

Using ChatGPT as a writing assistant: A study on essay quality development among Indonesian EFL University Students

Celiacika Gustisiwi Puteri, Bambang Yudi Cahyono, Utami Widiati* and Nunung Suryati

Department of English Language Education, Faculty of Letters, Universitas Negeri Malang, Jl. Semarang No. 5, Malang, East Java 65145, Indonesia

ABSTRACT

AI-driven tools, such as ChatGPT, can be used to provide feedback on the content and language aspects of essay writing. This study investigates the effect of utilizing ChatGPT as a writing assistant on the quality of essays produced by undergraduate Indonesian EFL students. The study employed a quasi-experimental design, involving 35 participants divided into an experimental group (N = 18) and a control group (N = 17). Pre-test and post-test scores were compared statistically to measure essay quality. Additionally, a questionnaire was used to explore the perceptions of students in the experimental group on ChatGPT as a writing assistant. Findings indicated significant differences in essay quality between the two groups from the pre-test to the post-test. However, post-test scores showed no significant difference between the experimental and the control groups. The survey results showed that students in the experimental group expressed positive perceptions of using ChatGPT, as measured across four categories: usefulness, ease of use, ease of learning, and satisfaction. The study concludes that although ChatGPT did not produce a statistically significant increase in essay quality when compared to the conventional method alone, the findings of student perceptions suggest that it could be used as a secondary tool. This is evidenced by a survey showing that access to ChatGPT improved the writing experience with real-time and tailored feedback, which increased student engagement, confidence, and writing autonomy. Thus, EFL instructors can adopt the use of ChatGPT in their essay writing courses in conjunction with the conventional method.

Keywords: Artificial intelligence; ChatGPT; EFL students; essay quality

Received:

6 January 2025

Revised:

30 August 2025

Accepted:

16 September 2025

Published:

30 September 2025

How to cite (in APA style):

Puteri, C. G., Cahyono, B. Y., Widiati, U., & Suryati, N. (2025). Using ChatGPT as a writing assistant: A study on essay quality development among Indonesian EFL University Students. *Indonesian Journal of Applied Linguistics*, 15(2), 332-346.
<https://doi.org/10.17509/3ebmy158>

INTRODUCTION

Within the domain of argumentative writing, which requires a higher level of critical thinking and linguistic competence (Hillocks, 2009), the capacity of ChatGPT to scaffold the writing process has been recognized as particularly beneficial (Abramson, 2023; Miao & Holmes, 2021). ChatGPT brings much-needed innovation to the teaching of argumentative writing due to the following advantages. First, it helps students in brainstorming and idea generation (Baskara, 2023), allowing for a

well-crafted thesis and logical progression of arguments from the start. Second, it assists in the development of content by expanding upon arguments, providing relevant support for claims, and minimizing opposition through well-articulated counterarguments (Kim, 2024; Song & Song, 2023). Third, it serves as a beneficial editing and proofreading tool (Song & Song, 2023; Vovk & Kryvosyia, 2024), providing instantaneous feedback on grammar, word choice, and sentence structure, allowing students to access this support

*Corresponding author
Email: utami.widiati.fs@um.ac.id

and improve their writing, thereby increasing linguistic accuracy. Ultimately, it facilitates reflection and revision, enabling students to identify imperfections and make effective changes independently (Baskara, 2023).

Additionally, it provides solutions to age-old dilemmas faced by students and teachers alike. The capability to brainstorm ideas, help structure an argument, and provide feedback on content and sentence-level edits enables the chatbot to serve as a writing guide (Marzuki et al., 2023). Moreover, ChatGPT assumes different roles throughout the writing process, not just in pre-writing, but also in post-writing (Su et al., 2023). During pre-writing activities, students can utilize ChatGPT to brainstorm ideas, create outlines, and receive assistance with structuring their essays. This early-stage assistance helps cultivate well-organized and logically coherent compositions. Moreover, in the post-writing phase, it offers students the opportunity to refine their work by identifying counterarguments.

This writing genre is often considered challenging. Students experience this in content areas as they must form their own logical claims, meaning their ideas need to be coherent and consistent, following the same line of thinking while providing supporting evidence for their beliefs. In essence, students are not merely offering ideas on a suggested topic; they are arguing that their ideas are the most effective and valid among all those taught thus far. Therefore, it requires higher-order thinking, such as elaboration, justification, and assessment (Noroozi et al., 2023), as well as critical reasoning to support such claims (Newell et al., 2011). Yet in real-world situations, students struggle when asked to write argumentative essays (Aertselaer & Dafouz-Milne, 2023; Dornbrack & Dixon, 2014; Latifi et al., 2021; Ranjbaran et al., 2023). For example, failure to include counterarguments is common (Ranjbaran et al., 2023). In addition, students struggle not only with creating a thesis but also with offering significant support, organizing their thoughts, and concluding effectively (Newell et al., 2011; Toan et al., 2020). Such struggles arise because students perceive the material as irrelevant and believe they lack the necessary interest or prior knowledge to evaluate the claims presented to them (Limón, 2001; Tandiana et al., 2017). Additionally, struggles also arise due to a lack of familiarity with certain lexical features and mastery of grammar (Chanie, 2013).

On one hand, a study found that ChatGPT can be easily incorporated into argumentative essay writing courses. Su et al. (2023) detail how the use of ChatGPT could assist in brainstorming, content adjustment, editing, and post-essay reflective exercises. Such implementation could reduce the burden on students' working memory and provide individualized feedback that would not be possible

in a traditional classroom setting. In addition, this study suggested using ChatGPT as a learning resource to mitigate the challenges of writing argumentative essays, particularly those related to ineffective organization and confusion. Therefore, Su et al. (2023) provide evidence that ChatGPT can serve as an additional resource to support positive outcomes in terms of writing quality and productivity.

On the other hand, another research presents a less favorable outcome. Bašić et al. (2023) studied the impact of ChatGPT on argumentative writing. While researchers found value in its use, it was determined that the group that utilized ChatGPT did not outperform the control group on any of the indicated measures. Thus, using ChatGPT did not produce higher-quality essays, swifter writing, or a more authentic rendering of work. Instead, these shortcomings stemmed from students' unfamiliarity with ChatGPT, which confused them and hindered its usefulness. However, this notion stresses the importance of training when using AI features like ChatGPT for written discourse. Su et al. (2023) and Bašić et al. (2023) present different prospective results in using ChatGPT to relieve struggles associated with writing an argumentative essay. While Su et al. (2023) found great success, Bašić et al. (2023) found no significant success at all. Other studies have emphasized particular benefits, including enhancements in idea development (Alhammad, 2024), language accuracy (Gao, 2024), and the logical coherence and content depth of student writing (Xu et al., 2024). While they assessed ChatGPT for argumentative writing purposes and also examined it for general language proficiency across various EFL contexts, there has yet to be research conducted within an Indonesian EFL setting concerning its specific impact on argumentative writing. Thus, given that argumentative essay writing is a complex genre, and made more challenging in Indonesian context due to the limited academic vocabulary, general rhetorical differences, and varied digital literacy levels, this quasi-experimental study intends to investigate whether ChatGPT as a writing assistant significantly improves this genre's quality when produced by Indonesian undergraduate EFL students and subsequently provides recommendations for pedagogical integration based on findings.

In light of the background and the purpose, this study aims to answer three research questions:

1. Is there a significant difference in the quality of argumentative essay writing between EFL students taught using ChatGPT and those taught using conventional methods?

2. Which aspects of argumentative writing show the most significant and least improvement among students using ChatGPT as a writing assistant, compared to those in the control group?
3. What are students' perceptions of using ChatGPT to assist them in writing an argumentative essay?

METHOD

Research design and subjects

This research aimed to investigate whether the writing quality of EFL students differed between those who utilized ChatGPT as a writing tool and those who did not. A quantitative method was employed with a quasi-experimental design (Creswell, 2014). A quasi-experimental design is suitable for this study due to access to intact classroom groups and naturalistic settings, which preclude the random selection of participants. Although this presents limitations for the study, the pre-test and post-test designs allow for comparable analyses of learning gain within and between groups, thereby enhancing the study's validity. The participants in this study were two classes of EFL students from the English Department of a private university in East Java, Indonesia, who took an Essay Writing course. Students were in their third semester of undergraduate studies, as the Essay Writing course is a required course at that time.

There were two distinct essay types taught in that course: the opinion essay and the argumentative essay. The selection of subjects was purposive, considering their enrollment in the course and prior experiences in writing classes (Tongco, 2007).

To determine which class would serve as the experimental and control groups, the researchers used purposive sampling. In this case, the class with the lower pre-test scores was recommended by the classroom teacher to be assigned as the experimental group. This decision may have been based on the pedagogical consideration that students with lower English proficiency could benefit most from additional instructional support. This approach aligns with a substantial body of research that emphasizes the importance of prioritizing targeted interventions for lower-achieving learners as a critical strategy for reducing performance gaps and promoting educational equity (Nickow et al., 2020; Tomlinson, 2017). Providing such focused support is often seen as an ethical imperative to ensure that all students have the opportunity to succeed (Gorski, 2018). Furthermore, both groups had approximately the same number of students, with 18 in the experimental group and 17 in the control group, totaling 35 students, as detailed in Table 1.

Table 1

Demography of the Research Subjects

Group	Gender	N	Total
Experimental	Female	15	18
	Male	3	
Control	Female	11	17
	Male	6	
Total			35

Research Instruments

Argumentative Essay Writing Prompt

A writing prompt was designed and administered both before (pre-test) and after (post-test) the treatments to both the experimental and control groups. It includes the students' identification information, instructions, a list of topics, and assessment criteria. The instructions provided are designed to guide students in structuring their argumentative essay, detailing the required number of paragraphs and the key elements to include. Students were allotted 100 minutes to complete the task.

Before beginning their essays, students were required to choose one of three provided topics. Moreover, the assessment criteria were included to guide students on what aspects of their writing they should focus on. Additionally, students were asked to write their argumentative essays on the same topic selected during the pre-test. In other words, the

topic for both the pre-test and post-test was the same to ensure a fair comparison of their writing improvement, focusing solely on their development in argumentative skills rather than their familiarity with different topics.

Argumentative Essay Scoring Rubric

A scoring rubric was adapted from the "ACT Writing Test Scoring Rubric" (ACT, 2016) to assess argumentative writing. The developed rubric features a five-point scale to evaluate writing quality across six key aspects: thesis statement, development and support, counterarguments, organization, language use, and vocabulary. The thesis statement aspect evaluates the clarity and relevance of the main argument, as well as the specificity and suitability of the supporting arguments related to the topic. The development and support aspect emphasizes the depth and coherence of idea development, as well as the robustness of

supporting evidence. Counterarguments evaluate the writer's capacity to identify and address opposing perspectives, ensuring effective rebuttal through arguments presented in the thesis statement. The organization aspect assesses the overall structure of the essay. The language use aspect evaluates the writer's proficiency in language, encompassing sentence structure, grammar, and mechanics. The vocabulary aspect emphasizes the variety and accuracy of word selection, assessing the writer's effectiveness in using vocabulary to communicate meaning and sustain a suitable tone throughout the essay.

Closed-ended Questionnaire

Students in the experimental group completed a Likert-scale questionnaire, adapted from Lund (2001) and the Technology Acceptance Model (TAM) framework by Davis (1989), to examine student acceptance of ChatGPT. The items were modified to focus on the essential points that would effectively reveal how ChatGPT influenced students' progress in developing an argumentative essay.

The selection of the questionnaire components was theoretically grounded in the TAM framework, which posits that Perceived Usefulness and Perceived Ease of Use are the primary determinants of technology acceptance. These were supplemented with constructs of *Ease of Learning* and *Satisfaction* to gain a more holistic view of the student experience. In the context of this study, each component is defined as follows:

1. *Usefulness*: Students' perception of how much ChatGPT helped them improve the quality of their argumentative essays.
2. *Ease of Use in Writing Argumentative Essays*: The degree to which students found the process of interacting with ChatGPT for specific writing tasks to be free of effort.
3. *Ease of Learning*: How quickly and easily students felt they could become skillful at using ChatGPT effectively for their academic writing needs.
4. *Satisfaction*: The overall positive or negative feelings students experienced as a result of using ChatGPT as a writing assistant.

The questionnaire measures these four components. Each component is assessed using five response alternatives: "strongly disagree," "disagree," "neutral," "agree," and "strongly agree" with corresponding numerical values of 1, 2, 3, 4, and 5, respectively. This scale was designed to capture students' attitudes regarding the impact of

ChatGPT as a writing assistant on their argumentative essay writing process.

Data Collection

The process of data collection involves two classes: the experimental group, which utilizes ChatGPT as a writing assistant, and the control group, which receives conventional writing instruction without the aid of ChatGPT. For this study, the conventional method is defined as a teacher-centered instructional approach that relies on direct explanation, whole-class and small-group discussions, and peer feedback, without the integration of AI-powered writing tools.

Over the course of nine sessions of the research process, both classes participated in pre- and post-tests assessing their argumentative essay writing skills. The initial session included a pre-test, during which students from both classes wrote an argumentative essay. Subsequently, the treatments spanned sessions 2 to 8. In explaining the use of ChatGPT, the teacher reminded the students to use ChatGPT only as an assistant in writing, not as a tool to generate ideas to copy and paste. The ninth session consisted of a post-test to assess potential differences in students' argumentative essay quality after the treatments, as well as a closed-ended questionnaire distributed online through Google Forms. Each session spanned 100 minutes, consisting of a pre-activity (15 minutes), a while-activity (75 minutes), and a post-activity (10 minutes). The topics of the meetings were tailored to the course profile using Refining Composition Skills (Smalley et al., 2001). In the post-activity for each meeting, students in the experimental group received assistance from ChatGPT and the teacher to provide feedback on their argumentative essay writing; meanwhile, students in the control group received feedback from the teacher and their peers. After all the students received their feedback, they were required to revise their drafts. The final complete draft was submitted to the teacher in the last meeting. Table 2 outlines the classroom activities in which ChatGPT was used in the experimental group, while the conventional method was employed in the control group.

Data Analysis

Several statistical methods were applied to determine the impact of ChatGPT on the quality of students' argumentative essays. These methods included scoring, descriptive statistics, normality testing (Shapiro-Wilk test), paired samples t-test, homogeneity of variance test (Levene's test), and independent samples test. Since the data were normally distributed and homogenous, further parametric analysis was conducted. The descriptive statistics of the analysis results, including the pre-test and post-test means, are presented in the Appendix..

Table 3
Descriptive Statistics

Group	N	Pre-test M (SD)	Post-test M (SD)	Mean Gain	N-Gain
Experimental Group	18	33.72 (7.35)	64.56 (11.82)	30.84	0.46
Control Group	17	34.35 (8.25)	58.65 (10.49)	24.30	0.37

Moreover, the magnitude of the improvements was analyzed to investigate the extent of the changes for each argumentative aspect by measuring the means and gains. In addition, the data obtained from the questionnaire were analyzed using measures of central tendency. To understand the

students' perceptions of using ChatGPT in writing argumentative essays, the responses to the five options provided in the questionnaire items were analyzed based on the values assigned to each option. As displayed in Table 4, the average scores for each item were interpreted (Pimentel, 2010).

Table 4
The Interpretation of the Average Scores

Range	Interpretation
4.21 – 5.00	Very Positive
3.41 – 4.20	Positive
2.61 – 3.40	moderate
1.81 – 2.60	Negative
1.00 – 1.80	Very Negative

FINDINGS

This section presents the study's findings, organized to answer the research questions. First, the quality of students' argumentative essay writing is analyzed through a comparison of pre-test and post-test scores for both experimental and control groups. Second, the magnitude of improvements in specific aspects of argumentative writing is examined to highlight the instructional impact of the intervention. Lastly, students' perceptions of the integration of ChatGPT

into the argumentative essay writing process are explored to provide insights into their experiences and attitudes toward this innovative approach.

Students' Argumentative Essay Writing Quality

To determine whether there was a significant difference between the pre-test and post-test scores of the students in the two groups, as shown in Table 5, paired and independent samples t-tests were carried out.

Table 5
Independent Samples Test of Pre-Test Scores

	Sig. (2-tailed)	Mean Difference	Remark
Equal variances assumed	.812	-.631	Not Significant

Table 5 presents the results of the independent samples test for the pre-test scores of the two groups. The p-value of .812 and the mean difference of -.631 indicate no statistically significant difference between the pre-test scores of the experimental and control groups. This result

confirms that both groups started from a similar baseline level of performance before the interventions were applied. A comparison of the pre-test and post-test results for the two groups is presented in Table 6.

Table 6
Comparison of the Pre-Test and Post-Test of the Two Groups

		Mean	Std. Dev.	Sig. (2-tailed)	Remark
Pair 1 (Experimental)	Pre-test and Post-test	30.83	12.655	.000	Significant
Pair 2 (Control)	Pre-test and Post-test	24.29	10.129	.000	Significant

Table 6 illustrates the paired samples test results comparing the pre-test and post-test scores within the experimental and control groups. The experimental group showed a mean improvement of 30.83 (SD = 12.655), while the control group exhibited a mean improvement of 24.29 (SD = 10.129). The p-values (.000) for both groups indicate statistically significant differences between the pre-test and post-test scores, suggesting that the

instructional interventions had a significant impact on students' performance in both groups.

Further analysis focused on how the students' writing scores improved after the intervention. This was conducted by grouping the students in the experimental group into two groups: high and low writing ability, based on the pre-test scores. The average scores of students with low and high writing abilities, as shown in the pre-test, were compared to

their scores in the post-test. The scores in terms of writing ability showed that students with lower average scores in the pre-test achieved higher

average scores in the post-test than those with higher average scores in the pre-test (see Table 7).

Table 7

Effect of Use of ChatGPT on the Students of Low and High Ability in Writing Based on the Pre-Test Scores

Group of Students	N	Pre-Test M	Post-Test M	Gain
Low Ability in Writing	9	27.33	62.33	35.00
High Ability in Writing	9	40.11	67.88	27.77

Table 7 compares how students with different initial writing abilities responded to the use of ChatGPT. Students in the low-ability group showed greater improvement, with their average scores rising from 27.33 to 62.33, resulting in a gain of 35.00. Meanwhile, students in the high-ability group improved from 40.11 to 67.88, representing a 27.77-point increase. This suggests that ChatGPT provided more noticeable benefits for students who started with lower writing abilities. It establishes the notion that focused intervention is necessary to achieve educational equity (Gorski, 2018). Thus, in this situation, it seems that ChatGPT acted as an excellent technological scaffold. It helps students function within their Zone of Proximal Development and simultaneously address issues they would not be able to address independently (Vygotsky, 1978). Therefore, the success of the lower-ability group not only suggests the power of such a tool but also highlights the importance of focusing interventions on those who would benefit most. If teachers were aware of these distinctions, they would be better equipped to implement more personalized teaching measures. For instance, it

would be recommended to use ChatGPT as a stable tool for lower-ability students in writing, allowing them to develop their fundamental writing skills. For higher-ability students in writing, teachers could promote a more temperate and serious use of ChatGPT, utilizing it to assess ideas, refine arguments, and enhance academic writing style. Furthermore, more intensive use of ChatGPT for structured sessions could help all students learn to utilize the tool better.

Magnitude of Improvements in Argumentative Aspects

The magnitude of improvements in each aspect of argumentative writing for both the experimental and control groups is presented in Table 8. It includes the mean scores and percentage gains for six key aspects: thesis statement, development and support, counterarguments, organization, language use, and vocabulary. The total gains for both groups are also provided, offering a comprehensive view of how each instructional approach impacted specific areas of argumentative writing.

Table 8

Magnitude of Improvements in each Argumentative Writing Aspect

Aspect	Experimental Group		Control Group	
	Mean	% Gain	Mean	% Gain
Thesis Statement	2.12	22.89	1.56	21.31
Development and Support	1.33	14.36	1.12	15.30
Counterarguments	1.69	18.25	1.26	17.21
Organization	1.76	19.00	1.27	17.35
Language Use	1.19	12.85	1.18	16.12
Vocabulary	1.17	12.63	0.93	12.70
Total	9.26	100	7.32	100

The experimental group demonstrated relatively higher gains in argumentation-related aspects, as evidenced by the higher gains in thesis statement (22.89%), counterarguments (18.25%), and organization (19.00%) compared to the control group. Conversely, the control group showed an advantage in language-related aspects, particularly in language use (16.12%), vocabulary (12.70%), and development and support (15.30%).

This finding can be directly attributed to the nature of the intervention in each group. In the experimental group, students were empowered by using ChatGPT to compose an argumentative essay. The students prompted the AI to brainstorm their

pros and cons, generate outlines, and provide examples of logical fallacies. The efficacy of ChatGPT in quickly generating various arguments and associated logical constructs and structures supports the creation of such essay elements as a thesis statement and organization. In contrast, the students in the control group participated in group discussions and received direct teacher feedback, providing them with more opportunities for in-depth language refinement. The teacher and peers provided contextual and nuanced feedback on word choice and sentence structure, which explains why this group excelled more in language use and vocabulary. For example, the improvement in

Student 17's thesis writing is particularly illustrative of the impact of ChatGPT-assisted interventions. In the pre-test, her thesis statement received a score of 2/5. The transition is evident in her writing before and after the interventions.

Before (Pre-test):

In my opinion, monitoring children's online activities by parents does not have to be done secretly because monitoring carried out without the child knowledge can raise ethical issues and damage the relationship of trust between parents and children.

The statement was functional but basic. It expressed a clear opinion, but it lacked engagement with counterarguments, failed to establish logical complexity, and relied on a simple cause-and-effect structure.

After (Post-test):

While parent's concerns about their children's online safety are understandable, covert surveillance can lead to unintended consequences, such as damaged relationship, decreased trust, and stunted emotional development.

The post-test thesis statement received a score of 4/5 for argumentative sophistication. This demonstrates a significant improvement of two points, reflecting the student's ability to construct a nuanced argument. The new thesis acknowledges counterarguments, employs parallelism in listing consequences, and shows logical consistency. The improvement indicates that the student transitioned from producing a one-sided opinion to crafting a more complex, credible, and rhetorically sophisticated argument.

The mechanism for this improvement can be traced to specific feedback from ChatGPT during guided practice. During the intervention, Student 17 practiced with a different topic and produced a flawed thesis as follows: Preserving protected

natural areas is essential to maintain ecological balance, protect fragile ecosystems, and threaten the habitat of endangered species. ChatGPT provided the following specific feedback:

Logical inconsistency: It claims preservation "threatens" endangered species' habitats, which contradicts the intended argument.

Parallelism: The first two verbs ("maintain," "protect") are positive actions, while "threaten" breaks the logical and grammatical flow.

By internalizing this feedback, Student 17 was able to construct the post-test thesis statement that adhered to principles of logical consistency and grammatical parallelism. However, the post-test still contained a minor lexical error: "convert surveillance" instead of the correct term "covert surveillance." This malapropism illustrates that while ChatGPT effectively scaffolded higher-order argumentative skills, it did not fully remediate fine-grained lexical and proofreading issues, aligning with quantitative findings that gains in language use and vocabulary were lower than in argumentation. Overall, Student 17's case demonstrates that ChatGPT's feedback can facilitate substantial growth in thesis sophistication and argumentative reasoning; however, learners may still require additional support to refine their precise language use. The improvement from 2/5 to 4/5 quantifies this progress.

Students' Perceptions on The Integration of ChatGPT in Argumentative Essay Writing

Table 9 provides a general overview of the students' perceptions. Tables 10 to 13 then scrutinize each category: usefulness, ease of use in writing an argumentative essay, ease of learning, and satisfaction, which includes satisfaction in terms of engagement, confidence, and writing autonomy, among others.

Table 9

Overall Analysis of Students' Perceptions

Category	Number of Items	Total Score	Average
Overall Analysis	20	74.22	3.71

The overall analysis implies that students generally view ChatGPT positively in their argumentative writing process. A score of 3.711

suggests that most students neither strongly agree nor disagree with the effectiveness of the tool, indicating a moderately positive reception.

Table 10

Perceived Usefulness of ChatGPT

No.	Item	Average
1	ChatGPT is useful	4.06
2	ChatGPT gives me more control over the writing activities	3.67
3	ChatGPT makes the writing process easier to get done	3.89
4	ChatGPT saves me time when I use it	3.83
5	ChatGPT meets my needs	3.78
Total Score		19.22
$\bar{X}$		3.84

As depicted in Table 10, the average score of 3.84 in perceived usefulness indicates that students consider ChatGPT relatively effective in supporting their writing activities, particularly in crafting argumentative essays. The questionnaire focused on the role of ChatGPT as a tool to facilitate the writing process, such as saving time, simplifying tasks, and

meeting their needs, rather than directly assessing its effectiveness in improving the quality of their essays. The score being closer to 4.0 than to the neutral midpoint (3.0) highlights students' positive perceptions of ChatGPT as an aid in managing and completing their writing tasks.

Table 11

Perceived Ease of Use in Writing Argumentative Essays with ChatGPT

No.	Item	Average
1	The suggestions provided by ChatGPT helped me structure my argumentative essay	3.67
2	I found ChatGPT easy to use when writing my argumentative essay	3.61
3	The suggestions from ChatGPT helped me develop strong evidence in my argumentative essay	3.56
4	I found it easy to understand the feedback from ChatGPT	3.61
5	I found it easy to incorporate the feedback from ChatGPT into my writing	3.61
Total Score		18.06
$\bar{X}$		3.61

Table 11 shows that all the students have a positive perception of the ease of use in writing argumentative essays with ChatGPT. This is indicated by the average score of 3.61. In other words, the students value how ChatGPT helped

them structure their argumentative essay and develop strong evidence to support their argument. They also found that ChatGPT can be used efficiently, and its feedback is easy to understand and to integrate into writing.

Table 12

Perceived Ease of Learning

No.	Item	Average
1	I found it easy to learn how to use ChatGPT	3.94
2	I quickly became comfortable using ChatGPT for my writing needs	3.67
3	I can remember easily how to use ChatGPT	4.00
4	I was able to understand the functionalities of ChatGPT without difficulty	3.89
5	Learning to interact with ChatGPT effectively enhanced my writing experience	3.78
Total Score		19.28
$\bar{X}$		3.86

The highest average score of 3.86, as shown in Table 12, indicates that students generally find it easy to learn ChatGPT. This suggests that the ChatGPT interface is not difficult to master.

Moreover, a score of 4.00 suggests that the tool is not only easier to learn but also has the potential to lead students to high-quality writing.

Table 13

Students' Satisfaction with ChatGPT in Assisting Argumentative Essay Writing

No.	Item	Average
1	I am satisfied with the quality of responses provided by ChatGPT when writing my argumentative essay	3.72
2	ChatGPT provided reliable suggestions that I could depend on in my argumentative writing	3.67
3	I felt confident using ChatGPT as a resource for my argumentative essay	3.39
4	I believe my writing skills have improved because of the assistance from ChatGPT	3.50
5	I felt that I could not effectively write an argumentative essay without the help of ChatGPT	3.39
Total Score		17.67
$\bar{X}$		3.53

Table 13 reflects a positive satisfaction yet cautious response with an average score of 3.53. The fact that this score is the lowest among the other aspects implies certain limitations. Specifically, the lower scores on the *confidence* item (3.39) and the item measuring dependency (3.39) highlight a dilemma: while the tool is helpful, some students worry about becoming overly reliant on it.

While the quantitative data indicate benefits, the perceptual data also highlight challenges

students faced. The high satisfaction scores, especially regarding confidence, suggest that students did not readily accept AI suggestions. To address the potential for direct copying (copy-paste) and to ensure ChatGPT was used as an assistant, the pedagogical intervention employed was crucial. The teacher explicitly framed ChatGPT as a brainstorming tool, not a content writer. During the writing process, the experimental group was required to revise their drafts based on feedback

from ChatGPT, as well as from peers, similar to the control group. To facilitate this critical engagement, students had to present orally, in front of their peers, the prompts they selected and explain how they either adjusted or negated the response from the AI in their final draft. Thus, students were empowered to critically assess AI suggestions, make decisions independently regarding writing, and avoid the potential pitfalls of overdependence.

DISCUSSION

The study examined the impact of integrating ChatGPT on the argumentative writing quality of EFL students. The findings from the data analysis contribute valuable knowledge to the existing literature on the use of ChatGPT as a writing assistant tool. To begin with, the statistical analysis lends weight to the research findings of Bašić et al. (2023), which indicate that the argumentative essays from students in the experimental group did not show a significant effect in terms of content quality and did not outperform those in the control group. However, an important distinction to note is that the previous study did not involve pre- and post-test assessments; instead, their analysis was based on a single task, which limits comparisons over time. Still, the use of ChatGPT conveyed an advantageous impact on improving the quality of academic writing (Mahapatra, 2024; Yan, 2023).

When examining each aspect of argumentative writing, both the experimental and control groups showed improvements across all aspects, although the gains differed. Aspects directly related to argumentation—such as the thesis statement, organization, and counterarguments—showed the most improvement, with the experimental group achieving higher scores in these areas. This suggests that ChatGPT helped students become more familiar with the structure of an argumentative essay, indicating that teachers can utilize ChatGPT as a writing assistant during the drafting stage to help students internalize essay structures rather than merely memorizing them superficially. This chatbot enabled students to improve several aspects of their writing process. For instance, the experimental group demonstrated improvements in brainstorming and ideation abilities (Atlas, 2023; Petraki, 2024), was taught to organize their writing effectively (Marzuki et al., 2023), and showed better ability to acknowledge and respond to counterclaims in their work (Su et al., 2023). The control group, however, performed better on components related to language, including language use, vocabulary, and development and support. While development and support are significant aspects of effective argumentation, they are still reliant on language, as students need to generate well-developed body paragraphs to support their claims. Without a sufficient vocabulary and a solid understanding of

language, it becomes challenging for students to elaborate on and effectively support their ideas. This may explain why the control group performed better in the development and support aspects, underscoring the need for teachers to integrate language-focused instruction explicitly alongside the use of AI.

One of the benefits of this chatbot is its capability to assist students in brainstorming ideas to develop well-reasoned arguments. In this capacity, the students were able to craft strong thesis statements that effectively presented their arguments on the given topics. Furthermore, ChatGPT could also play devil's advocate, thus allowing students to not only learn their own perspectives (Farrokhnia et al., 2024) but also to understand both their own arguments and counterarguments. This is a crucial component of the argumentative writing process. Additionally, it is recognized for generating accurate, specific, thorough, and coherent responses (Susnjak, 2022). Its answers often include relevant examples, detailed explanations, and clear organization, incorporating transitions and a conclusion. This response format not only aids in understanding but also reinforces students' memory of essay structures, helping them apply these organizational principles to their own writing. On the contrary, some students in the control group struggled to maintain a consistent structure in their essays. Their arguments were often scattered, lacked clear topic sentences, and omitted conclusions altogether, highlighting that integrating AI tools like ChatGPT with explicit teacher modeling of essay organization can accelerate students' ability to produce cohesive and logically ordered essays.

Moreover, Susnjak (2022) analyzed responses from ChatGPT and highlighted their clear responses. The AI provides replies that are easy to understand, well-organized, coherent, and grammatically accurate. ChatGPT also provides appropriate vocabulary to meet expectations for natural language responses (Farrokhnia et al., 2024; Wang et al., 2023) as well as personalized feedback (Farrokhnia et al., 2024; Shen et al., 2023). Such functionality enables ChatGPT to offer seemingly accurate and increasingly reliable answers. Yet, research reveals that many students treat AI as a generator of content rather than a collaborative learning partner (Liu et al., 2023; Tu, 2024). For example, in the current research study, some participants confessed to using ChatGPT-generated responses without fact-checking or rewording them in their own voice. This is a reflective learning response to a larger phenomenon in which academic writing continues to employ AI-generated tools to create concepts, draft, or edit essays, despite ethical concerns. This phenomenon of over-reliance is found in the experimental group, who scored lower than the control group in terms of language. Rather than integrating ChatGPT as a foundation for self-

directed learning, students often take its responses at face value. They miss opportunities to learn from how ChatGPT refines their grammar or improves their text.

Additionally, ChatGPT can serve as a valuable source of feedback (Zhai, 2023), helping students develop ideas and refine their argumentative writing. (Kumar, 2023) explains that ChatGPT gives quick answers for academic writing, making it easier for students to collect information (Farrokhnia et al., 2024). However, outputs from this chatbot are not always accurate (Tlili et al., 2023). Based on the research results and observations, students tend to depend heavily on ChatGPT, accepting its outputs without double-checking or analyzing them further. This raises the risk of unintentional plagiarism, primarily when students reproduce AI-generated text verbatim without attribution or revision. If students could utilize ChatGPT more effectively—viewing it as a learning aid rather than a substitute for their own effort—they would likely achieve better outcomes in language learning (Cotton et al., 2024; Zhai, 2023).

A significant concern, however, is that this AI may foster dependence and distract students from learning (Bašić et al., 2023). Students may not need to analyze or contemplate information if it is provided without challenge, and thus, they run the risk of becoming lazy thinkers who cannot discern because everything is done for them. Therefore, they will undoubtedly struggle with assessments that require independent growth and creativity or a clear and deep understanding of the project at hand (Bašić et al., 2023; Farrokhnia et al., 2024; J. Gao et al., 2023). The ease with which ChatGPT responds to prompts may discourage students from conducting their own research and exercising their critical thinking skills (Kasneci et al., 2023; Liu et al., 2023). Therefore, EFL educators must holistically approach AI literacy and ethical writing practices by teaching students not only how to paraphrase suggestions but also how to fact-check accuracy and engage critically before accepting AI feedback unquestioningly. Researchers note that although ChatGPT helps meet complex learning outcomes, it is less effective in fundamental learning tasks requiring higher-order thinking (Rudolph et al., 2023).

Furthermore, the questionnaire data revealed that students expressed a generally positive perception of the impact of ChatGPT on their writing skills. This positive attitude was evident across the four evaluated categories: usefulness, ease of use in writing argumentative essays, ease of learning, and overall satisfaction. Students have determined that ChatGPT is a helpful tool for learning how to write, as it makes the process easier and meets their need for efficiency and effectiveness (Yan, 2023). Therefore, with the assistance of

ChatGPT, students were more equipped to try this writing task, and they anticipated such assistance because essays of the argument are judged to be complex writing tasks (Aertselaer & Dafouz-Milne, 2023; Dornbrack & Dixon, 2014; Latifi et al., 2021; Noroozi et al., 2023). Furthermore, ease of use emerged as one of the most critical categories praised by students, primarily for the assistance in helping them format their essays. This aligns with the findings that show the experimental group outperformed the control group in the organization category. The organizational suggestions from ChatGPT were essential in helping students order their ideas in a clear and logical order (Zhao et al., 2024).

Additionally, another significant finding was that ChatGPT aided students in their critical thinking efforts during the pre-writing stage. Students received varying answers to a posed question and weighed the advantages and disadvantages of a specific issue (Farrokhnia et al., 2024). Exposure to differing viewpoints better enabled students to consider what would serve well as their own arguments or potential counterarguments. Coupled with collaborative group activities, this allows students to engage meaningfully with the information provided by ChatGPT, thereby formulating their argumentative essay more successfully. This resulted in higher scores in critical categories of argumentation, including thesis statements and counterarguments. These two components differentiate an argument from a point of view piece. For example, the inclusion of counterarguments and their discussion demonstrates an awareness of differing opinions, which not only reinforces one's argument but also elevates the argument itself.

Recognizing the benefits, however, ChatGPT did not outperform traditional methods in every component of argument writing. For instance, the control group excelled with the development and support component, which suggests they were better able to find relevant examples to support their claims. This finding indicates that while ChatGPT excels at providing scaffolding for brainstorming and organizing (Petraki, 2024), it does not fully grasp the nuances necessary for analytical consideration to develop claims with supporting evidence. This could be because teacher-led discussion in a control group afforded students more opportunities to engage and critically assess with peers and the instructor (Lam et al., 2018). As such, this interaction was productive in helping them better understand how to choose and apply relevant evidence. Conversely, using ChatGPT provided students with a plethora of information, yet in a general sense, it did not enable students to ground their ideas in specific examples or nuanced information more suitable for the argument (Kumar, 2023). Thus, these findings suggest that ChatGPT

serves as a supplementary resource for learning how to write an argument. While it successfully addresses construction and organization components, its use should be in addition to teacher-led and peer-generated facilitated discussions for a holistic approach to learning the skill (Mahapatra, 2024). Teachers must train students to assess and critically dispute AI's offerings based on specific needs for argumentation. In other words, the AI and the inclusivity stand to bring a lot to the table. While this research aimed to determine the efficiency of using AI versus traditional means, it is essential to note that numerous extraneous factors can impact students' learning. For example, what they know, how they reinforce what they have learned, and their motivation are all intrapersonal factors that significantly influence how any student will approach a task. Students with greater prior knowledge may have an easier time agreeing or disagreeing with prompts provided by ChatGPT or a teacher-led discussion, as they can better situate their new findings within pre-existing notions (Shloul et al., 2024). Moreover, the extent to which students review and revisit materials taught in class for reinforcement impacts their ability to generalize these ideas to a writing assignment (Kinyanjui et al., 2015). Finally, motivation is critical. More motivated students will likely spend more time polishing their arguments than those who are not motivated, regardless of the tool or method used (Alzubi & Nazim, 2024; Duță, 2015).

Additionally, one of the most substantial benefits of using ChatGPT is its intuitive user interface (UI) and engaging user experience (UX). All students agreed that the ChatGPT layout facilitated easy learning, allowing them to quickly learn how to navigate the chatbot and understand its features, thereby grasping what it had to offer them. This contributed to a sense of convenience and enhanced their overall writing experience. This is a significant finding that notes usability in educational technologies goes a long way toward improving student engagement and productivity, particularly when accompanied by usable UI/UX designs (Miya & Govender, 2022; Pandita & Kiran, 2023). Additionally, students were generally satisfied with the response quality that ChatGPT generated; they viewed it as reliable, albeit fallible. However, this perspective offers encouraging prospects for future use, indicating that more precise prompt adjustments could enhance accuracy. Xiao and Zhi (2023) state that the more detailed the prompt, the higher the quality of the response. Therefore, students must learn how to develop precise prompts. A highlighted experience in prompt development positively reinforces critical thinking skills (Guo & Lee, 2023). This approach encourages students to engage actively in the writing process. ChatGPT should only be used as a tool to stimulate their ideas rather than as a substitute for independent learning. Lastly,

many students reported increased confidence in their writing abilities due to ChatGPT, as evidenced by higher scores and their positive responses on questionnaires, which indicated that they appreciated this new feature. Yet some students became codependent on the feature, worrying that they could no longer write without it. This codependency suggests a downside of AI integration, as some students may prioritize convenience over skill development (Liu et al., 2023). For many EFL students, the biggest obstacle is language ability. Students noted difficulties with vocabulary, grammar, and structural problems—students had issues with subject-verb agreement and spelling. Thus, while ChatGPT provided commentary on linguistic concerns, many students did not engage with the proposed changes, resulting in similar mistakes being made. The need to encourage student awareness for independent learning was evident from this study.

CONCLUSION

ChatGPT can assist EFL students in enhancing the quality of their argumentative essays, implying that this AI should be considered a complementary tool rather than a replacement. Also, this study highlights the potential of ChatGPT to strengthen argumentation aspects. Additionally, ChatGPT is quite easy for students to learn, making it a suitable option even for those new to AI-based writing tools. Students acknowledge ChatGPT as beneficial for the writing process, as it saves time; however, they also note some difficulties in processing due to the ease of use within this writing process. This means that although ChatGPT can help at various stages of writing an argumentative essay, it has its limits, especially when it comes to formulating strong arguments. The information and data it gave writers were often too vague for them to expand upon. In addition, the moderate level of satisfaction also suggests that while ChatGPT is practical and helpful, it may be lacking certain upgrades that would enable students to complete higher-level work more effectively.

Educators can offer a range of solutions to the challenges that students encounter when composing their own argumentative essays. First, continuous feedback from a peer review system or annotated drafts allows students to rework their thoughts and structures. Second, lessons designed for multiple drafts and revisions remind students of the importance of revision in achieving high-quality essays. Third, lessons on time management help students accommodate the needs for all stages of the writing process. Fourth, scaffolded instruction that focuses on single elements—such as thesis construction, argument development, and counterargument inclusion—before blending them into a full essay leads to a clearer understanding.

Moreover, when students are allowed to reflect upon their mistakes through feedback and proper corrections are offered as enhancements for clear expectations, they adopt a more beneficial mindset. In addition, when assignments require higher-order thinking and creativity—such as oral presentations or tangible projects—students will be more engaged, as they'll be contributors instead of mere receivers. In doing so, they also develop essential soft skills, such as communication and collaboration. Thus, increased engagement not only helps students learn the material more effectively but also equips them with transferable skills that benefit them outside the classroom. Lastly, implementing formative evaluations, such as reflective journals, class discussions, or portfolios, provides a comprehensive picture of how students are performing. It focuses on their growth during the learning process, rather than just their outcomes.

ACKNOWLEDGEMENTS

The researchers would like to express their deepest gratitude to the Master's to Doctoral Education for Excellent Graduates (*Pendidikan Magister menuju Doktor untuk Sarjana Unggul/PMDSU*) Program and the Directorate of Research, Technology, and Community Service (*DRTPM*) of the Republic of Indonesia for funding this research. This research was funded under Decree Number 0459/E5/PG.02.00/2024 and Research Contract Number 066/E5/PG.02.00.PL/2024.

REFERENCES

- Abramson, A. (2023, June). *How to use ChatGPT as a learning tool*. American Psychological Association.
<https://www.apa.org/monitor/2023/06/chatgpt-learning-tool>
- ACT. (2016). *The ACT writing test scoring rubric*. ACT.
<https://www.act.org/content/dam/act/unsecured/documents/Writing-Test-Scoring-Rubric.pdf>
- Aertselaer, J. N., & Dafouz-Milne, E. (2008). Argumentation patterns in different languages: An analysis of metadiscourse markers in English and Spanish texts. In M. Pütz & J. N. Aertselaer (Eds.), *Developing contrastive pragmatics* (pp. 87–102). Mouton de Gruyter.
<https://doi.org/doi:10.1515/9783110207217.1.87>
- Alhammad, A. I. (2024). The impact of ChatGPT in developing Saudi EFL learners' literature appreciation. *World Journal of English Language, 14*(2), 331–338.
<https://doi.org/10.5430/wjel.v14n2p331>
- Alzubi, A. A. F., & Nazim, M. (2024). Students' intrinsic motivation in EFL academic writing: Topic-based interest in focus. *Heliyon, 10*(1), e24169.
<https://doi.org/10.1016/j.heliyon.2024.e24169>
- Atlas, S. (2023). *ChatGPT for higher education and professional development: A guide to conversational AI*. DigitalCommons@URI, University of Rhode Island.
https://digitalcommons.uri.edu/cba_facpubs/548
- Bašić, Ž., Banovac, A., Kružić, I., & Jerković, I. (2023). ChatGPT-3.5 as writing assistance in students' essays. *Humanities and Social Sciences Communications, 10*, 750.
<https://doi.org/10.1057/s41599-023-02269-7>
- Baskara, F. R. (2023). Integrating ChatGPT into EFL writing instruction: Benefits and challenges. *International Journal of Education and Learning, 5*(1), 44–55.
<https://doi.org/10.31763/ijelev.v5i1.858>
- Chanie, B. S. (2013). Assessing pre-engineering students' writing errors at Bahir Dar University, Ethiopia. *Journal of Media and Communication Studies, 5*(3), 20–24.
<https://doi.org/10.5897/JMCS2013.0350>
- 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. (2014). *Research design: Qualitative, quantitative and mixed methods approaches* (4th ed.). Sage.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.
<https://doi.org/10.2307/249008>
- Dornbrack, J., & Dixon, K. (2014). Towards a more explicit writing pedagogy: The complexity of teaching argumentative writing. *Reading & Writing, 5*(1), 1–8.
<https://doi.org/10.4102/rw.v5i1.40>
- Duță, N. (2015). Importance of the motivation of students for learning - Premise for academic performances. Inventory of reasons for which students learn. *Euromentor Journal, 6*(2), 54–72.
<https://www.proquest.com/docview/1695791924?sourcetype=Scholarly%20Journals>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International, 61*(3), 460–474.
<https://doi.org/10.1080/14703297.2023.2195846>
- Gao, J., Zhao, H., Yu, C., & Xu, R. (2023). *Exploring the feasibility of ChatGPT for event extraction*.
<https://doi.org/10.48550/arXiv.2303.03836>

- Gao, X. (2024). The impact of AI on the accuracy of second language learners in acquiring non-finite verb structures. *Communications in Humanities Research, 61*, 1–6. <https://doi.org/10.54254/2753-7064/2025.18139>
- Gorski, P. C. (2018). *Reaching and teaching students in poverty: Strategies for erasing the opportunity gap* (2nd ed.). Teachers College Press.
- Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for enhancing critical thinking skills. *Journal of Chemical Education, 100*(12), 4876–4883. <https://doi.org/10.1021/acs.jchemed.3c00505>
- Hillocks, G. (2009). *Teaching argument writing, grades 6-12*. Heinemann.
- 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*, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kim, W. G. (2024). A study on prompting for argumentative writing in the AI era: Focusing on ChatGPT writing instruction methods utilizing metacognition. *The Korean Journal of Literacy Research, 15*(3), 193–223. <https://doi.org/10.37736/KJLR.2024.06.15.3.07>
- Kinyanjui, M. W., Aloka, P., Mutisya, S. K., Ndeke, F. N., & Nyang'ara, N. M. (2015). Classroom instruction reinforcement strategies and factors that influence their implementation in Kenyan primary schools. *Journal of Educational and Social Research, 5*(3), 267–278. <https://doi.org/10.5901/jesr.2015.v5n3p267>
- Kumar, A. H. S. (2023). Analysis of ChatGPT tool to assess the potential of its utility for academic writing in Biomedical domain. *BEMS Reports, 9*(1), 24–30. <https://doi.org/10.5530/bems.9.1.5>
- Lam, Y. W., Hew, K. F., & Chiu, K. F. (2018). Improving argumentative writing: Effects of a blended learning approach and gamification. *Language Learning & Technology, 22*(1), 97–118. <https://doi.org/10.1111/bjet.13054>
- Latifi, S., Noroozi, O., & Talaei, E. (2021). Peer feedback or peer feedforward? Enhancing students' argumentative peer learning processes and outcomes. *British Journal of Educational Technology, 52*(2), 768–784. <https://doi.org/10.1111/bjet.13054>
- Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction, 11*(4–5), 357–380. [https://doi.org/10.1016/S0959-4752\(00\)00037-2](https://doi.org/10.1016/S0959-4752(00)00037-2)
- Liu, M., Ren, Y., Nyagoga, L. M., Stonier, F., Wu, Z., & Yu, L. (2023). Future of education in the era of generative artificial intelligence: Consensus among Chinese scholars on applications of ChatGPT in schools. *Future in Educational Research, 1*(1), 72–101. <https://doi.org/10.1002/fer3.10>
- Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability Interface, 8*(2), 3–6. https://www.researchgate.net/publication/230786746_Measuring_Usability_with_the_USE_Questionnaire
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments, 11*(9), 1–18. <https://doi.org/10.1186/s40561-024-00295-9>
- 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), 1–17. <https://doi.org/10.1080/2331186X.2023.2236469>
- Miao, F., & Holmes, W. (2021). *International forum on AI and the futures of education, developing competencies for the AI era, 7-8 December 2020: Synthesis report*. United Nations Educational, Scientific and Cultural Organization. <https://unesdoc.unesco.org/ark:/48223/pf0000377251>
- Miya, T. K., & Govender, I. (2022). UX/UI design of online learning platforms and their impact on learning: A review. *Research in Business & Social Science, 11*(10), 316–327. <https://doi.org/10.20525/ijrbs.v11i10.2236>
- Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly, 46*(3), 273–304. <https://doi.org/10.1598/rrq.46.3.4>
- Nickow, A. J., Oreopoulos, P., & Quan, V. (2020). *The impressive effects of tutoring on pre-k-12 learning: A systematic review and meta-analysis of the experimental evidence* (NBER Working Paper No. 27476). National Bureau of Economic Research. <https://doi.org/10.3386/w27476>
- Noroozi, O., Banihashem, S. K., Biemans, H. J. A., Smits, M., Vervoort, M. T. W., & Verbaan, C.-L. (2023). Design, implementation, and evaluation of an online supported peer feedback module to enhance students' argumentative essay quality. *Education and*

- Information Technologies*, 28(10), 12757–12784. <https://doi.org/10.1007/s10639-023-11683-y>
- Pandita, A., & Kiran, R. (2023). The technology interface and student engagement are significant stimuli in sustainable student satisfaction. *Sustainability*, 15(10), 7923. <https://doi.org/10.3390/su15107923>
- Petraki, E. (2024). Exploring the integration of generative AI in assessment in a tertiary context: A case study. In T. Cochrane, V. Narayan, E. Bone, C. Deneen, M. Saligari, K. Tregloan, & R. Vanderburg (Eds.), *Proceedings ASCILITE 2024* (pp. 308–318). <https://doi.org/10.14742/apubs.2024.1537>
- Pimentel, J. L. (2010). A note on the usage of Likert scaling for research data analysis. *USM R&D Journal*, 18(2), 109–112. https://www.researchgate.net/publication/331231816_A_note_on_the_usage_of_Likert_Scaling_for_research_data_analysis
- Ranjbaran, F., Babaei, M., Khaneh, M. P. A., Gohari, M., Ghorbani, B. D., Kerman, N. T., Banihashem, S. K., & Noroozi, O. (2023). Students' argumentation performance in online learning environments: Bridging culture and gender. *International Journal of Technology in Education*, 6(3), 434–454. <https://doi.org/10.46328/ijte.460>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, 6(1), 242–263. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Shen, Y., Heacock, L., Elias, J., Hentel, K. D., Reig, B., Shih, G., & Moy, L. (2023). ChatGPT and other large language models are double-edged swords. *Radiology*, 307(2), 1–4. <https://doi.org/10.1148/radiol.230163>
- Shloul, T. A., Mazhar, T., Abbas, Q., Iqbal, M., Ghadi, Y. Y., Shahzad, T., Mallek, F., & Hamam, H. (2024). Role of activity-based learning and ChatGPT on students' performance in education. *Computers and Education: Artificial Intelligence*, 6, 100219, 1–18. <https://doi.org/10.1016/j.caeai.2024.100219>
- Smalley, R. L., Ruetten, M. K., & Kozyrev, J. (1995). *Refining composition skills: Rhetoric and grammar*. Heinle & Heinle.
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, 100752. <https://doi.org/10.1016/j.asw.2023.100752>
- Susnjak, T. (2022). ChatGPT: The end of online exam integrity? *ArXiv*. <https://doi.org/10.48550/arXiv.2212.09292>
- Tandiana, S. T., Abdullah, F., & Komara, U. (2017). Digital writing tools: Teaching argumentative essays beyond the traditional frontiers. *Proceedings of the 10th CONAPLIN and ICOLLITE Conferences*, 1, 336–343. <https://doi.org/10.5220/0007167003360343>
- 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), 1–24. <https://doi.org/10.1186/s40561-023-00237-x>
- Toan, L. A., Dan, T. C., Da, T. C., Chau, P. T. H., Ngoan, C. T., Thien, M. B., & Anh, P. H. D. (2020). A study about EFL English-major students' challenges in writing argumentative essays at Soc Trang Teachers' Training College, Vietnam. *RSU International Research Conference*, 1544–1558. <https://doi.org/10.14458/RSU.res.2020.10>
- Tomlinson, C. A. (2017). *How to differentiate instruction in academically diverse classrooms* (3rd ed.). ASCD.
- Tongco, M. D. C. (2007). Purposive sampling as a tool for informant selection. *Ethnobotany Research and Applications*, 5, 147–158. <https://ethnobotanyjournal.org/index.php/era/article/view/126>
- Tu, D. H. M. (2024). Navigating the convenience rap with ChatGPT and Google Translate: The risks of homogenization in translation teaching for Vietnamese students. *International Journal of Linguistics, Literature and Translation*, 7(8), 244–255. <https://doi.org/10.32996/ijllt.2024.7.8.28>
- Vovk, O., & Kryvosyia, D. (2024). Pros and cons of using ChatGPT in academic writing instruction. *Bulletin of the Cherkasy Bohdan Khmelnytsky National University. Series Pedagogical Sciences*, 1(3), 52–60. <https://doi.org/10.31651/2524-2660-2024-3-52-60>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wang, F.-Y., Miao, Q., Li, X., Wang, X., & Lin, Y. (2023). What does ChatGPT say: The DAO from algorithmic intelligence to linguistic intelligence. *IEEE/CAA Journal of Automatica Sinica*, 10(3), 575–579. <https://doi.org/10.1109/JAS.2023.123486>
- Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, 8(3), 212. <https://doi.org/10.3390/languages8030212>

Xu, J., Gao, J., Gong, X., Shadiev, R., & Li, Y. (2024). The impact of using ChatGPT on EFL students' writing. *IEEE International Conference on Advanced Learning Technologies (ICALT)*, 126–128. <https://doi.org/10.1109/ICALT61570.2024.00043>

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 28(11), 13943–13967. <https://doi.org/10.1007/s10639-023-11742-4>

Zhai, X. (2023). ChatGPT user experience: Implications for education. *SSRN 4312418*. <https://doi.org/10.2139/ssrn.4312418>

Zhao, T., Chance, R., Buckhalter, C., & Wang, G. (2024). Impact of ChatGPT on student writing in construction management: Analyzing literature and countermeasures for Writing Intensive courses. *Proceedings of 60th Annual Associated Schools of Construction International Conference Impact*, 5, 339–348. <https://doi.org/10.29007/4d7m>

APPENDIX

Classroom Procedure

Stage (Time Allocation)	
Experimental Group	Control Group
Meeting 1 – Pre-Test	
Meeting 2 – Introduction to Argumentative Writing	
Meeting 3 – Explaining the organization of an argumentative essay using model essays	
Meeting 4 – Brainstorming	
The teacher introduced brainstorming techniques. Students then used ChatGPT with specific prompts (e.g., “Provide three arguments for the topic [topic name]” or “What are some counterarguments against [a specific claim]?”) to generate ideas, evaluate multiple perspectives, and refine their thesis.	The teacher facilitated a small-group discussion in which students brainstormed arguments collectively and selected thesis ideas, providing feedback and examples throughout the session.
Meeting 5 – Introduction Development	
The teacher explained the structure of an introductory paragraph. Students then used ChatGPT with prompts like “Generate an engaging hook for an essay about [topic name]” to explore suggestions for crafting hooks and transitioning into their thesis.	The teacher explained the components of the introductory paragraph, provided examples of hooks and background information, and explained their functions, guiding students step-by-step in writing their introductions.
Meetings 6 and 7 – Body Paragraphs	
The teacher reviewed the structure of body paragraphs. Students then used ChatGPT to organize ideas and incorporate counterarguments using prompts such as “Help me structure a paragraph for the argument that [argument]” or “How can I effectively rebut the counterargument that [counterargument]?”.	The teacher explained the structure of body paragraphs, provided examples of justifications and counterarguments, and encouraged peer collaboration to develop and refine arguments.
Meeting 8 – Conclusion Writing	
The teacher delivered a lesson on writing effective conclusions. Students used ChatGPT to rephrase their thesis statements and craft impactful closing statements with prompts like “Rephrase this thesis in a new way: [original thesis]” or “Suggest a final thought for an essay about [topic name]”.	The teacher guided students in restating their thesis, discussed strategies for writing a compelling conclusion, and provided feedback on their drafts during a plenary session.
Meeting 9 – Post-Test	