

INTEGRATING CHATGPT INTO EFL CRITICAL READING: EFFECTS ON STUDENTS' ANALYTICAL THINKING

Ilham Sahabuddin^{1*}, Asriadi Rasyid², & Sianna³

^{1,3}Universitas Muhammadiyah Parepare, Parepare, Indonesia

²Institut Teknologi dan Bisnis Nobel Indonesia, Makassar, Indonesia

*ilham852803@gmail.com

ABSTRACT

This study examined the integration of ChatGPT into critical reading instruction and its association with students' analytical thinking in an EFL higher education context. The study was motivated by the limited empirical evidence on how ChatGPT supports critical reading skills, particularly in identifying arguments, evaluating evidence, making inferences, comparing perspectives, and constructing analytical responses. A quantitative approach with a pre-experimental one-group pretest-posttest design was employed. The study was conducted at Universitas Muhammadiyah Parepare and involved 10 fourth-semester students of the English Education Study Program enrolled in a Critical Reading course. Data were collected through analytical thinking tests administered before and after the intervention, using a critical reading rubric covering argument identification, fact-opinion distinction, evidence evaluation, inference, perspective comparison, idea connection, and evidence-based response construction. The data were analyzed using descriptive statistics, Shapiro-Wilk normality test, paired sample t-test, normalized gain, and Cohen's d. The findings showed that the mean score increased from 58.50 in the pretest to 74.50 in the posttest, with a mean gain of 16.00. The normalized gain was 0.39, indicating moderate improvement. The paired sample t-test revealed a significant difference, $t(9) = 13.74$, $p < 0.001$, and Cohen's $d = 4.35$ indicated a very large effect size. The results indicated that ChatGPT integration was associated with improved students' analytical thinking in critical reading.

Keywords: AI-Assisted Language Learning, Analytical Thinking, ChatGPT, Critical Reading, EFL Learning

INTRODUCTION

The rapid development of generative artificial intelligence (AI) has produced a profound shift in higher education, particularly in English as a Foreign Language (EFL) teaching and learning. Tools such as ChatGPT can generate explanations, summarize complex texts, translate unfamiliar passages, provide examples, and respond interactively to learners' questions. These affordances are pedagogically significant because EFL students in higher education are expected not only to decode English texts but also to interpret, evaluate, and respond to academic ideas in intellectually responsible ways (Zaimoğlu & Dağtaş, 2025). Recent scholarship indicates that ChatGPT can enrich language learning through immediate

feedback, adaptive interaction, and learner-centered support; however, its value depends strongly on the quality of instructional design, teacher guidance, and students' ability to evaluate AI-generated responses (C. Meniado, 2023; Kohnke et al., 2023; Lo et al., 2024). Therefore, ChatGPT should not be treated merely as a technological innovation in EFL classrooms, but as a pedagogical issue related to cognitive engagement, academic literacy, and higher-order thinking.

In academic literacy practices, many EFL students already use ChatGPT to translate difficult passages, explain unfamiliar vocabulary, summarize articles, simplify theoretical concepts, and generate possible answers to reading questions. These practices may reduce linguistic barriers and increase students' confidence when they engage with English academic texts (Lo et al., 2024; Shafiee Rad, 2025). Nevertheless, easier access to explanations does not automatically lead to deeper reading. When students accept AI-generated responses without checking textual evidence, questioning assumptions, or comparing alternative interpretations, ChatGPT may unintentionally reinforce passive comprehension, dependency, and surface-level learning. This concern is particularly relevant in higher education, where students are expected to develop independent judgment rather than merely reproduce information. Studies on generative AI in education emphasize that large language models require critical AI literacy, fact-checking strategies, and human oversight because their outputs may be incomplete, biased, overgeneralized, or contextually inaccurate (Kasneci et al., 2023; Pujiastuti et al., 2025). The meta-analysis on ChatGPT, learning performance, and higher-order thinking also reminds researchers to interpret claims about AI effectiveness cautiously and to rely on context-sensitive empirical evidence rather than broad technological optimism (Liang & Wu, 2024; Shang et al., 2025).

Critical reading is one of the most important academic competencies for EFL students because it requires readers to move beyond literal comprehension toward interpretation, evaluation, inference, and reasoned judgment (Dorgham & Obiad, 2025; Niculescu & Dragomir, 2023). In critical reading, students are expected to identify an author's main claim, examine how arguments are organized, evaluate the relevance and sufficiency of evidence, distinguish facts from opinions, recognize bias, infer implicit meanings, and formulate responses grounded in textual support. These activities are cognitively demanding for EFL learners because they must simultaneously manage linguistic challenges, disciplinary concepts, and argumentative structures. Consequently, critical reading cannot be reduced to understanding vocabulary or locating explicit information. It is a reflective and evaluative practice through which learners interrogate texts, compare perspectives, and construct informed interpretations.

Analytical thinking is central to this process. It enables students to break down complex information, classify ideas, identify relationships among claims and evidence, detect inconsistencies, examine assumptions, and draw logical conclusions. In higher education, analytical thinking also supports self-regulated learning because students must monitor the quality of their own interpretations and revise them when new evidence appears (Dresel et al., 2015). In EFL reading instruction, the connection between critical reading and analytical thinking is particularly strong: students cannot evaluate an academic text critically unless they can analyze how meaning is constructed, how arguments are justified, and how

evidence functions within the text. Recent studies on EFL reading and higher-order thinking similarly show that students need explicit instruction, guided questioning, and reflective tasks to develop interpretive, evaluative, and analytical reading abilities (Akbar et al., 2025; Ledesma Acosta et al., 2025; Liang & Wu, 2024; Liu & Puteh, 2025)

The integration of ChatGPT into critical reading instruction can be understood through this relationship between reading, analysis, and guided inquiry. ChatGPT may function as a dialogic scaffold that helps learners ask questions, compare interpretations, request counterarguments, clarify concepts, and revise their initial responses (Shoufan, 2023; Stojanov, 2023). From a sociocultural perspective, learning is mediated by tools, interaction, and purposeful guidance; therefore, AI can support cognitive development when it is embedded in meaningful pedagogical activity rather than used as a replacement for thinking (Alkhudiry, 2022; Kasneci et al., 2023). In a critical reading class, for example, students may use ChatGPT to generate possible interpretations of a paragraph, but they must then verify those interpretations against textual evidence. They may ask ChatGPT to identify an author's claim, but they must evaluate whether the answer is accurate, whether the cited evidence is relevant, and whether alternative readings are possible. In this sense, ChatGPT can become a tool for inquiry, critique, and metacognitive reflection when students are trained to interrogate its responses rather than consume them passively.

However, the pedagogical impact of ChatGPT is not inherent in the technology itself. The same tool that can support analytical reading may also weaken it when tasks are poorly designed. If students use ChatGPT only to obtain summaries or ready-made answers, the tool may bypass the very cognitive processes that critical reading is supposed to develop. Conversely, when ChatGPT is integrated through structured prompts, evidence-based verification, collaborative discussion, and reflective evaluation, it can stimulate students to engage more actively with texts. This distinction is supported by recent studies suggesting that ChatGPT is most productive when it is used within explicit learning designs that require learners to evaluate, revise, and justify AI-assisted outputs (Liang & Wu, 2024; Shang et al., 2025). Therefore, the central question is not whether ChatGPT should be used in EFL reading classes, but how it can be integrated in ways that strengthen students' analytical agency.

Previous research has increasingly examined the role of ChatGPT and other AI technologies in language education. Kohnke et al. (2023) highlight ChatGPT's potential to support interaction, language practice, and learner autonomy while emphasizing the importance of teacher control and ethical awareness (Kohnke et al., 2023). Liang & Wu (2024), Lo et al. (2024), and Shang et al. (2025), through a rapid review, reports that ChatGPT may enhance learning efficiency and provide personalized support, but its effectiveness depends on task design and students' ability to evaluate AI outputs (Liang & Wu, 2024; Lo et al., 2024; Shang et al., 2025). Similarly, Akinsemolu and Onyeaka (2025) discuss AI's contribution to automated feedback, adaptive instruction, and intelligent tutoring in language learning, although they also note concerns about reliability and learner dependency (Akinsemolu & Onyeaka, 2025). Meniado (2023) reports that ChatGPT can support English language teaching, learning, and assessment by providing scaffolding, feedback, and language practice opportunities (C. Meniado, 2023). Other studies have documented ChatGPT's value in writing instruction, including idea generation, grammar

revision, organization, and feedback (Mahapatra, 2024), as well as broader applications in vocabulary, grammar, speaking, storytelling, and autonomous language learning (Eragamreddy, 2025; Slamet, 2024). These studies show that AI-assisted language learning is a rapidly expanding field with clear pedagogical potential.

Despite this growing body of research, important gaps remain. Systematic reviews indicate that most empirical studies on ChatGPT in ESL/EFL education have concentrated on writing, while reading, speaking, and listening remain comparatively under-researched (Lo et al., 2024). This imbalance is theoretically and pedagogically important because reading comprehension support is not the same as critical reading development. A tool may help students understand difficult vocabulary or summarize an article without necessarily helping them evaluate argument quality, identify bias, assess evidence, or construct independent interpretations. Existing studies on AI-supported reading tend to emphasize comprehension, engagement, or self-regulated learning (Shafiee Rad, 2025), whereas fewer studies examine how ChatGPT can be structured to support critical reading as an analytical and evaluative practice. Similarly, although studies on critical reading in EFL contexts stress the importance of higher-order thinking (Dorgham & Obiad, 2025; Ledesma Acosta et al., 2025), they often do not focus specifically on generative AI as a dialogic tool for argument analysis and evidence evaluation.

Recent empirical work has begun to address this issue, but the field is still developing. Liang and Wu (2024) found that ChatGPT-supported pedagogical design can help EFL learners reflect on critical thinking processes, although its effects depend on how students balance AI assistance with human cognition (Liang & Wu, 2024). Shang et al. (2025) further suggest that competitive questioning between learners and ChatGPT may enhance reading comprehension and critical thinking by encouraging students to question, compare, and justify responses (Shang et al., 2025). These studies are valuable, but they also indicate the need for more contextual evidence from diverse EFL higher education settings. In particular, there remains limited empirical attention to how ChatGPT-assisted critical reading instruction influences students' analytical thinking in Indonesian higher education, especially within English Education programs where students are preparing to become future language teachers.

This study responds to that gap by focusing on students in the Critical Reading course at the English Education Study Program of Universitas Muhammadiyah Parepare. In this context, students are increasingly exposed to generative AI in academic activities, yet many still require systematic guidance to transform AI use from a convenience tool into a learning resource for analysis, evaluation, and reflection. The study is therefore significant for three reasons. First, it addresses an urgent pedagogical problem: the growing use of ChatGPT among EFL students without sufficient instructional frameworks for critical and ethical engagement. Second, it contributes to the literature on critical reading by examining analytical thinking as a measurable and teachable outcome of AI-assisted instruction. Third, it offers contextual evidence from an Indonesian EFL higher education setting, thereby expanding the geographical and institutional diversity of research on AI-assisted language learning.

The novelty of this study lies in its positioning of ChatGPT not as a source of answers, but as a structured dialogic scaffold for critical reading. Unlike studies that focus mainly on writing assistance, general perceptions of AI, or broad language learning benefits, this study investigates how ChatGPT can be integrated into reading instruction to help students analyze arguments, evaluate evidence, identify assumptions, make inferences, compare perspectives, and formulate analytical responses. This focus distinguishes the study from previous research because it connects three domains that are often discussed separately: critical reading pedagogy, analytical thinking development, and generative AI integration. By examining their intersection, the study provides a more nuanced understanding of how AI-assisted instruction can either support or hinder higher-order reading practices depending on the pedagogical conditions under which it is implemented.

Accordingly, this study aims to examine the impact of integrating ChatGPT into critical reading instruction on EFL students' analytical thinking. More specifically, it investigates how ChatGPT-assisted instruction helps students understand academic texts critically, evaluate arguments, identify textual evidence, make inferences, detect weaknesses in reasoning, and develop analytical responses to reading materials. The findings are expected to provide theoretical, pedagogical, and practical contributions. Theoretically, the study enriches discussion on AI-assisted language learning by clarifying the relationship between generative AI, critical reading, and analytical thinking. Pedagogically, it offers insights into how lecturers can design prompts, tasks, and classroom interactions that guide students toward responsible AI use. Practically, it may help English Education programs prepare students to become critical, ethical, and reflective users of AI in academic and professional literacy practices.

METHODS

Research Design

This study employed a quantitative approach using a pre-experimental one-group pretest-posttest design. The design was selected because the study aimed to examine measurable changes in EFL students' analytical thinking before and after the implementation of ChatGPT-assisted critical reading instruction. In this design, the same group of students completed a pretest, participated in six meetings of structured instructional treatment, and completed a posttest after the intervention. Although the absence of a control group restricts causal generalization, the design was appropriate for identifying score differences associated with the instructional treatment and for providing preliminary evidence regarding the pedagogical value of ChatGPT in critical reading instruction (Kohnke et al., 2023; Lo et al., 2024).

Research Setting and Participants

The study was conducted in the Critical Reading course of the English Education Study Program at Universitas Muhammadiyah Parepare. This context represents an EFL higher education setting in which students are required to read academic texts, evaluate arguments, examine evidence, and construct reasoned responses in English. The participants were 10

fourth-semester students enrolled in the selected class. Total sampling was used because all students in the class were involved as research participants. The participants were relevant to the study because they were actively engaged in critical reading activities and were situated in an academic environment where generative AI tools have increasingly become part of language learning practices (C. Meniado, 2023; Zaimoğlu & Dağtaş, 2025).

Instructional Intervention and ChatGPT Integration

The intervention was implemented over six meetings through lecturer-guided ChatGPT-assisted critical reading activities. ChatGPT was used as a dialogic scaffold, not as a substitute for students' independent reasoning. Each meeting followed a consistent instructional sequence: (1) the lecturer introduced the reading objective and analytical focus; (2) students read an academic text or selected passage; (3) students interacted with ChatGPT using prompt-guided tasks; (4) students compared ChatGPT-generated responses with textual evidence; (5) the lecturer facilitated discussion on the accuracy, relevance, and limitations of the AI output; and (6) students revised their analytical responses based on their own judgment.

The six meetings were organized progressively. Meeting 1 introduced responsible use of ChatGPT, critical reading indicators, and prompt-writing procedures. Meeting 2 focused on identifying main ideas, authorial claims, and arguments. Meeting 3 addressed facts, opinions, assumptions, and claims. Meeting 4 trained students to evaluate supporting evidence and the strength of reasoning. Meeting 5 emphasized inference, alternative perspectives, and possible bias. Meeting 6 required students to synthesize ideas across text sections and produce evidence-based analytical responses. Examples of prompts used during the intervention included: "Identify the writer's main claim and explain the textual evidence that supports it"; "List possible assumptions behind the author's argument and justify each with clues from the text"; "Ask three questions that help evaluate the strength of the evidence"; and "Provide an alternative interpretation of the argument, then compare it with the information in the text." Students were required to treat ChatGPT responses as provisional suggestions and to verify them through close reading. This procedure reflects sociocultural views of learning tools as mediators of cognitive development when used purposefully under guidance (Alkhudiry, 2022) and aligns with calls for pedagogical control in AI-assisted language learning (Kohnke et al., 2023; Slamet, 2024).

Research Instruments and Assessment Rubric

The main instrument was an analytical thinking test embedded in critical reading tasks and administered as both pretest and posttest. The tests required students to analyze academic texts by identifying main ideas and arguments, distinguishing facts from opinions, assumptions, and claims, evaluating evidence, making logical inferences, comparing perspectives, connecting ideas across text sections, and constructing evidence-based responses. The pretest and posttest were developed using comparable task formats and the same analytical indicators to allow meaningful comparison of students' performance before and after treatment.

A scoring rubric was used to ensure systematic and transparent assessment. The rubric consisted of five major criteria: (1) identification of main ideas and arguments, (2) distinction among facts, opinions, assumptions, and claims, (3) evaluation of evidence and reasoning, (4) inference and interpretation, and (5) construction of evidence-based analytical responses. Each criterion was rated on four performance levels, ranging from limited, developing, proficient, to advanced performance, and the total score was converted into a 0-100 scale. The rubric descriptors guided raters in assigning scores consistently and reduced subjectivity in evaluating open-ended analytical responses.

Data Collection Procedures

Data collection was conducted in several stages. First, the researcher developed the pretest, posttest, scoring rubric, and ChatGPT-assisted reading tasks based on the objectives of critical reading instruction. Second, the instruments were reviewed through expert judgment to ensure content relevance and alignment with analytical thinking indicators. Third, the students completed the pretest to determine their initial analytical thinking ability. Fourth, the researcher implemented six meetings of ChatGPT-assisted critical reading instruction. Fifth, after the intervention, the students completed the posttest to measure their analytical thinking after treatment. Finally, the pretest and posttest scores were collected, coded, and prepared for statistical analysis.

Data Analysis Techniques

The data were analyzed using descriptive and inferential statistics. Descriptive statistics were used to calculate the mean, standard deviation, minimum score, maximum score, and percentage of improvement. A normality test was conducted to determine the appropriate inferential procedure (Downie, 2016). If the data were normally distributed, a paired sample t-test was used to examine whether there was a significant difference between pretest and posttest scores (Ledolter et al., 2020). If the data were not normally distributed, the Wilcoxon signed-rank test was used as a non-parametric alternative (Nahm, 2016). Cohen's d was calculated to determine the effect size of the intervention, while gain score or normalized gain analysis was used to classify the level of students' improvement after the ChatGPT-assisted instruction (Samaray, 2025).

Validity, Reliability, and Ethical Considerations

Content validity was established through expert judgment by specialists in English language teaching, critical reading, and educational assessment. The experts reviewed the alignment among the test tasks, rubric criteria, intervention objectives, and analytical thinking indicators. Revisions were made when instructions, prompt wording, text difficulty, or rubric descriptors required clarification. Reliability was examined through internal consistency analysis using Cronbach's alpha where item-level data were applicable. For rubric-based scoring, inter-rater reliability was considered by comparing raters' scores and resolving discrepancies through discussion until scoring agreement was achieved (Musikin et al., 2023; Nguyen & Trinh, 2025). Ethical principles were maintained by obtaining informed consent, protecting students' identities, using data only for research purposes, and

reporting results anonymously. Students were also instructed not to enter personal or sensitive information into ChatGPT.

RESULTS

Overview of the Results

This section presents the quantitative results of the study by comparing students' analytical thinking scores before and after the implementation of ChatGPT-assisted critical reading instruction. The analysis was conducted using pretest and posttest scores obtained from 10 fourth-semester EFL students enrolled in the Critical Reading course at Universitas Muhammadiyah Parepare. The results are organized into descriptive statistics, score improvement, normality testing, paired sample t-test results, effect size, and a summary of the main findings. These analyses were used to determine whether students' analytical thinking in critical reading improved after participating in structured learning activities supported by ChatGPT, as described in the research methodology.

Descriptive Statistics of Students' Analytical Thinking Scores

The descriptive statistics of students' pretest and posttest scores are presented in Table 1. The table includes the number of participants, mean score, standard deviation, minimum score, and maximum score for both tests.

Table 1. Descriptive statistics of pretest and posttest scores

Test	N	Mean	SD	Minimum	Maximum
Pretest	10	58,50	3,03	54	63
Posttest	10	74,50	5,60	66	82

Table 1 shows that the students' mean score increased from the pretest to the posttest. The pretest mean was 58.50, while the posttest mean reached 74.50, indicating an average improvement of 16.00 points after the instructional intervention. The minimum score also increased from 54 in the pretest to 66 in the posttest, while the maximum score rose from 63 to 82. These descriptive results indicate that students' analytical thinking performance in critical reading tasks was higher after the implementation of ChatGPT-assisted instruction.

The standard deviation also increased from 3.03 in the pretest to 5.60 in the posttest. This suggests that while the overall scores improved, the distribution of students' posttest scores became slightly more varied. Nevertheless, the posttest scores were consistently higher than the pretest scores, indicating a positive pattern of improvement across participants.

Improvement in Students' Analytical Thinking

To examine individual improvement, gain scores and normalized gain scores were calculated for each student. The results are shown in Table 2.

Table 2. Students' pretest, posttest, and gain scores

Student	Pretest	Post test	Gain Score	N-Gain
S1	54	66	12	0.26
S2	57	74	17	0.40
S3	60	78	18	0.45
S4	62	73	11	0.29
S5	58	79	21	0.50
S6	55	70	15	0.33
S7	61	81	20	0.51
S8	59	75	16	0.39
S9	56	67	11	0.25
S10	63	82	19	0.51

As shown in Table 2, all students demonstrated higher scores in the posttest than in the pretest. The gain scores ranged from 11 to 21 points, indicating that every participant experienced measurable improvement after the ChatGPT-assisted critical reading instruction. The highest gain score was obtained by S5, with an increase of 21 points, while the lowest gain scores were recorded by S4 and S9, each with an increase of 11 points. The mean gain score was 16.00. In addition, the average normalized gain was 0.39, which falls into the moderate improvement category. This result indicates that students' analytical thinking improved after the intervention, although the level of improvement remained moderate when measured using normalized gain analysis.

Normality Test

Before conducting the inferential test, the normality of the pretest, posttest, and difference scores was examined using the Shapiro-Wilk test. This test was selected because the number of participants was relatively small.

Table 3. Shapiro-Wilk normality test

Variable	Statistic	df	Sig.
Pretest	0.970	10	0.892
Posttest	0.948	10	0.649
Difference Score	0.922	10	0.375

Table 3 shows that the significance values for the pretest, posttest, and difference score were all higher than 0.05. The pretest obtained a significance value of 0.892, the posttest obtained 0.649, and the difference score obtained 0.375. These results indicate that the data were normally distributed. Therefore, the use of a paired sample t-test was considered appropriate for examining the difference between pretest and posttest scores.

Paired Sample t-test

A paired sample t-test was conducted to determine whether the difference between students' pretest and posttest scores was statistically significant. The result is presented in Table 4.

Table 4. Paired sample t-test result

Comparison	Mean Difference	t	df	Sig. (2-tailed)
Posttest – Pretest	16.00	13.74	9	< 0.001

The result of the paired sample t-test showed a statistically significant difference between the pretest and posttest scores, $t(9) = 13.74$, $p < 0.001$. Since the p-value was lower than the significance level of 0.05, the difference between the two sets of scores was statistically significant. This finding indicates that students' analytical thinking scores increased significantly after they participated in ChatGPT-assisted critical reading instruction.

Effect Size

To determine the magnitude of the improvement, Cohen's d was calculated. The result is shown in Table 5.

Table 5. Effect size of the intervention

Effect Size Measure	Value	Interpretation
Cohen's d	04.35	Very large effect

The Cohen's d value of 4.35 indicates a very large effect. Based on the general interpretation of Cohen's d, values of 0.20, 0.50, and 0.80 represent small, medium, and large effects, respectively. Therefore, the obtained value suggests that the difference between pretest and posttest scores was substantial.

Summary of Findings

The results of the study indicate that students' posttest scores were higher than their pretest scores after the implementation of ChatGPT-assisted critical reading instruction. All 10 students showed improvement, with gain scores ranging from 11 to 21 points. The average normalized gain was 0.39, indicating moderate improvement in students' analytical thinking. The Shapiro-Wilk test showed that the data were normally distributed, allowing the use of a paired sample t-test. The inferential analysis revealed a statistically significant difference between pretest and posttest scores, $t(9) = 13.74$, $p < 0.001$. In addition, the effect size was very large, with Cohen's d = 4.35. Overall, the findings suggest that the integration of ChatGPT into critical reading instruction was associated with an improvement in EFL students' analytical thinking.

DISCUSSION

The results show a clear improvement in students' analytical thinking after the six-meeting ChatGPT-assisted critical reading intervention. Before the treatment, students' pretest mean score was 58.50, suggesting that their ability to analyze academic texts was still limited. After the intervention, the mean score increased to 74.50, with all ten students achieving higher posttest scores. This pattern, supported by a moderate normalized gain (0.39) and a significant paired-sample t-test result, indicates that the improvement was not incidental. More importantly, the findings suggest that lecturer-guided use of ChatGPT helped students engage with reading texts more critically by analyzing arguments, checking evidence, questioning assumptions, and developing more reasoned responses.

The improvement can be understood through the nature of the tasks implemented during the intervention. Students were not asked to use ChatGPT only to summarize passages or produce ready-made answers. Instead, they were guided to use AI-generated responses as provisional inputs that had to be examined, questioned, compared with textual evidence, and revised. This process is central to critical reading because students must identify claims, distinguish facts from opinions and assumptions, evaluate evidence, make inferences, recognize bias, and formulate evidence-based responses (Dorgham & Obiad, 2025; Niculescu & Dragomir, 2023). In this sense, ChatGPT appears to have supported analytical thinking not by replacing students' reasoning, but by stimulating a cycle of questioning, verification, and reflection. The tasks also correspond to Bloom's revised taxonomy because students were encouraged to progress from understanding the text to analyzing arguments, evaluating evidence, and producing justified interpretations (Bolat, 2023).

For EFL learners, this support is pedagogically important. Critical reading in English requires students to manage vocabulary, academic discourse, argumentative structure, and disciplinary ideas at the same time. ChatGPT may reduce some of these linguistic burdens by providing clarification, alternative explanations, guiding questions, and examples. Yet the contribution of the tool lies not only in making texts easier to understand, but in enabling students to allocate more cognitive attention to higher-order operations such as comparing perspectives, detecting weak reasoning, and connecting evidence with claims. This interpretation is consistent with sociocultural perspectives that view learning tools as mediators of cognitive development when they are used purposefully under instructional guidance (Alkhudiry, 2022; Kasneci et al., 2023). It also supports the argument that generative AI is most educationally valuable when students are trained to interrogate its output rather than accept it passively (Kohnke et al., 2023; Liang & Wu, 2024a; Shang et al., 2025).

The findings align with earlier studies showing that ChatGPT and other AI-assisted tools can promote language learning through immediate feedback, adaptive interaction, personalized support, learner autonomy, and engagement (C. Meniado, 2023; Kohnke et al., 2023; Lo et al., 2024). They also support research indicating that generative AI can help students develop ideas, improve performance, and receive scaffolded assistance when tasks are pedagogically meaningful (Akinsemolu & Onyeaka, 2025; Mahapatra, 2024; Shafiee Rad,

2025). Nevertheless, the present study extends this literature by focusing specifically on critical reading and analytical thinking, areas that remain less frequently examined than writing, grammar, vocabulary, and general language practice (Lo et al., 2024). The result therefore contributes contextual evidence from Indonesian EFL higher education and shows that ChatGPT can be positioned as a dialogic scaffold for argument analysis, evidence evaluation, and reflective interpretation.

The study also has implications for EFL teaching practice. Lecturers should not integrate ChatGPT as an answer-generating tool, but as part of a carefully designed sequence of reading, prompting, verification, discussion, and revision. Practical activities may include asking students to compare ChatGPT's interpretation with textual evidence, evaluate whether an AI-generated claim is supported by the passage, identify assumptions in both the author's argument and the AI response, generate counterquestions, and revise their analysis after peer and lecturer feedback. Such activities can strengthen students' critical AI literacy, which is increasingly necessary because AI outputs may be incomplete, biased, inaccurate, or overgeneralized (Kasneci et al., 2023; Pujiastuti et al., 2025). Thus, ChatGPT integration should be accompanied by explicit instruction on evidence checking, source awareness, ethical use, and independent judgment.

Despite these contributions, the findings should be interpreted cautiously. The one-group pretest-posttest design does not allow the improvement to be attributed solely to ChatGPT. Students' progress may also have been influenced by lecturer guidance, repeated exposure to critical reading tasks, familiarity with the test format, motivation, or regular classroom practice. The small sample size also limits generalizability, and the very large effect size should be read as preliminary evidence rather than definitive proof of effectiveness. Future studies should involve larger samples, control or comparison groups, qualitative data, and longer interventions to examine how students actually interact with ChatGPT during critical reading. Overall, the study suggests that ChatGPT can support EFL students' analytical thinking when it is used as a guided, reflective, and evidence-based learning tool rather than as a substitute for critical reading itself.

CONCLUSION

The study examined the integration of ChatGPT into critical reading instruction and its association with EFL students' analytical thinking in a higher education context. Within a quantitative pre-experimental one-group pretest-posttest design, the study involved ten fourth-semester students in the Critical Reading course at the English Education Study Program of Universitas Muhammadiyah Parepare. The findings indicate that students' analytical thinking scores improved after the ChatGPT-assisted instructional intervention, as reflected in the increase from the pretest mean score of 58.50 to the posttest mean score of 74.50, with a mean gain of 16.00. All participants showed score improvement, while the normalized gain of 0.39 indicated a moderate level of improvement. The paired sample t-test also showed a statistically significant difference between pretest and posttest scores, $t(9) = 13.74$, $p < 0.001$, and Cohen's d of 4.35 suggested a very large effect.

These results suggest that ChatGPT-assisted critical reading instruction was associated with improved analytical thinking among the participating EFL students. The study therefore

addresses its objective by showing that, when used through structured and lecturer-guided activities, ChatGPT may support students in identifying arguments, examining evidence, making inferences, comparing perspectives, and developing more analytical responses to texts. Theoretically, the study contributes to the growing discussion on AI-assisted language learning by positioning ChatGPT not merely as a comprehension aid, but as a dialogic scaffold for higher-order reading engagement. Practically, the findings imply that EFL lecturers can integrate ChatGPT into critical reading instruction through guided prompts, evidence-checking activities, reflective questioning, and evaluation of AI-generated responses.

Within the limitations of this study, however, the findings should be interpreted cautiously. The use of a single group without a control group and the small number of participants limit the extent to which the improvement can be attributed exclusively to ChatGPT or generalized to broader EFL populations. Other instructional, motivational, or contextual factors may also have contributed to the observed gains. Further research is needed using larger samples, control or comparison groups, mixed-methods designs, longitudinal data, and analysis of students' written analytical responses. Overall, the findings suggest that ChatGPT has promising pedagogical value in EFL critical reading instruction, provided that its use is carefully designed, ethically guided, and directed toward strengthening not replacing students' independent analytical thinking.

ACKNOWLEDGMENT

The author extends heartfelt appreciation to the English Education Study Program of Universitas Muhammadiyah Parepare for the support and opportunity provided throughout the conduct of this research. Sincere thanks are also addressed to the fourth-semester students who actively took part in the Critical Reading course and contributed meaningful data to this study. The author is equally grateful to colleagues, expert validators, and academic supervisors for their valuable suggestions, guidance, and encouragement during the research process. Lastly, the author acknowledges the integration of ChatGPT as a learning support tool in this study, which assisted the critical reading activities while encouraging students to maintain independent analytical thinking.

REFERENCES

- Akbar, F. F., Wilany, E., & Siahaan, S. (2025). Higher-order thinking skills (HOTS) in reading comprehension of EFL class. *Celtic : A Journal of Culture, English Language Teaching, Literature and Linguistics*, 12(2), 730–753. <https://doi.org/10.22219/celtic.v12i2.41568>
- Akinsemolu, A. A., & Onyeaka, H. (2025). The Role of Artificial Intelligence in Transforming Language Learning: Opportunities and Ethical Considerations. *Journal of Language and Education*, 11(1), 148–152. <https://doi.org/10.17323/jle.2025.22118>
- Alkhudiry, R. (2022). The Contribution of Vygotsky's Sociocultural Theory in Mediating L2 Knowledge Co-Construction. *Theory and Practice in Language Studies*, 12(10), 2117–2123. <https://doi.org/10.17507/tpls.1210.19>

- Bolat, H. (2023). Examining the Learning Outcomes Included in Science and Art Centers' Summer School Support and Development Course of Thinking Education Workshop Program According to Bloom's Taxonomy. *Open Journal for Educational Research*, 7(2), 93–106. <https://doi.org/10.32591/coas.ojer.0702.03093b>
- C. Meniado, J. (2023). The Impact of ChatGPT on English Language Teaching, Learning, and Assessment: A Rapid Review of Literature. *Arab World English Journal*, 14(4), 3–18. <https://doi.org/10.24093/awej/vol14no4.1>
- Dorgham, R., & Obiad, L. Bin. (2025). Utilizing an Instructional Higher-Order Thinking-Based Strategy to Improve EFL Critical Reading Skills. *Journal of Language Teaching and Research*, 16(4), 1373–1382. <https://doi.org/10.17507/jltr.1604.30>
- Downie, N., & S. A. (2016). Descriptive and Inferential Statistics. In *Quantitative Geography: The Basics* (pp. 75–95). SAGE Publications Ltd. <https://doi.org/10.4135/9781473920446.n5>
- Dresel, M., Schmitz, B., Schober, B., Spiel, C., Ziegler, A., Engelschalk, T., Jöstl, G., Klug, J., Roth, A., Wimmer, B., & Steuer, G. (2015). Competencies for successful self-regulated learning in higher education: structural model and indications drawn from expert interviews. *Studies in Higher Education*, 40(3), 454–470. <https://doi.org/10.1080/03075079.2015.1004236>
- Eragamreddy, Dr. N. (2025). Interactive AI-Driven Storytelling for Language Development. *International Journal of Social Science Humanity & Management Research*, 04(02). <https://doi.org/10.58806/ijsshmr.2025.v4i2n08>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, 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>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Ledesma Acosta, B. V., Martínez Mora, G. G., Medina Castro, G. J., & Lozano Alvarado, C. I. (2025). Project-Based learning for Academic reading: Enhancing higher-order thinking Skills in EFL Learners. *Revista Científica Multidisciplinar G-Nerando*, 6(1). <https://doi.org/10.60100/rcmg.v6i1.538>
- Ledolter, J., Gramlich, O. W., & Kardon, R. H. (2020). Parametric Statistical Inference for Comparing Means and Variances. *Investigative Ophthalmology & Visual Science*, 61(8), 25. <https://doi.org/10.1167/iovs.61.8.25>

- Liang, W., & Wu, Y. (2024). Exploring the use of ChatGPT to foster EFL learners' critical thinking skills from a post-humanist perspective. *Thinking Skills and Creativity*, 54, 101645. <https://doi.org/10.1016/j.tsc.2024.101645>
- Liu, X., & Puteh, M. (2025). A Systematic Literature Review on Enhancing Critical Thinking Skills in EFL Reading. *World Journal of English Language*, 15(8), 239. <https://doi.org/10.5430/wjel.v15n8p239>
- Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: a systematic review of empirical studies. *Smart Learning Environments*, 11(1), 50. <https://doi.org/10.1186/s40561-024-00342-5>
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: a mixed methods intervention study. *Smart Learning Environments*, 11(1), 9. <https://doi.org/10.1186/s40561-024-00295-9>
- Musikin, M., Mohd Matore, M. E. a. k. a. E., & Mohamad, N. (2023). Expert Evaluation on the Content Validity of the Novice Teachers' Assessment Inventory (InPGN). *Environment-Behaviour Proceedings Journal*, 8(SI15), 107–112. <https://doi.org/10.21834/e-bpj.v8iSI15.5087>
- Nahm, F. S. (2016). Nonparametric statistical tests for the continuous data: the basic concept and the practical use. *Korean Journal of Anesthesiology*, 69(1), 8. <https://doi.org/10.4097/kjae.2016.69.1.8>
- Nguyen, T. B. T., & Trinh, T. H. (2025). A Proposed Instrument to Investigate University Students' Critical Reading Strategy Use in English Reading Comprehension. *VNU Journal of Foreign Studies*, 41(4), 151–169. <https://doi.org/10.63023/2525-2445/jfs.ulis.5480>
- Niculescu, B.-O., & Dragomir, I.-A. (2023). Critical Reading - A Fundamental Skill for Building 21st Century Literacy. *International Conference KNOWLEDGE-BASED ORGANIZATION*, 29(2), 215–220. <https://doi.org/10.2478/kbo-2023-0060>
- Pujiastuti, I., Damaianti, V. S., Mulyati, Y., Sastromihardjo, A., & Lestari, D. (2025). Ketergantungan penggunaan AI pada pendidikan tinggi: Ancaman terhadap keterampilan membaca teks akademik. *Diglosia: Jurnal Kajian Bahasa, Sastra, Dan Pengajarannya*, 8(2), 473–484. <https://doi.org/10.30872/diglosia.v8i2.1243>
- Samaray, S. (2025). Penerapan Artificial Intelligence-ChatGPT dalam Pembelajaran Matematika Diskrit. *SABER: Jurnal Teknik Informatika, Sains Dan Ilmu Komunikasi*, 3(1), 263–272. <https://doi.org/10.59841/saber.v3i1.2283>
- Shafiee Rad, H. (2025). Reinforcing L2 reading comprehension through artificial intelligence intervention: refining engagement to foster self-regulated learning. *Smart Learning Environments*, 12(1), 23. <https://doi.org/10.1186/s40561-025-00377-2>
- Shang, J., Huang, Y., Xu, M., Huang, Y., Shen, X., Wang, G., Wang, Y., & Zhang, L. (2025). Competition between human learners and ChatGPT: enhancing university EFL students'

reading comprehension and critical thinking through competitive questioning.
Computer Assisted Language Learning, 1–35.
<https://doi.org/10.1080/09588221.2025.2577356>

Shoufan, A. (2023). Exploring Students' Perceptions of ChatGPT: Thematic Analysis and Follow-Up Survey. *IEEE Access*, 11, 38805–38818.
<https://doi.org/10.1109/ACCESS.2023.3268224>

Slamet, J. (2024). Potential of ChatGPT as a digital language learning assistant: EFL teachers' and students' perceptions. *Discover Artificial Intelligence*, 4(1), 46.
<https://doi.org/10.1007/s44163-024-00143-2>

Stojanov, A. (2023). Learning with ChatGPT 3.5 as a more knowledgeable other: an autoethnographic study. *International Journal of Educational Technology in Higher Education*, 20(1), 35. <https://doi.org/10.1186/s41239-023-00404-7>

Zaimoğlu, S., & Dağtaş, A. (2025). Teacher Cognition and Practices in Using Generative AI Tools to Support Student Engagement in EFL Higher-Education Contexts. *Behavioral Sciences*, 15(9), 1202. <https://doi.org/10.3390/bs15091202>