

AI-Mediated Collaborative Writing: Evidence for Enhanced Engagement and Writing Performance in EFL Learning

¹Dina Salma, ¹*Lusia Maryani Silitonga, ¹Arso Setyaji

¹Master of English Education, Universitas Persatuan Guru Republik Indonesia Semarang (UPGRIS), Jl. Sidodadi Timur No. 24, Dr. Cipto, Semarang, Indonesia

*Corresponding Author e-mail: lusiamaryanisilitonga@upgris.ac.id

Received: April 2026; Revised: May 2026; Accepted: May 2026; Published: June 2026

Abstract

Ninth-grade English as a Foreign Language (EFL) learners encounter significant difficulties in developing ideas into coherent paragraphs, maintaining grammatical accuracy, particularly consistent sentence structure, and selecting contextually appropriate vocabulary. These challenges are often accompanied by low writing confidence and limited active participation, especially in collaborative learning settings, which hinder the development of effective written communication. Despite the growing integration of technology in language learning, empirical evidence regarding the effectiveness of AI-assisted tools in overcoming these challenges remains limited, particularly at the secondary level in Indonesian contexts. This study aimed to investigate the effect of AI-supported collaborative writing through ChatGPT on students' engagement and writing performance. This study employed a mixed-methods approach with a quasi-experimental design involving 72 ninth-grade students from SMPIT Az-Zahra Demak, divided into an experimental group (n = 36) that used ChatGPT and a control group (n = 36) that used conventional instruction. Data were collected through writing tests, engagement questionnaires measuring behavioral, emotional, and cognitive dimensions, and semi-structured interviews. Quantitative data were analyzed using ANCOVA, while qualitative findings were analyzed through thematic analysis using MAXQDA. The findings demonstrated that the experimental group consistently outperformed the control group across all dimensions of engagement, with large effect sizes observed. In terms of writing performance, significant improvements were found, particularly in text organization and language use. These findings indicate that AI-supported collaborative writing can effectively enhance both student engagement and writing performance. Therefore, integrating AI tools such as ChatGPT into collaborative writing activities offers a promising pedagogical approach for EFL instruction at the secondary school level.

Keywords: Artificial intelligences; Collaborative writing; Students' engagement; Writing performance

How to Cite: Salma, D., Silitonga, L.M., & Setyaji, A. (2026). AI-Mediated Collaborative Writing: Evidence for Enhanced Engagement and Writing Performance in EFL Learning. *Journal of Language and Literature Studies*, 6(2), 727-739. doi: <https://doi.org/10.36312/jolls.v6i2.4975>



<https://doi.org/10.36312/jolls.v6i2.4975>

Copyright© 2026, Salma et al.

This is an open-access article under the CC-BY-SA License.



INTRODUCTION

The increasing presence of digital technologies has reshaped how teaching and learning are conducted in contemporary classrooms, moving conventional teacher-centered instruction toward more interactive and student-centered learning environments. (Allayarova, 2025; Wang et al., 2024). This transformation is closely associated with the demands of 21st-century skills, which emphasize critical thinking, collaboration, creativity, and digital literacy (Allayarova, 2025; Angguni et al., 2026). In this context, learning is no longer viewed as a passive process of knowledge transmission but rather as an active and constructive process in which learners engage meaningfully with content, peers, and digital tools (Swai, 2025). Consequently, educational systems are increasingly

required to adopt innovative technologies that can support personalized, flexible, and engaging learning experiences (Djibran et al., 2024; OECD, 2023).

In line with these developments, artificial intelligence (AI) has become a pivotal innovation in both technological advancement and educational practices. AI enables machines to simulate human behavior, adapt to new inputs, and learn from prior experiences, making it increasingly relevant across various sectors, particularly in education. In the rapid growth of technology, the integration of AI into the educational context has become increasingly unavoidable (King, 2023). This trend is further reinforced by the national digitalization agenda promoted by Kemendikbudristek RI, 2022), which emphasizes the use of digital platforms to enhance educational quality and promote meaningful learning in the classroom. Similarly, UNESCO (2023) highlights the significance of integrating generative AI in ethical and pedagogically appropriate ways to support personalized and effective learning experiences.

Among various AI tools, ChatGPT has attracted considerable attention because of its potential to support teaching and learning processes. Previous studies have demonstrated that ChatGPT can assist in managing instructional resources and enhance learning efficiency (Yu, 2023). Furthermore, it encourages learner autonomy through personalized, interactive assistance tailored to students' individual needs (Baidoo-Anu & Owusu Ansah, 2023). As an instructional assistant, ChatGPT can offer feedback, guide students through learning materials, and maintain their engagement throughout the learning process (Amin, 2023). In academic writing contexts, existing studies suggest that ChatGPT facilitates idea generation, organization of ideas, and overall writing performance (Bibi & Atta, 2024; Imran & Almusharraf, 2023).

However, most of these studies primarily focused on the individual use of ChatGPT as a learning support tool, often emphasizing its efficiency and convenience rather than its pedagogical integration. While previous research highlights its potential to enhance writing performance and learner autonomy, it tends to overlook how students engage critically with AI-generated content during the learning process. Moreover, limited attention has been paid to the social and interactive dimensions of learning, particularly how AI tools, such as ChatGPT, can be integrated into collaborative learning approaches to promote deeper understanding and active knowledge construction.

Recent literature has highlighted the expanding role of ChatGPT in education. Systematic reviews indicate that ChatGPT supports adaptive learning, enhances student engagement, and promotes interactive learning environments (Albadarin et al., 2024; Bettayeb et al., 2024). Its application in higher education has also extended to various academic activities, including writing, research, and instructional support (Bhullar et al., 2024; Ijaz & Yadegaridehkordi, 2024). Empirical studies have revealed that ChatGPT contributes to improved writing quality, supports idea development, and assists in evaluating written responses (Jauhiainen & Guerra, 2025; Rejeb et al., 2024). Additionally, it facilitates personalized learning and encourages greater learner autonomy (Asy'ari & Sharov, 2024; Jo, 2024). Nonetheless, these studies predominantly examine ChatGPT as an individual learning tool, with considerably less attention given to its potential as a mediating tool within socially constructed, collaborative learning environments, a gap that warrants systematic empirical investigation.

Despite these advantages, several concerns have arisen regarding the use of ChatGPT in educational settings. One major issue relates to academic integrity, particularly the risk of plagiarism, as students may become overly dependent on AI-generated content (King, 2023). In addition, ChatGPT may produce inaccurate or repetitive responses and potentially limit students' creativity if it is not used appropriately (Biswas, 2023). These concerns highlight the importance of implementing appropriate

pedagogical strategies to ensure that AI tools are used effectively and responsibly, particularly in writing instruction (Shakil & Siddiq, 2024).

Writing is commonly regarded as one of the most difficult skills for learners in EFL contexts, as it demands the simultaneous management of multiple linguistic, cognitive, and affective processes. In the context of ninth-grade EFL learners, these challenges manifest in several interconnected ways: difficulty in generating and developing ideas into well-structured paragraphs, inconsistent grammatical accuracy, particularly in tense usage and sentence construction, limited vocabulary that restricts expressive range, and low confidence that discourages active participation in writing tasks, especially in collaborative settings (Hyland, 2003). In many Indonesian junior secondary EFL classrooms, these difficulties persist despite the increasing availability of digital learning tools. Students frequently lack sufficient opportunities for meaningful writing practice, and the absence of structured peer interaction further restricts their ability to collaboratively develop ideas and improve their written output. Therefore, writing is not merely a mechanical activity. It is a complex, recursive process involving planning, drafting, and revising that requires both cognitive support and social interaction. AI-based tools like ChatGPT can address the cognitive dimensions of this process, while collaborative learning frameworks provide the social scaffolding necessary for deeper engagement and knowledge construction.

Pedagogical approaches that promote active and critical learning are essential in response to these challenges. One such approach is collaborative learning (CL), which is widely acknowledged for its effectiveness in fostering meaningful learning through social interactions. Grounded in sociocultural theory, particularly Lev Vygotsky's work, learning is considered a socially mediated process where knowledge is developed through interaction with others. Collaborative learning emphasizes peer interaction, allowing students to exchange ideas and negotiate meaning, thereby deepening their understanding of the subject matter. It also facilitates scaffolding, as learners support each other in completing tasks that may be difficult to accomplish individually. Furthermore, collaborative learning enhances critical thinking by promoting students' ability to assess ideas, give feedback, and consider various perspectives (Basoeki & Wu, 2022; Johnson et al., 2010). Previous studies have also indicated that collaborative learning enhances student engagement and academic achievement (Suciati et al., 2024; Wu et al., 2024). These characteristics make collaborative learning particularly suitable for supporting complex processes, such as writing.

Despite the increasing amount of research on AI in education and collaborative learning, existing studies remain fragmented. Research on ChatGPT tends to focus on its role as an individual writing aid, emphasizing efficiency, feedback provision, and learner autonomy, with limited attention to how it can function as a mediating tool within collaborative learning frameworks (Baidoo-Anu & Owusu Ansah, 2023; Jauhiainen & Guerra, 2025). Conversely, studies on collaborative learning highlight peer interaction and engagement, but rarely incorporate AI as an integral part of the collaborative process (Aljabri, 2024; Alzubi et al., 2024). Although several recent studies have started to investigate the integration of AI and collaborative learning in EFL contexts, this line of inquiry remains nascent, with most studies examining engagement and writing performance as separate outcomes rather than simultaneously (Suciati et al., 2024; Wang, 2024). This gap is particularly pronounced at the junior secondary level in Indonesian EFL settings, where empirical evidence on AI-mediated collaborative writing remains scarce. The present study addresses this gap by examining ChatGPT not merely as an individual writing assistant but as a mediating tool embedded within a collaborative writing framework while simultaneously investigating its impact on both student engagement and writing performance.

Therefore, this study examines the effects of AI-mediated collaborative writing through ChatGPT on students' engagement and writing performance among ninth-grade EFL learners at SMPIT Az-Zahra Demak. To address this aim, the following research questions were formulated: (1) Does AI-mediated collaborative writing significantly improve students' behavioral, cognitive, and emotional engagement compared to conventional instruction? (2) Does AI-mediated collaborative writing significantly improve students' writing performance, particularly in text organization and language use? (3) How do students perceive the use of ChatGPT as a mediating tool in collaborative writing activities?

METHOD

Research Design

This study applied a mixed-methods approach using a quasi-experimental design, particularly a non-equivalent control group design, in which participants were not randomly assigned to groups but were drawn from intact classes as they naturally existed within the school setting (Thomas, 2024). The overall research framework is presented in Figure 1. The study involved two groups: a control and an experimental group. The independent variable was the type of instructional approach: conventional collaborative learning and AI-mediated collaborative writing. The dependent variables were student engagement and writing performance. Both groups received similar instructional procedures in terms of tasks, duration, and learning objectives; however, they differed in their use of learning tools. The experimental group used an AI chatbot (ChatGPT) to support collaborative writing activities.

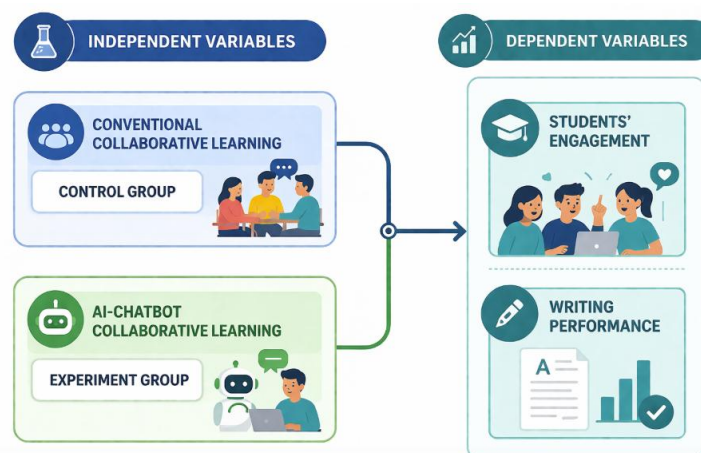


Figure 1. Research Framework

Participants

The participants consisted of 72 ninth-grade students (aged 14–15 years) enrolled at SMPIT Az-Zahra Demak, comprising both male and female students, with females in the majority. Purposive sampling was used to select the participants, as ninth-grade students were considered the most appropriate cohort given their completion of two years of foundational English writing instruction under Kurikulum Merdeka, greater cognitive maturity, and developing, though not yet proficient, writing competencies in idea organization, grammatical accuracy, and vocabulary use. Since only two ninth-grade classes existed at the school, both were included without further selection: one class served as the control group ($n = 36$) while the other functioned as the experimental group ($n = 36$), assigned based on existing class schedules. Students' prior familiarity with ChatGPT

varied; consequently, a brief orientation session was provided to the experimental group before the intervention.

Research Instrument

This study employed three instruments: a questionnaire, a writing test, and a semi-structured interview. Student engagement was measured using a questionnaire adapted from Reeve & Tseng (2011). The questionnaire contained 17 statements divided into three categories: five items for behavioral engagement, four items for emotional engagement, and eight items for cognitive engagement. Responses were scored using a five-point Likert scale, ranging from strongly disagree (1) to strongly agree (5). The questionnaire was administered in Bahasa Indonesia to ensure clarity for participants. All items were found to be valid, with validity coefficients ranging from .601 to .682, exceeding the critical r -value of .329 ($p < .05$). The reliability coefficients were .657 for the behavioral engagement, .689 for the emotional engagement, and .789 for the cognitive engagement, indicating acceptable internal consistency for all dimensions.

Writing performance was assessed using a writing test evaluated using an analytical scoring rubric adapted from (Silitonga et al., 2023). The rubric evaluated five aspects of writing, namely content, organization, vocabulary, language use, and mechanics. The rubric was explained before implementation to ensure students' understanding. Several modifications were also introduced to match the learners' proficiency level, including simplifying certain criteria and reducing the required word count to 75–100 words. All rubric components yielded valid correlation values above the critical r -value of .329 ($p < .05$), and the rubric demonstrated strong internal consistency with a Cronbach's alpha of .965. Scoring was conducted independently by two raters, the researcher and the English subject teacher, to ensure objectivity and minimize evaluator bias.

Students' perceptions were explored through semi-structured interviews adapted from (Rahimi & Fathi, 2022). The interview instrument consisted of six questions administered to students in the experimental group. Several adjustments were made to align the interview questions with the context of the present study. Participants were selected based on their performance, representing high, medium, and low achievers. All interviews were recorded with participants' consent and transcribed for thematic analysis. Instrument testing was conducted to evaluate validity and reliability. Pearson's product-moment correlation was used to examine validity, while reliability was measured using Cronbach's alpha. The results showed that all questionnaire items were valid ($p < .05$) and reliable ($\alpha > 0.60$).

Data collection was completed in three stages over 11 sessions, each lasting 80 minutes. The overall procedural framework is presented in Figure 2. In the pre-intervention phase (Session 1), both groups completed an engagement questionnaire as a preliminary measurement of students' engagement. During the intervention phase (Sessions 2–8), both groups received equivalent instructional content covering academic writing, vocabulary, the simple present tense, and the use of articles, delivered through collaborative learning activities. The instruction was organized into three cycles of pre-writing and peer-feedback activities on progressively varied topics: Best Friend, Pet, and Favourite Thing. In each pre-writing session, students produced an essay of 75–100 words, while peer-feedback sessions involved group discussions of four students, rubric-guided peer review, and whole-class error discussion. Although both groups followed identical instructional procedures and covered the same topics, they differed in the writing support tool they employed. The control group utilized Google Translate throughout the writing activities, whereas the experimental group integrated ChatGPT as a mediating tool to support vocabulary exploration, idea generation, drafting, and peer review.

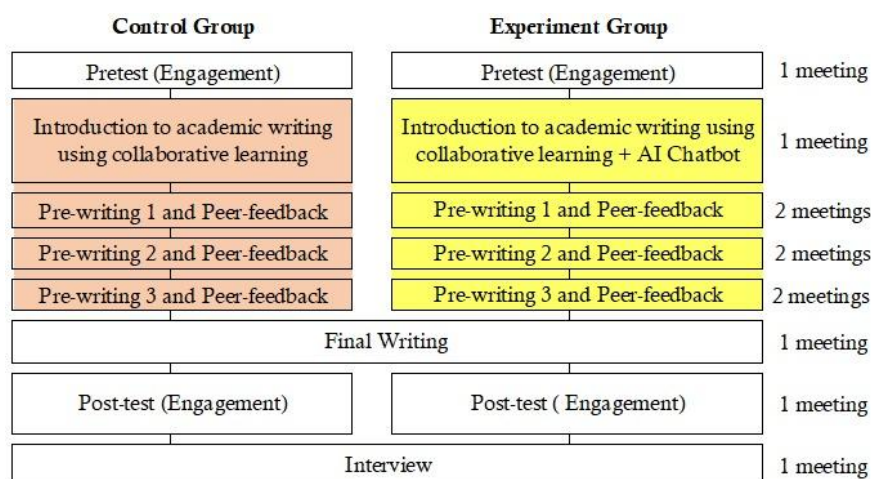


Figure 2. Experiment Procedure

In the post-intervention phase (Sessions 10–11), the engagement questionnaire was re-administered as a post-test to both groups. Writing performance was assessed based on the final writing task using an analytical rubric. Semi-structured interviews were then carried out with six selected students from the experimental group representing different levels of writing performance. With participants' permission, all interviews were recorded and later transcribed for qualitative analysis.

Data Analysis

The data were analyzed through several stages. First, instrument testing was performed to examine the validity and reliability of the instruments. Pearson's Product-Moment correlation was used to examine validity, whereas reliability was measured through Cronbach's alpha coefficient. The instrument was regarded as valid when the significance value was lower than 0.05 and reliable when Cronbach's alpha was higher than 0.70. After confirming the instrument quality, assumption testing was performed. Data normality was assessed using the Shapiro–Wilk test with a significance level of 0.05. The data were categorized as normally distributed when the significance value exceeded 0.05. Furthermore, homogeneity of variance was tested using Levene's test to confirm the homogeneity of variance across groups. Once the assumptions were met, the main statistical analyses were conducted.

For the quantitative strand, ANCOVA was employed to analyze differences in students' engagement across three dimensions, behavioral, cognitive, and emotional, between the control and experimental groups, by using pre-test engagement scores as covariate variables to statistically control for baseline differences between groups. This approach was selected because ANCOVA increases statistical precision by accounting for pre-existing variation, thereby allowing a more precise estimation of the treatment effect on post-test engagement scores. ANOVA was applied to examine differences in writing performance between the two groups. Since no pre-test writing score was administered, writing performance was measured through a composite score derived from the average of three in-process writing tasks and one final writing task.

For the qualitative strand, the interview recordings were transcribed and analyzed through thematic analysis based on Braun & Clarke (2006) Six-phase framework including data familiarization, initial coding, theme identification, theme review, theme definition, and report writing. MAXQDA software was utilized to support the coding process and theme organization. The qualitative findings were subsequently integrated with the quantitative results at the interpretation stage to obtain a broader understanding of students' experiences with AI-mediated collaborative writing.

RESULTS AND DISCUSSION

Results

This study examined the impact of AI-mediated collaborative writing on students' engagement and writing performance among ninth-grade EFL learners. The findings revealed that the experimental group, which employed ChatGPT as a supporting learning tool, showed greater progress than the control group in all dimensions of engagement, including behavioral, cognitive, and emotional engagement, as well as in writing performance, especially in organization and language use. These findings were supported by descriptive statistics, inferential analyses, and qualitative themes derived from student interviews. Nevertheless, the results should be interpreted within the context of the present study and do not imply that AI-mediated collaborative writing will be equally effective across all EFL classrooms or instructional settings.

Quantitative Results

The quantitative data were analyzed to investigate the effect of the instructional treatment on students' engagement and writing performance. Before the primary analysis, assumption testing was carried out to determine whether the data fulfilled the required statistical assumptions. The Shapiro–Wilk test results showed that the data were normally distributed since all significance values exceeded 0.05. Furthermore, the Levene's test results indicated homogeneous data distribution ($p > 0.05$). Therefore, the data met the assumptions required for further statistical analysis.

Table 1. Descriptive results for students' engagement

Variable	Control				Experiment			
	Pre test		Post test		Pre test		Post test	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Behaviour	19.25	2.046338	20.16667	2.179449	16.41667	0.924211	23.75	0.759203
Emotional	14.94444	1.884799	15.55556	1.77082	13.13889	0.713083	18.86111	0.854816
Cognitive	30.33333	2.718251	30.69444	3.828399	26.41667	1.187317	38.13889	1.134136

Table 2. ANCOVA result for students' engagement

Variable	SS	df	Mean Square	F	P	Partial η^2
Behaviour	162.300	1	162.300	59.773	.000	.544
Emotional	157.412	1	157.412	79.001	.000	.579
Cognitive	592.204	1	592.204	71.613	.000	.626

Tables 1 and 2 show the results of students' engagement. Descriptive statistics showed that the experimental group experienced a notable increase in all dimensions of engagement compared with the control group. In the behavioral dimension, the mean score of the experimental group showed improvement from 16.41 to 23.75, while the control group showed only a slight improvement, from 19.25 to 20.17. Similarly, in the emotional dimension, the experimental group improved from 13.13 to 18.86, whereas the control group increased marginally from 14.94 to 15.56. A more substantial improvement was also observed in the cognitive dimension, with the experimental group rising from 26.41 to 38.13, compared with a minimal increase in the control group from 30.33 to 30.69.

Furthermore, the ANCOVA results revealed a statistically significant difference between the control and experimental groups in all engagement dimensions. The behavioral dimension showed a significant effect ($F = 59.773$, $p < 0.05$, partial $\eta^2 = 0.544$), followed by the emotional ($F = 79.001$, $p < 0.05$, partial $\eta^2 = 0.579$) and cognitive dimensions ($F = 71.613$, $p < 0.05$, partial $\eta^2 = 0.626$). These findings indicate that the use of the AI chatbot in collaborative writing activities had a statistically significant effect on all dimensions of students' engagement, with large effect sizes observed across the behavioral, emotional, and cognitive aspects.

Table 3. Descriptive result for writing skill

Variable	Control			Experiment		
	Sum	Mean	SD	Sum	Mean	SD
Writing skill	1719.00	11.9375	1.36064	1804.00	12.5278	1.29115

Table 4. ANOVA result for writing skill

Variable	SS	df	Mean Square	F	p
Writing skill	6.272	1	6.272	3.565	.043

$P < 0.05$.

Tables 3 and 4 present the results of students' writing performance. Descriptive statistics indicated that the experimental group obtained a higher mean score ($M = 12.52$, $SD = 1.29$) than the control group ($M = 11.93$, $SD = 1.36$). ANOVA results showed a statistically significant difference between the two groups ($F = 3.565$, $p < 0.05$). This finding suggests that instructional treatment using an AI chatbot significantly affected students' writing performance compared to the conventional method.

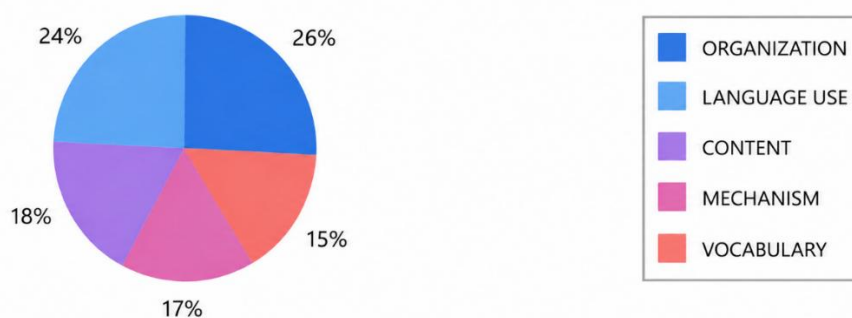


Figure 3. Writing Performance

Qualitative Results

To enrich the quantitative results, qualitative findings from the semi-structured interview were examined using thematic analysis assisted by MAXQDA. The analysis identified five main aspects of students' writing performance: organization, language use, content, mechanics, and vocabulary, as illustrated in Figure 3. Among these, organization emerged as the most dominant aspect (26%), followed by language use (24%); content and mechanics contributed moderately, accounting for 18% and 17%, respectively, while vocabulary was the least represented aspect (15%). These findings provide deeper insights into how the integration of the AI chatbot facilitated improvements in students' writing performance, particularly in organizing ideas and enhancing language use, which is consistent with the quantitative results.

Discussion

This study aimed to investigate the impacts of AI-mediated collaborative writing on students' engagement and writing performance. The result showed that the integration of the AI chatbot significantly improved students' engagement across behavioral, emotional, and cognitive dimensions and their writing performance. The findings imply that the use of AI in collaborative activities may encourage more active classroom interaction and support students' learning progress.

From a theoretical perspective, these findings are grounded in Vygotsky's (1978) sociocultural theory, which highlights the role of social interaction and cultural mediation in the learning process. In this study, ChatGPT functioned as a mediational tool within

the Zone of Proximal Development (ZPD), providing scaffolding that enabled students to accomplish writing tasks that might have been difficult to complete independently. Specifically, ChatGPT supported idea generation, paragraph organization, vocabulary selection, and grammatical accuracy, areas that are consistently identified as primary challenges for EFL learners at the secondary level (Hyland, 2003; Moses & Mohamad, 2019). Importantly, this scaffolding did not replace peer interaction but complemented it, as students continued to engage in group discussions and peer feedback activities throughout the intervention. This combination of AI-mediated support and social interaction reflects the dual cognitive and social dimensions of effective writing instruction.

The improvements observed across all three engagement dimensions are in line with previous studies showing the positive impact of technology-enhanced learning on student motivation, participation, and confidence (Albadarin et al., 2024; Bettayeb et al., 2024). The cognitive engagement gains, which yielded the largest effect size in this study (partial $\eta^2 = .626$), suggest that ChatGPT encouraged students to think more deeply about their writing, evaluate vocabulary choices, reconsider sentence structures, and reflect on peer feedback. This finding extends prior research by Baidoo-Anu & Owusu Ansah (2023), who emphasized the potential of ChatGPT to support learner autonomy, demonstrating that this autonomy extends to collaborative settings where students collectively negotiate meaning with AI assistance. Similarly, the emotional and behavioral engagement gains suggest that students developed greater confidence and willingness to participate actively in writing tasks. Similar findings were also reported by Asy'ari and Sharov (2024), who stated that AI-supported learning reduces writing anxiety and fosters more positive attitudes toward writing tasks.

Beyond engagement, the effects of AI-mediated collaborative writing also extended to students' writing performance. The findings showed that the experimental group performed significantly better on the post-test than the control group ($F = 3.565$, $p = .043$), with qualitative findings further revealing that the most prominent improvements occurred in organization (26%) and language use (24%). These results are consistent with Bibi and Atta (2024) and Imran and Almusharraf (2023), who found that ChatGPT facilitates idea development and structural coherence in EFL writing. However, this study further demonstrates that these improvements can be achieved not only through individual use of AI but also through AI-mediated collaborative writing at the junior secondary level, a context that has received considerably less empirical attention in the existing literature. The qualitative data further illuminate this finding: students reported that ChatGPT helped them explore vocabulary, construct grammatically accurate sentences, and organize their ideas more systematically, which suggests that AI tools can effectively support scaffolding in writing instruction (Dao & Dan, 2024).

Despite these promising findings, several important considerations warrant attention. First, there is a potential risk of students becoming overdependent on AI-generated content, whereby students might rely excessively on ChatGPT rather than developing their own writing competencies (King, 2023). To mitigate this risk, structured usage guidelines were implemented in this study; however, long-term dependence remains a concern that future implementations should address through explicit instruction on critical engagement with AI output. Second, unequal digital literacy among students may have influenced the extent to which individual learners benefited from ChatGPT, as students with limited prior experience with AI tools may have required additional support to use the tool effectively. Third, ethical considerations surrounding AI use in educational settings, including concerns related to academic integrity and data privacy, must be carefully managed when integrating such tools into classroom instruction (UNESCO, 2023).

In terms of pedagogical implications, this study suggests that integrating AI chatbots into collaborative writing activities can be an effective strategy for enhancing both engagement and writing skills. Teachers may consider incorporating AI-based tools as scaffolding support to facilitate students' learning processes and encourage more active participation in classrooms. Nevertheless, several limitations should be acknowledged in this study. First, the number of participants was relatively limited and drawn from only one educational setting. This may reduce the generalizability of the finding. Furthermore, the intervention was conducted over a relatively short period, making it difficult to fully identify the long-term effect of AI-assisted learning. Therefore, future studies are recommended to involve a larger participant group and a longer intervention period for broader insights into the effectiveness of AI chatbot integration in the educational context.

CONCLUSION

Taken together, the results of this study affirm that purposeful integration of ChatGPT within collaborative writing frameworks represents a viable and effective pedagogical approach for addressing the multidimensional challenges of EFL writing at the junior secondary level. Beyond confirming significant gains in engagement and writing performance, the results highlight the importance of combining AI-mediated support with peer interaction, a combination that appears to address both the cognitive and social dimensions of writing development simultaneously. Students' positive perceptions of ChatGPT as a mediating tool further reinforce its pedagogical value, particularly in supporting idea generation, vocabulary exploration, and the peer review process. These outcomes, however, should be interpreted within the study's contextual boundaries, given its limited number of participants, single educational setting, and relatively brief intervention period. Future research should examine this approach across larger and more diverse samples, extended timeframes, varied writing genres, and different AI tools to establish a more generalizable and robust evidence base for AI-mediated collaborative writing in EFL pedagogy.

RECOMMENDATION

Based on the findings of this study, future research is recommended to involve larger and more diverse participants as well as extended intervention periods, to more rigorously examine the long-term effects of AI-mediated collaborative writing on students' engagement and writing performance. Further studies are encouraged to explore the implementation of AI chatbots in different writing genres and language skills to obtain a broader understanding of their effectiveness. However, several potential barriers should be considered, such as the limited sample size, the short duration of the intervention, and differences in students' digital literacy and access to technology, which could influence the effectiveness of AI use in educational settings.

ACKNOWLEDGMENT

The researcher would also like to express sincere appreciation to Universitas Persatuan Guru Republik Indonesia Semarang (UPGRIS) Semarang for the academic support provided during the research process. The author also thanks the principal, students, and staff of SMPIT Az-Zahra Demak for support and cooperation during this research process. The researcher would also like to extend sincere gratitude to the supervisors for their valuable guidance and constructive suggestions throughout this study.

DECLARATION OF USING AI TOOLS

ChatGPT (OpenAI, GPT-4) was utilized in this study in two capacities. First, it served as the primary instructional tool for the experimental group, supporting collaborative writing activities including idea generation, vocabulary exploration, drafting, and peer review

processes. Second, ChatGPT was used by the authors during manuscript preparation to assist with language refinement of selected sections. All AI-supported content underwent further review and revision by the authors, who take full responsibility for the final manuscript's accuracy, integrity, and originality.

REFERENCES

- Albadarin, Y., Saqr, M., Pope, N., & Tukiainen, M. (2024). A systematic literature review of empirical research on ChatGPT in education. *Discover Education*. <https://doi.org/10.1007/s44217-024-00138-2>
- Aljabri, S. (2024). Engagement and preferences: EFL students' use of collaborative retrieval practice. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2394741>
- Allayarova, S. U. (2025). Transformation in 21st-century education: Digital pedagogy and learner-centered teaching trends. *International Journal of Social Sciences & Interdisciplinary Research*, 14(06), 132–146.
- Alzubi, A. A., Nazim, M., & Ahamad, J. (2024). Examining the effect of a collaborative learning intervention on EFL students' English learning and social interaction. *Journal of Pedagogical Research*, 8(2), 26–46. <https://doi.org/10.33902/JPR.202425541>
- Amin, M. Y. M. (2023). AI and ChatGPT in language teaching: Enhancing EFL classroom support and transforming assessment techniques. *International Journal of Higher Education Pedagogies*, 4(4), 1–15. <https://doi.org/10.33422/ijhep.v4i4.554>
- Angguni, P. E., Saputri, M., & Sugesti, N. (2026). The relationship between digital literacy and self-perceived academic. *Journal of Language and Literature Studies*, 6(1), 43–54. <https://doi.org/10.36312/jolls.v6i1.4496>
- Asy'ari, M., & Sharov, S. (2024). Transforming education with ChatGPT: Advancing personalized learning, accessibility, and ethical AI integration. *International Journal of Essential Competencies in Education*, 3(2), 119–157. <https://doi.org/10.36312/ijece.v3i2.2424>
- Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52–62. <https://doi.org/10.61969/jai.1337500>
- Basoeki, O. de H., & Wu, T.-T. (2022). Enhancement of reading comprehension skills in collaborative setting: A preliminary research on students' perception. *ICITL*, 372–381. <https://doi.org/10.1007/978-3-031-15273-3>
- Bettayeb, A. M., Talib, M. A., Altayasinah, A. Z. S., & Dakalbab, F. (2024). Exploring the impact of ChatGPT: conversational AI in education. *Frontiers in Education*, <https://doi.org/10.3389/feduc.2024.1379796>
- Bhullar, P. S., Joshi, M., & Chugh, R. (2024). ChatGPT in higher education: Synthesis of the literature and a future research agenda. *Education and Information Technologies*, 29(21), 21501–21522. <https://doi.org/10.1007/s10639-024-12723-x>
- Bibi, Z., & Atta, A. (2024). The role of ChatGPT as an AI English writing assistant : A study of students' perceptions, experiences, and satisfaction level. *Annals of Human and Social Sciences*, 5(1), 433–443. [https://doi.org/10.35484/ahss.2024\(5-I\)39](https://doi.org/10.35484/ahss.2024(5-I)39)
- Biswas, S. (2023). ChatGPT and the future of medical education. *Academic Medicine*, 98(8), 867–868. <https://doi.org/10.1148/radiol.223312>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>

- Dao, V. T. A., & Dan, T. C. (2024). Secondary school EFL students' difficulties and solutions to learning English paragraph writing. *European Journal of Applied Linguistics Studies*, 7(2), 82–117. <https://doi.org/10.46827/ejals.v7i2.552>
- Darus, S. (2009). Error analysis of the written English essays of secondary school students in Malaysia : A case study. *European Journal of Social Science*, 8(3), 483–495.
- Djibran, A. K. S., Subiyanto, P., & Rahayu, N. S. (2024). Transforming education in the digital age : How technology affects teaching and learning methods. *Journal of Pedagogi : Jurnal Pendidikan*, 1(3), 141–155. <https://doi.org/10.62872/s13px737>
- Hassan, A. (2026). Leveraging artificial intelligence (AI) to enhance student engagement and academic performance in higher education. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-026-13946-w>
- Hyland, K. (2003). *Second language writing*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511667251>
- Ijaz, M., & Yadegaridehkordi, E. (2024). ChatGPT in higher education : A systematic literature review and research challenges. *International Journal of Educational Research*, 127. <https://doi.org/10.1016/j.ijer.2024.102411>
- Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4). <https://doi.org/10.30935/cedtech/13605>
- Jauhiainen, J. S., & Guerra, A. G. (2025). Generative AI in education: ChatGPT-4 in evaluating students ' written responses. *Innovations in Education and Teaching International*, 62(4), 1377–1394. <https://doi.org/10.1080/14703297.2024.2422337>
- Jo, H. (2024). From concerns to benefits : a comprehensive study of ChatGPT usage in education. *International Journal of Educational Technology in Higher Education*, 21(35). <https://doi.org/10.1186/s41239-024-00471-4>
- Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). Cooperative learning returns to college: what evidence is there that it works? *Change: The Magazine of Higher Learning*, 30(4), 26–35. <https://doi.org/10.1080/00091389809602629>
- Kemendikbudristek RI. (2022). Panduan pemanfaatan teknologi digital dalam pembelajaran. <https://kemendikdasmen.go.id/pengumuman/14932-skb-pedoman-pemanfaatan-dan-pembelajaran-teknologi-digital-dan-ka>
- King, M. R. (2023). A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and Molecular Bioengineering*, 16(1), 1–2. <https://doi.org/10.1007/s12195-022-00754-8>
- Long, D. Y., Wang, S., Rashid, S. M., & Lu, X. T. (2026). Artificial intelligence in higher education: A systematic review of its impact on student engagement and the mediating role of teaching methods. *Frontiers in Education*, 10, 1–18. <https://doi.org/10.3389/feduc.2025.1648661>
- 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>
- Moses, R. N., & Mohamad, M. (2019). Challenges faced by students and teachers on writing skills in ESL contexts : A literature review. *Creative Education*, 10, 3385–3391. <https://doi.org/10.4236/ce.2019.1013260>
- OECD. (2023). OECD digital education outlook 2023: Towards an effective digital education ecosystem. *OECD*. https://www.oecd.org/en/publications/oecd-digital-education-outlook-2023_c74f03de-en.html

- Rahimi, M., & Fathi, J. (2022). Exploring the impact of wiki-mediated collaborative writing on EFL students' writing performance, writing self-regulation, and writing self-efficacy: a mixed methods study. *Computer Assisted Language Learning*, 35(9), 2627–2674. <https://doi.org/10.1080/09588221.2021.1888753>
- Reeve, J., & Tseng, C. M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology*, 36(4), 257–267. <https://doi.org/10.1016/j.cedpsych.2011.05.002>
- Rejeb, A., Rejeb, K., Appolloni, A., Treiblmaier, H., & Iranmanesh, M. (2024). Exploring the impact of ChatGPT on education: A web mining and machine learning approach. *The International Journal of Management Education*, 22(1). <https://doi.org/10.1016/j.ijme.2024.100932>
- Shakil, E., & Siddiq, A. (2024). ESL teachers' perceptions about ChatGPT as a threat to analytical writing abilities of ESL learners at graduate level. *Pakistan Languages and Humanities Review*, 8(1), 115–128. [https://doi.org/10.47205/plhr.2024\(8-i\)10](https://doi.org/10.47205/plhr.2024(8-i)10)
- Silitonga, L. M., Hawanti, S., Aziez, F., Furqon, M., Zain, D. S. M., Anjarani, S., & Wu, T. T. (2023). The impact of AI chatbot-based learning on students' motivation in English writing classroom. In *Lecture Notes in Computer Science* (vol. 14099, pp. 542–549). Springer. https://doi.org/10.1007/978-3-031-40113-8_53
- Suciati, S., Silitonga, L. M., Wiyaka, Huang, C.-Y., & Anggara, A. A. (2024). Enhancing engagement and motivation in English writing through AI: The impact of ChatGPT-supported collaborative learning. *Springer Nature Switzerland AG*, 1, 205–214. https://doi.org/10.1007/978-3-031-65884-6_21
- Swai, C. T. (2025). A systematic review of classroom technologies supporting student centered teaching. *Discover Education*, 4, Article 564. <https://doi.org/10.1007/s44217-025-00932-6>
- Thomas, L. (2024). *Quasi-experimental design | definition, types & examples*. Scribbr. <https://www.scribbr.com/methodology/quasi-experimental-design/>
- UNESCO. (2023). *Guidance for generative AI in education and research*. <https://doi.org/10.54675/EWZM9535>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wang, S. (2024). The dimensions and dynamism of group engagement in computer-mediated collaborative writing in EFL classes. *SAGE Open*, 1–14. <https://doi.org/10.1177/21582440241227707>
- Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems With Applications*, 252, 124167. <https://doi.org/10.1016/j.eswa.2024.124167>
- Wu, T. T., Silitonga, L. M., & Murti, A. T. (2024). Enhancing English writing and higher-order thinking skills through computational thinking. *Computers and Education*, 213, 105012. <https://doi.org/10.1016/j.compedu.2024.105012>
- Yu, H. (2023). Reflection on whether ChatGPT should be banned by academia from the perspective of education and teaching. *Frontiers in Psychology*, 01–12. <https://doi.org/10.3389/fpsyg.2023.1181712>