025). Teaching and learning argumentative writing as critical thinking in an EFL composition classroom. *Learning Culture and Social Interaction*, 51, Article 100891. <https://doi.org/10.1016/j.lcsi.2025.100891>
- Werdiningsih, I., Marzuki, N., & Rusdin, D. (2024). Balancing AI and authenticity: EFL students' experiences with ChatGPT in academic writing. *Cogent Arts and Humanities*, 11(1). <https://doi.org/10.1080/23311983.2024.2392388>
- Woldt, J. L., & Nenad, M. W. (2021). Reflective writing in dental education to improve critical thinking and learning: A systematic review. *Journal of Dental Education*, 85(6), 778–785. <https://doi.org/10.1002/jdd.12561>
- Xu, T., & Jumaat, N. F. (2024). ChatGPT-empowered writing strategies in EFL students' academic writing: Calibre, challenges and chances. *International Journal of Interactive Mobile Technologies (ijIM)*, 18(15), 95–114. <https://doi.org/10.3991/ijim.v18i15.49219>
- Yildirim, B., Özkahraman, S., & Karabudak, S. S. (2011). The critical thinking teaching methods in nursing students. *International Journal of Research in Business and Social Science*, 2(24), 174–182.
- Yin, X., & Dou, K. (2025). An AI-assisted critical thinking intervention to enhance undergraduate EFL learners' writing proficiency. *Studies in Educational Evaluation*, 86, Article 101480. <https://doi.org/10.1016/j.stueduc.2025.101480>
- Yin, X., Bin Mohd Saad, M. R., & Halim, H. B. A. (2024). Technology-enhanced social learning (TSL) to foster critical thinking dispositions and thinking in writing. *Cogent Education*, 11(1), Article 2341584. <https://doi.org/10.1080/2331186X.2024.2341584>
- Yin, X., Saad, M. R. B. M., & Halim, H. B. A. (2023). A systematic review of critical thinking instructional pedagogies in EFL writing: What do we know from a decade of research. *Thinking Skills and Creativity*, 49, Article 101363. <https://doi.org/10.1016/j.tsc.2023.101363>
- Zhai, N., & Ma, X. (2022). The effectiveness of automated writing evaluation on writing quality: A meta-analysis. *Journal of Educational Computing Research*, 61(4), 875–900. <https://doi.org/10.1177/07356331221127300>
- Zhang, Z., Aubrey, S., Huang, X., & Chiu, T. K. F. (2025). The role of generative AI and hybrid feedback in improving L2 writing skills: A comparative study. *Innovation in Language Learning and Teaching*, 1–19. <https://doi.org/10.1080/17501229.2025.2503890>