



The Effectiveness of Reflective Interaction with ChatGPT in Teaching Argumentative Text Writing at Junior High School

Moh Taufiq Hamsah¹, Syamsudduha², Muhammad Saleh³

^{1,2,3} Pendidikan Bahasa, Universitas Negeri Makassar, Makassar

Article Info	Abstract
Received: 2026-04-07 Revised: 2026-05-19 Accepted: 2026-09-20 Keywords: <i>Reflective Interaction, Chatgpt, Writing, Argumentative Text, Indonesian Language Learning</i> DOI: 10.24256/ideas.v14i2.11555 Corresponding Author: Moh Taufiq Hamsah mohtaufiq.240001301030@student.unm.ac.id Pendidikan Bahasa, Universitas Negeri Makassar, Makassar	<i>This study aimed to describe the results of argumentative text writing instruction through reflective interaction with peers, describe the results of argumentative text writing instruction through reflective interaction with ChatGPT, and examine the effectiveness of reflective interaction with ChatGPT in teaching argumentative text writing to students at SMP Negeri 8 Makassar. In this study, reflective interaction with ChatGPT refers to a dialogic process between students and ChatGPT in evaluating writing, obtaining feedback, understanding errors, and gradually and independently improving arguments. This study employed a quantitative approach using a quasi-experimental design in the form of a nonequivalent control group design. The sample consisted of 80 students, including 40 students from class IX.12 as the experimental class, which received treatment through reflective interaction with ChatGPT, and 40 students from class IX.10 as the control class, which received treatment through reflective interaction with peers. Data were collected through an argumentative text writing test and analyzed using descriptive statistics and the Mann-Whitney U test. The results showed that the mean score of the control class was 85.00, which was categorized as good, while the mean score of the experimental class was 92.88, which was categorized as very good. The Mann-Whitney U test showed a significance value of $0.004 < 0.05$, indicating a significant difference in students' argumentative writing results between the two classes. These findings indicate that ChatGPT can serve as a digital reflective partner that helps students evaluate content, clarify arguments, improve writing</i>

structure, and develop critical thinking skills in writing instruction at the junior high school level. Therefore, reflective interaction with ChatGPT provides a practical contribution to the development of AI-assisted writing instruction that is more interactive, responsive, ethical, and centered on students' learning needs.

1. Introduction

Writing is one of the essential language skills in Indonesian language learning because it enables students to express ideas, opinions, and feelings in a structured manner. Writing skills are not acquired naturally, but through a continuous process of learning, practice, feedback, and revision (Septiwi et al., 2024). This ability requires mastery of linguistic aspects, such as word choice, sentence structure, spelling, and paragraph coherence, as well as cognitive ability to organize ideas logically so that the writing can be understood by readers (Indah et al., 2021; Yasmin et al., 2025).

Writing argumentative texts is a skill that trains students to express opinions logically and critically, supported by relevant evidence. In argumentative texts, students need to understand the main elements, namely claims, evidence, and reasoning, so that the opinions presented can be rationally accepted (Candara, 2024). Students also need to understand the stages of composing argumentative texts, starting from introducing the topic, presenting arguments, using evidence, and drawing convincing conclusions Hidayati, (2023).

Various studies have shown that students' writing skills still face several challenges. Sari et al. (2020) found that 42.9% of students at SMPN 14 Kota Bengkulu had not achieved the minimum mastery criterion in writing skills. The difficulties experienced by students included problems in starting their writing, determining topics and ideas, limited vocabulary, inappropriate use of terms, incoherent sentence construction, weak logical thinking skills, and low reading experience. Other findings also revealed that approximately 60% of students had not met the minimum standard in writing, particularly in the aspects of vocabulary use, idea development, and language accuracy, as reported by Yulianti et al. (2024).

These problems indicate that students need a learning support tool that can help them develop ideas, improve writing structure, and evaluate their writing independently. One alternative that can be utilized is artificial intelligence technology in the form of ChatGPT. ChatGPT can be used as a learning assistant to help students understand concepts, obtain information, develop ideas, and receive concise explanations (Ramadhan et al., 2024). In writing instruction, ChatGPT can serve as a medium for reflective interaction between students and technology, allowing students to evaluate their writing, obtain feedback, and make gradual improvements (Rahmawati et al., 2025). These problems show that writing instruction requires a strategy that not only assigns writing tasks, but also helps

students evaluate and improve the thinking process behind their writing.

In this context, reflection becomes important because it helps students understand the weaknesses in their writing and determine appropriate improvement strategies. Dewey, as cited in Suphari and Chinokul (2023), views reflection as an active, continuous, and careful thinking process that occurs when individuals encounter a problem and analyze the experience to gain new understanding. Meanwhile, Schön, as cited in Marlina et al. (2023), distinguishes reflection into reflection-in-action, which occurs while a person is engaged in an activity, and reflection-on-action, which takes place after an activity has been completed in order to evaluate the results and determine subsequent actions. In writing instruction, both forms of reflection can occur when students review their writing, understand the reasons for its weaknesses, and revise it based on the feedback they receive.

Reflective interaction with ChatGPT allows students to understand errors in their writing, enrich their ideas, examine the logic of their arguments, and improve the structure of their texts more independently. ChatGPT can also help students improve grammar, review the coherence of arguments, and critically assess the quality of their writing (Werdiningsih et al. 2024). This process is in line with the concept of self-regulated learning, which refers to students' ability to regulate, monitor, and evaluate their own learning process (Teng & Zhang, 2020). Technological support in the reflection process can also increase students' learning engagement because the feedback provided is immediate, personalized, and non-judgmental Sari & Rahman, (2024).

According to Alexander (2025), ChatGPT functions as a reflective partner that can facilitate students' reflective thinking processes through feedback, critical questions, and the review of ideas. ChatGPT provides students with opportunities to continuously improve their writing through practice and personalized feedback that is suited to individual needs. Thus, students not only learn better writing techniques but also develop their critical thinking skills throughout the writing process.

According to Sok and Heng (2023), reflective interaction is also a form of new pedagogical practice resulting from the use of AI in education. Sok and Heng explain that ChatGPT can support reflective learning practices through automatic feedback, error analysis, and dialogue-based guidance, enabling both teachers and students to introspect on the learning strategies they use. Therefore, the success of reflective interaction largely depends on users' ability to interpret and follow up on the information provided by ChatGPT.

According to Robillos (2024), ChatGPT serves as a learning partner that assists students in writing, from planning and drafting to revising and editing. This is in line with Li et al. (2023), who state that ChatGPT is not merely a text-based program, but functions as an artificial cognitive agent capable of imitating the way

humans think and write reflectively. Li et al. further emphasize that ChatGPT can generate texts that resemble human writing, with logical, cohesive, and well-reasoned argumentative structures.

Artificial intelligence-assisted learning, or AI-assisted learning, is a learning approach that utilizes artificial intelligence technology to support students' learning processes through the provision of information, guidance, feedback, and more personalized learning assistance. In the educational context, the use of large language models such as ChatGPT creates opportunities for students to receive support in understanding materials, developing ideas, composing texts, and revising learning outcomes in a more interactive manner. Kasneci et al. (2023) explain that large language models have the potential to support more personalized learning experiences, enhance learning interaction, and assist teachers and students in reading, writing, and critical thinking activities. However, the use of AI in learning still needs to be pedagogically directed so that students do not merely accept answers from technology, but are also able to assess, examine, and use feedback critically. This idea is in line with (Mollick and Mollick 2023), who position AI as a tutor, coach, mentor, or learning aid that still requires active human involvement as the controller of the learning process.

Formative feedback refers to information provided to learners during the learning process with the aim of helping them improve their understanding, strategies, or learning outcomes. Shute (2008) explains that formative feedback is information communicated to learners in order to change their thinking or learning behavior so that learning improvement can occur. Effective feedback should be specific, timely, supportive, and not merely judgmental of learning outcomes. In line with this view, Hattie and Timperley (2007) emphasize that feedback is one of the most influential factors in student learning and achievement, although its effectiveness depends on the type, timing, and manner in which it is delivered.

In writing instruction, self-regulated learning can be seen when students are able to plan the content of their writing, monitor the relevance of their arguments, evaluate the strength of their evidence, and revise their texts based on the feedback they receive. Panadero (2017) explains that self-regulated learning includes cognitive, metacognitive, behavioral, motivational, and emotional aspects of the learning process. Students' ability to use ChatGPT reflectively can therefore be understood as a form of independent learning, because students do not merely receive input from AI, but also interpret, select, and apply that input to improve their writing. Teng and Zhang (2020) also emphasize that self-regulated learning strategies in writing can help learners develop the ability to plan, monitor, and evaluate the writing process more consciously

The next comparison between previous studies and the present study is the study conducted by Khampusaen (2025), which shares similarities with this study because both examine the use of ChatGPT in developing argumentative writing skills. Both studies show that ChatGPT can help learners improve the quality of

their writing, particularly in developing arguments, organizing texts, integrating evidence, and using language effectively. However, there are several fundamental differences between Khampusaen's study and the present study. Khampusaen's study was conducted with EFL university students using a mixed-methods design, whereas the present study was conducted with ninth-grade students at SMP Negeri 8 Makassar using a quantitative quasi-experimental approach. In addition, Khampusaen's study emphasized the development of writing drafts, students' perceptions, and academic writing skills, while the present study focuses on the effectiveness of reflective interaction with ChatGPT in teaching argumentative text writing.

Yasmin et al. (2025) also share similarities with the present study because both employ a quasi-experimental approach and focus on argumentative writing skills. However, Yasmin et al.'s study was conducted in an English as a Foreign Language context with students in Pakistan, whereas the present study was conducted in the context of Indonesian language learning with students at SMP Negeri 8 Makassar. In addition, Yasmin et al.'s study placed greater emphasis on improving argumentative writing skills and error correction, while the present study emphasizes the use of ChatGPT as a reflective interaction partner in the process of evaluating and revising students' writing.

Based on the explanation above, this study focuses on the use of ChatGPT as a reflective tool in teaching argumentative text writing. Previous studies have widely discussed the use of ChatGPT in writing instruction at the senior high school and university levels, while its application at the junior high school level remains limited. This gap provides an important basis for conducting the present study. Therefore, this study aims to investigate the effectiveness of using ChatGPT through reflective interaction in improving junior high school students' ability to write argumentative texts. The findings of this study are expected to provide an alternative approach to writing instruction that is more interactive, responsive, and aligned with students' learning needs.

Reflective interaction with ChatGPT in this study is understood as a two-way communication process between students and ChatGPT to evaluate, analyze, and revise writing. Students remain the primary writers, while ChatGPT functions as a provider of feedback, a prompt for reflective questions, and a revision companion. Thus, students do not merely receive answers from ChatGPT; they interpret the feedback, consider its relevance, and improve their writing based on their own awareness. This distinction differentiates reflective interaction with ChatGPT from the passive use of ChatGPT to generate texts.

2. Method

The population of this study consisted of 482 ninth-grade students divided into 12 classes, while the research sample consisted of class IX.12, comprising 40 students as the experimental group, and class IX.10, comprising 40 students as the

control group. The experimental group received treatment in the form of reflective interaction with ChatGPT, while the control group used reflective interaction with peers. The sampling process was conducted through field observation and interviews with the ninth-grade homeroom teachers. The researcher also requested students' attendance lists as the basis for determining the research population. Based on the attendance lists, it was found that the ninth grade consisted of 12 classes. These data were then used by the researcher as the basis for determining the required research sample. Data were collected through an argumentative text writing test. The research instrument used was an argumentative text writing assessment rubric covering five aspects: relevance of content to the topic, clarity of arguments and evidence, vocabulary use, sentence structure, and appropriateness of arguments. The data were analyzed using descriptive and inferential statistics. The normality test was conducted using the Shapiro-Wilk test with the assistance of SPSS version 31. If the data were normally distributed, hypothesis testing was conducted using a t-test. However, if the data were not normally distributed, the Mann-Whitney U test was used at a significance level of 0.05

The implementation of reflective interaction with ChatGPT in the experimental class was carried out in several stages. In the initial stage, students were given an explanation of argumentative texts, including their definition, structure, composition, use of evidence, and examples. After that, students were asked to write an initial draft of an argumentative text independently. The initial draft was then used as the basis for reflective interaction with ChatGPT.

The interaction with ChatGPT was conducted using prompts prepared by the researcher to ensure that the feedback process was directed and focused. One of the prompts used was: "Evaluate the following argumentative text based on five aspects: relevance of content to the topic, clarity of arguments and evidence, vocabulary use, sentence structure, and appropriateness of arguments. The assessment should use a scale of 1 to 4, where 1 means poor, 2 means fair, 3 means good, and 4 means very good. Assess each aspect objectively by providing a score, the reason for the assessment, and appropriate suggestions for improvement." Through this prompt, ChatGPT provided feedback in the form of identifying weaknesses in the writing, suggestions for improving arguments, and comments on text structure, word choice, and sentence clarity.

After receiving feedback, students were guided to read and examine the feedback from ChatGPT critically. Students were not allowed to directly copy the output from ChatGPT, but were asked to use the feedback as a basis for reflection to improve their writing. The revision process was carried out by improving the content, clarifying the arguments, adding relevant evidence, improving paragraph structure, and constructing more effective sentences. The researcher acted as a facilitator who guided students to use ChatGPT wisely and in accordance with the learning objectives

Each learning session was conducted for 2×40 minutes. The first session

was used to introduce the material and prepare the initial draft. The second session was used for interaction with ChatGPT and feedback acquisition. The third session was used for feedback analysis and student reflection. The fourth session was used for writing revision. The final session was used to collect the students' final writing products. The students' final writings were then assessed based on the predetermined aspects of argumentative text writing assessment.

The validity and reliability of the argumentative text writing assessment rubric were also considered in this study. The rubric was adapted and modified from Utami et al. (2024). It was used to assess five main aspects of argumentative texts: relevance of content to the topic, clarity of arguments and evidence, vocabulary use, sentence structure, and appropriateness of arguments. The validity of the rubric in this study was based on content validity, as the assessment aspects were aligned with the characteristics of argumentative texts, which require content relevance, clarity of arguments, supporting evidence, language accuracy, and coherence among arguments.

The reliability of rubric use was maintained by providing clear assessment descriptors for each aspect using a scale of 1 to 4, where 1 = poor, 2 = fair, 3 = good, and 4 = very good. The same rubric was used to assess students' writing in both the experimental and control classes to ensure that the assessment process was more directed and consistent. In the control class, the rubric was used in peer assessment activities as part of reflective interaction, while in the experimental class, it was used as the basis for writing evaluation through reflective interaction with ChatGPT.

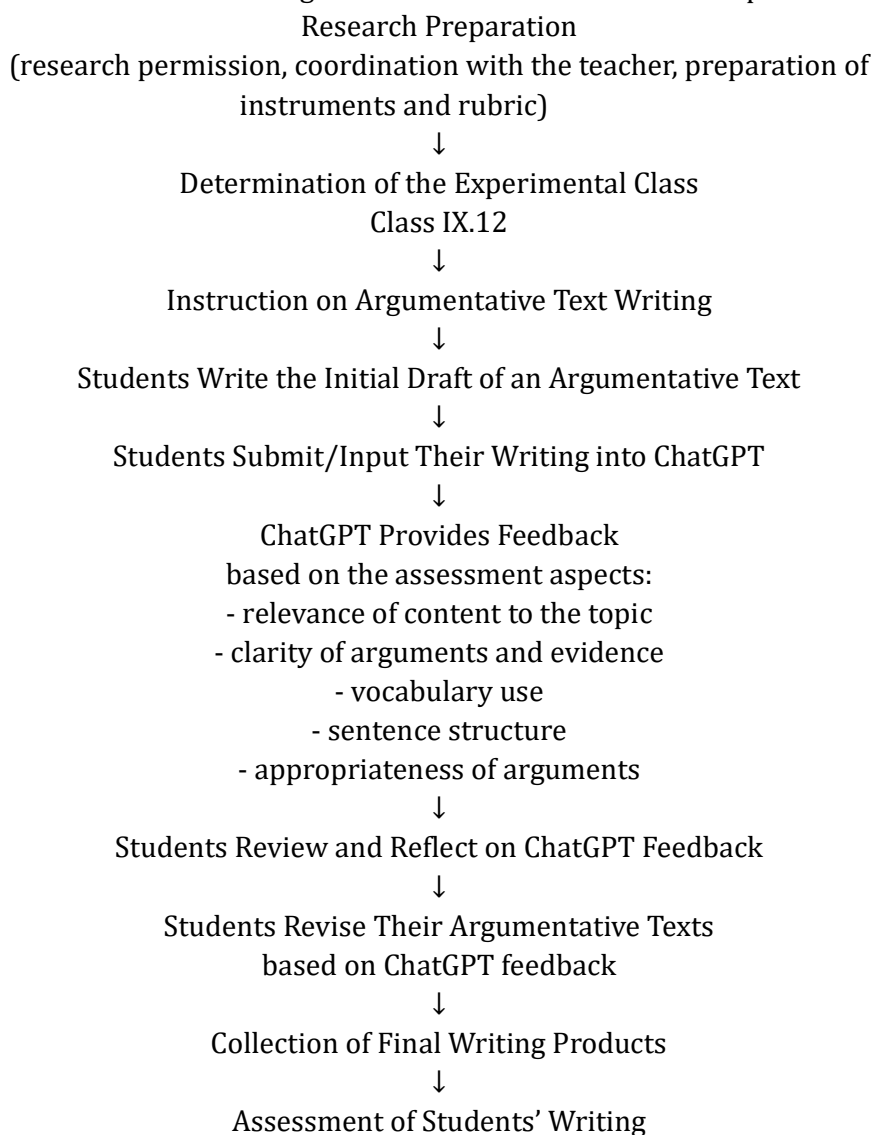
Peer reflection was implemented in the control class in this study. Students were first asked to write an argumentative text based on the structure of argumentative writing, namely the introduction, body, and conclusion. The introduction presented the topic or the initial statement of opinion. The body contained reasons, evidence, facts, or examples that supported the opinion, while the conclusion restated the opinion that had been presented.

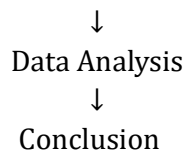
After writing their argumentative texts, students' writings were evaluated through reflective interaction with peers. At this stage, students read one another's texts and provided assessments based on the prepared assessment criteria. The evaluation was not conducted freely, but was guided by five main aspects: relevance of content to the topic, clarity of arguments and evidence, vocabulary use, sentence structure, and appropriateness of arguments. These assessment aspects provided students with clearer guidelines for reading and assessing their peers' writing. This process helped students identify whether the argumentative texts were relevant to the topic, contained clear arguments, were supported by relevant evidence, used appropriate vocabulary, and were constructed in effective sentences.

Ethical considerations in this study included school approval, informed consent, and the responsible use of AI by students. Based on interviews with teachers, the use of AI at SMP Negeri 8 Makassar is permitted in the learning

process; however, its implementation depends on each teacher's policy and is not used continuously. AI is used at certain times according to learning needs, particularly as a supporting tool for searching information and helping students understand the material. Thus, the school basically provides opportunities for the use of AI as a learning support medium, as long as its use remains guided and supervised by teachers.

The assessment in this study did not involve two independent raters to calculate inter-rater reliability statistically. In the control class, students conducted peer assessment using the predetermined assessment rubric. Students read one another's writing, assigned scores, and provided feedback based on the aspects of relevance of content to the topic, clarity of arguments and evidence, vocabulary use, sentence structure, and appropriateness of arguments. Meanwhile, in the experimental class, students used ChatGPT as a medium for reflective interaction to obtain feedback on their writing based on the same assessment aspects.





3. Result

The Results of Learning Argumentative Text Writing Through Reflective Interaction with Peers among Students of SMP Negeri 8 Makassar

The learning outcomes of writing argumentative texts through reflective interaction with peers are presented in the form of frequency and percentage distributions. The presentation of these data aims to identify the distribution of students' achievement scores based on the predetermined categories, ranging from very good, good, moderate, to poor. The description of the frequency and percentage of the learning outcomes can be seen in Table 1 below.

Table 1. Description of the Frequency and Percentage of the Results of Learning Argumentative Text Writing Through Reflective Interaction with Peers

No	Score Range	Category	Frequency	Percentage
1	91-100	Very Good	11	27,5 %
2	81-90	Good	15	37,5%
3	71-80	Moderate	9	22,5%
4	60-70	Poor	3	7,5%

Based on the description of frequency and percentage, it is known that out of a total of 40 samples, 11 students, or 27.5%, were in the very good category with a score range of 91–100. Furthermore, 15 students, or 37.5%, were in the good category with a score range of 81–90. This category had the highest frequency, indicating that most of the samples achieved good scores. In the moderate category, there were 9 students, or 22.5%, with a score range of 71–80. Meanwhile, the poor category had the lowest frequency, consisting of 3 students, or 7.5%, with a score range of 60–70. Thus, it can be concluded that the results of learning argumentative text writing through reflective interaction with peers were in the good category, with the highest frequency of 15 students and a percentage of 37.5%.

Table 2. Results of Descriptive Statistical Data Analysis of Argumentative Text Writing Learning Outcomes Through Reflective Interaction with Peers

Descriptive Statistics							
	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Error	Std. Deviation Statistic	Variance Statistic
Interaksi Reflektif dengan Teman Sejawat	40	50	100	85.00	2.010	12.710	161.538
Valid N (listwise)	40						

Based on the descriptive statistics table, it is known that the number of samples analyzed was 40 students. The minimum score for the reflective interaction with peers variable was 50, while the maximum score was 100. The mean score of 85.00 indicates that reflective interaction with peers was in the good

category. Therefore, it can be concluded that reflective interaction with peers in this study was categorized as good, with a mean score of 85.00.

The Results of Learning Argumentative Text Writing Through Reflective Interaction with ChatGPT among Students of SMP Negeri 8 Makassar

The learning outcomes of writing argumentative texts through reflective interaction with ChatGPT are presented in the form of frequency and percentage distributions. The presentation of these data aims to identify the distribution of students' achievement scores based on the predetermined categories, namely very good, good, moderate, and poor. The description of the frequency and percentage of the learning outcomes in writing argumentative texts through reflective interaction with ChatGPT can be seen in Table 2 below.

Table 3. Descriptive Frequency and Percentage of Argumentative Text Writing Learning Outcomes Through Reflective Interaction with ChatGPT

No	Rentang Nilai	Kategori	Frekuensi	Presentase
1	91-100	Very Good	20	50%
2	81-90	Good	19	47,5%
3	71-80	Moderate	1	2,5
4	60-70	Poor	-	-

Based on the descriptive frequency and percentage table of argumentative text writing learning outcomes through reflective interaction with ChatGPT, it was shown that 20 students, or 50%, obtained scores in the range of 91–100 and were categorized as very good. Furthermore, 19 students, or 47.5%, obtained scores in the range of 81–90 and were categorized as good. In the score range of 71–80, there was 1 student, or 2.5%, who was included in the moderate category. Meanwhile, no students obtained scores in the range of 60–70 or were categorized as poor. Therefore, it can be concluded that the results of learning argumentative text writing through reflective interaction with ChatGPT were categorized as very good, as most students were in the very good category, with a frequency of 20 students and a percentage of 50%.

Table 4. Results of Descriptive Statistical Data Analysis of Argumentative Text Writing Learning Outcomes Through Reflective Interaction with ChatGPT

Descriptive Statistics							
	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Error Std. Error	Std. Deviation Statistic	Variance Statistic
Interaksi Reflektif dengan ChatGPT	40	80	100	92.88	1.027	6.493	42.163
Valid N (listwise)	40						

Based on the descriptive statistics table, it is known that the number of samples analyzed was 40 students. The minimum score for the reflective interaction with ChatGPT variable was 80, while the maximum score was 100. The mean score of 92.88 indicates that the results of reflective interaction with ChatGPT were at a high level. Therefore, it can be concluded that the learning outcomes of

argumentative text writing through reflective interaction with ChatGPT were categorized as very good, with a mean score of 92.88.

Tabel 5. Descriptive Statistics of Argumentative Writing Scores in the ChatGPT Reflective Interaction and Peer Reflective Interaction Groups

Descriptive Statistics								
	N	Range	Minimum	Maximum	Sum	Mean	Std. Deviation	Variance
Interaksi Reflektif dengan Teman Sejawat	40	50.000	50	100	3400	85.00	12.710	161.538
Interaksi Reflektif dengan ChatGPT	40	20.000	80	100	3715	92.88	6.493	42.163
Valid N (listwise)	40							

Based on the descriptive statistics table, it is known that the number of participants in each group was 40 students. In the control class, which applied peer reflective interaction, the lowest score obtained by the participants was 50, while the highest score was 100, resulting in a score range of 50. The total score in this group was 3,400, with a mean score of 85.00. The standard deviation of 12.710 and variance of 161.538 indicate that the participants' scores in the peer reflective interaction group were relatively varied. This means that the participants' writing ability in this group was not entirely evenly distributed, as there was a considerable difference between students who obtained low scores and those who achieved high scores.

Normality Test Results

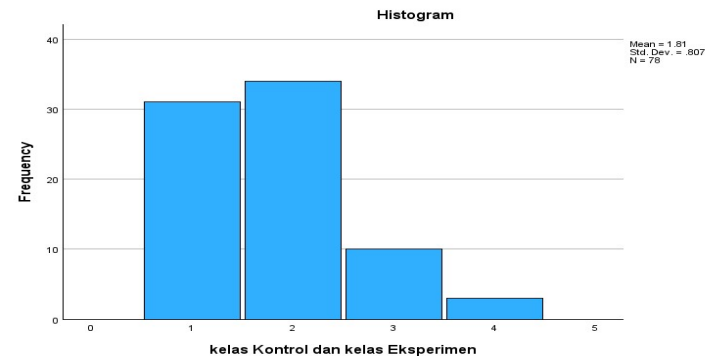
Table 6. Results of the normality test in the control class through reflective interaction with peers and the experimental class through reflective interaction with ChatGPT

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Interaksi reflektif dengan teman sejawat	.178	40	.003	.900	40	.002
Interaksi reflektif dengan ChatGPT	.239	40	<.001	.836	40	<.001

a. Lilliefors Significance Correction

Based on the normality test results in the *Tests of Normality* table, it is known that the number of samples in each class was 40 students. Since each class consisted of 40 students, the normality test was focused on the Shapiro-Wilk value. In the reflective interaction with peers class, the significance value was 0.002, while in the reflective interaction with ChatGPT class, the significance value was 0.001. These two values are smaller than 0.05. Therefore, it can be concluded that the data in both classes were not normally distributed.

Figure 1. Histogram of Argumentative Writing Scores in the Experimental and Control Classes



Based on the histogram of writing results in the control and experimental classes, the data distribution was mostly concentrated in the very good and good categories. This is indicated by the highest frequency in the good category, with 34 students, followed by the very good category, with 31 students. Meanwhile, 10 students were in the moderate category, and only 3 students were in the very poor category. The histogram shows that the data were not concentrated in a single category, but were distributed across several levels of ability. The mean score of 1.81, with a standard deviation of 0.807 and 78 valid data points, indicates that the variation in students’ writing ability was not very large. Therefore, it can be concluded that the writing ability of students in the control and experimental classes tended to be relatively homogeneous, with the results predominantly falling into the good and very good categories.

Mann-Whitney U Test Results

The Mann-Whitney U test is a non-parametric statistical test used to determine whether there is a significant difference between two independent groups, particularly when the data do not meet the assumption of normality or when the data are ordinal or continuous. Unlike the independent samples t-test, which compares means, the Mann-Whitney U test compares data ranks or the sum of ranks to determine whether one group tends to have higher or lower scores than the other group (Chicco et al., 2025).

The Mann-Whitney U test was conducted to compare the scores of the control class, which used reflective interaction with peers, and the experimental class, which used reflective interaction with ChatGPT in learning to write argumentative texts among ninth-grade students at SMP Negeri 8 Makassar. The results of the test are presented in Table 6.

Table 7. Results of the Mann-Whitney U Test for the Control Class Using Reflective Interaction with Peers and the Experimental Class Using Reflective Interaction with ChatGPT

Test Statistics^a

Interaksi reflektif dengan teman sejawat dan interaksi reflektif dengan ChatGPT	
Mann-Whitney U	505.000
Wilcoxon W	1325.000
Z	-2.901
Asymp. Sig. (2-tailed)	.004

a. Grouping Variable: Kelas

Based on the Test Statistics table, the Mann-Whitney U value was 505.000, the Z value was -2.901, and the Asymp. Sig. (2-tailed) value was 0.004. Since the significance value was lower than 0.05, the difference between the control class and the experimental class was statistically significant. This means that there was a significant difference between the argumentative writing results of students who used reflective interaction with ChatGPT and those who used reflective interaction with peers.

To determine the magnitude of the difference between the two groups, the effect size was calculated using the formula $r = Z/\sqrt{N}$. Based on the Z value of -2.901 and the total sample size of 80 students, the effect size value was 0.324. This value falls into the medium category. Therefore, the difference in argumentative writing results between the control class and the experimental class was not only statistically significant but also had a medium level of practical effect.

Based on these results, it can be concluded that reflective interaction with ChatGPT was more effective than reflective interaction with peers in teaching argumentative text writing to students at SMP Negeri 8 Makassar. Therefore, the alternative hypothesis (H1) was accepted, while the null hypothesis (H0) was rejected.

4. Discussion

Learning Argumentative Text Writing Through Reflective Interaction with Peers among Students of SMP Negeri 8 Makassar

The results of the study in the control class showed that learning argumentative text writing through reflective interaction with peers achieved fairly good results. This can be seen from the mean score of the control class, which was 85.00 and was categorized as good. This score indicates that most students were able to write argumentative texts in accordance with the assigned task.

The mean score of 85.00 indicates that reflective interaction with peers can support the process of evaluating students' writing. This is supported by Kafa et al. (2025), who explained that reflective interaction through peer feedback can significantly improve students' writing skills. Students become more critical in assessing written texts, more careful in correcting errors, and more responsible for the quality of their writing. Through the activity of reading each other's work,

students do not only act as writers but also as readers who evaluate their peers' texts. This activity provides an important learning experience because students can compare the way they construct their own argumentative texts with the writing of their peers Widayati et al. (2022).

Dharma et al. (2024) explained that reflective interaction with peers in writing instruction occurs when students examine each other's writing, provide feedback based on guidelines or checklists, discuss errors, ask for clarification, and decide whether peer suggestions should be used in revising their writing. This process allows students not only to receive comments passively but also to think critically about the quality of writing, both their own writing and that of their peers.

Dharma et al. (2024) also stated that peer feedback helps junior high school students improve their writing skills in terms of content, organization, grammar, vocabulary, and mechanics. In addition, it creates a more comfortable learning atmosphere because students can discuss their writing more openly with their peers. This finding is supported by Sari et al. (2024), who showed that the implementation of peer correction techniques can make students more active, confident, focused, and enthusiastic, as well as increase their interest and motivation in learning to write exposition texts.

Learning Argumentative Text Writing Through Reflective Interaction with ChatGPT among Students of SMP Negeri 8 Makassar

The results of the study in the experimental class showed that learning argumentative text writing through reflective interaction with ChatGPT achieved very good results. The mean score of the experimental class was 92.88, indicating that students' writing outcomes were in the very good category. The lowest score obtained by students was 80, while the highest score reached 100.

The standard deviation of the experimental class was 6.493, and the variance was 42.163, indicating that the distribution of students' scores was relatively small. This means that students' writing outcomes in the experimental class did not show substantial differences from one student to another. This finding suggests that learning through reflective interaction with ChatGPT provided strong support for students in writing argumentative texts. This finding is in line with Mahapatra (2024), who explained that ChatGPT can provide fast and contextual formative feedback, thereby helping students understand errors, improve sentence structure, and increase their awareness of the writing process.

In reflective interaction with ChatGPT during the process of learning argumentative text writing at school, ChatGPT can function as a learning partner that provides fast, personalized, and focused feedback. Through this interaction, students can receive guidance on the weaknesses in their writing, such as unfocused ideas, unclear arguments, insufficient supporting evidence, inappropriate word choice, or ineffective sentences Balci, (2024).

Through reflective interaction, ChatGPT can help students review the parts of

their writing so that their texts become more focused and well-directed. This is in line with Suh et al. (2025), who showed that the use of ChatGPT in argumentative writing can help students improve the clarity of ideas, logical coherence, and the use of evidence in their writing. This finding is also reinforced by Alnemrat (2025), who explained that AI-based feedback can help improve the quality of argumentative writing, particularly in strengthening arguments and improving linguistic aspects of writing.

In AI-based learning, Robillos (2024) explains that ChatGPT serves as a learning partner that assists students in writing, starting from planning, drafting, revising, to editing. This is in line with (Li et al. 2023), who state that ChatGPT is not merely a text-based program, but also functions as an artificial cognitive agent capable of imitating the way humans think and write reflectively. Li et al. further emphasize that ChatGPT can produce texts that resemble human writing, with logical, cohesive, and well-reasoned argumentative structures.

Suharmawan (2023) states that ChatGPT is an artificial intelligence system used for text-based interaction. In learning, ChatGPT functions as a learning aid that supports language translation, material recommendations, productivity improvement, and the provision of interactive learning resources. ChatGPT provides easier access to information, helps students formulate ideas, and facilitates problem-solving, academic task completion, and independent evaluation of the tasks they have completed.

The Effectiveness of Reflective Interaction with ChatGPT in Learning Argumentative Text Writing among Students of SMP Negeri 8 Makassar

Based on the research findings, reflective interaction with ChatGPT was proven to be effective in learning argumentative text writing among students of SMP Negeri 8 Makassar. This effectiveness can be seen from the comparison of the results between the experimental class and the control class. The experimental class, which used reflective interaction with ChatGPT, obtained a mean score of 92.88, while the control class, which used reflective interaction with peers, obtained a mean score of 85.00. This difference in mean scores indicates that students' writing outcomes in the experimental class were higher than those in the control class.

Furthermore, the Test Statistics results showed a Mann-Whitney U value of 505.000 and an Asymp. Sig. (2-tailed) value of 0.004. Since the significance value of 0.004 was lower than 0.05, there was a significant difference between the argumentative text writing outcomes of students in the experimental class and those in the control class.

Based on the results of the Mann-Whitney U test in this study, the alternative hypothesis (H1) was accepted, while the null hypothesis (H0) was rejected. This indicates that there was a significant difference between argumentative text writing instruction among students who used reflective interaction with ChatGPT and those who used reflective interaction with peers. Since the experimental class

obtained a higher mean score and mean rank, reflective interaction with ChatGPT can be considered more effective in learning argumentative text writing among students of SMP Negeri 8 Makassar.

ChatGPT functions as a reflective learning tool that provides immediate feedback, helps students identify weaknesses in their writing, improve sentence structure, develop arguments, and organize ideas more coherently. In this process, students do not merely receive final results but are also encouraged to evaluate and revise their writing gradually. Mahapatra (2024) stated that ChatGPT can function as a formative feedback tool in writing instruction and has a positive effect on students' writing skills. ChatGPT-based reflective feedback can improve writing quality, particularly in the aspects of essay organization, grammar, vocabulary, argumentation, coherence, clarity, and student engagement in the writing process (Nurjati et al., 2025).

Pancasakti et al. (2025) stated that ChatGPT helps school and university students develop ideas, improve text structure, and enhance the quality of their writing. Yasmin et al. (2025) explained that ChatGPT is effective in improving argumentative writing skills because it can support error correction and writing revision. In addition, Khampusaen (2025) revealed that the use of ChatGPT can improve the quality of argumentative writing, particularly in the aspects of argument construction, academic language use, and content development.

This finding is in line with Dewey's reflective learning theory, as cited in Suphari and Chinokul (2023), which views reflection as an active, continuous, and careful thinking process that occurs when individuals encounter a problem and analyze their experience to gain new understanding. Students who interact with ChatGPT do not merely receive feedback on their writing, but are also guided to review the weaknesses of their texts, understand errors in argument construction, improve the structure of their ideas, and develop their writing in a more directed manner.

This finding is also relevant to Schön's concept, as cited in Marlina et al. (2023), regarding *reflection-in-action* and *reflection-on-action*. *Reflection-in-action* occurs when students revise their writing during the writing process based on the feedback they receive, whereas *reflection-on-action* occurs when students review their completed writing to determine further improvements. In this context, ChatGPT functions as a digital reflective partner that helps students evaluate, revise, and develop argumentative texts in a more directed way, as long as it is used as a thinking aid rather than a substitute for students' own writing ability.

In writing instruction, *self-regulated learning*, or independent learning, can be seen when students are able to plan the content of their writing, monitor the relevance of their arguments, evaluate the strength of their evidence, and revise their texts based on the feedback they receive. Panadero (2017) explains that *self-regulated learning* includes cognitive, metacognitive, behavioral, motivational, and emotional aspects of the learning process. In this study, students' ability to use ChatGPT reflectively can be understood as a form of independent learning because

students do not merely receive input from AI, but also interpret, select, and apply that input to improve their argumentative writing. Teng and Zhang (2020) also emphasize that self-regulated learning strategies in writing can help learners develop the ability to plan, monitor, and evaluate the writing process more consciously

From the perspective of formative assessment, ChatGPT feedback functions as input during the learning process, rather than merely as a final assessment. Shute (2008) explains that formative feedback is information provided to learners during the learning process to help them improve their thinking, learning strategies, or learning outcomes so that learning improvement can occur. In this study, reflective interaction with ChatGPT serves as a form of formative feedback because students receive direct input on the argumentative texts they compose. This input helps students identify errors, improve text structure, clarify arguments, review the strength of evidence, and improve linguistic aspects in their writing.

In writing instruction, *self-regulated learning* is reflected in students' ability to plan the content of their writing, monitor the relevance of their arguments, evaluate the strength of their evidence, and revise their texts based on the feedback they receive. Panadero (2017) explains that *self-regulated learning* encompasses cognitive, metacognitive, behavioral, motivational, and emotional aspects of the learning process. In this study, students' ability to use ChatGPT reflectively can be understood as a form of self-regulated learning because they do not merely receive input from AI, but also interpret, select, and apply that input to improve their argumentative writing. Teng and Zhang (2020) further emphasize that self-regulated learning strategies in writing can help learners develop the ability to plan, monitor, and evaluate the writing process more consciously.

AI-assisted writing instruction can help improve the quality of students' writing. Ding, et al (2025) state that ChatGPT can help improve the mechanical aspects of writing, grammar, tone, and overall writing quality. This supports the findings of the present study, in which students who learned through reflective interaction with ChatGPT achieved higher writing outcomes than those who learned through reflective interaction with peers, with the experimental class obtaining a mean score of 92.88 and the control class obtaining a mean score of 85.00.

However, the use of ChatGPT in writing instruction still requires pedagogical guidance. Bašić et al. (2023) found that the use of ChatGPT does not always produce better writing than that of the control group, while Yoon et al. (2023) showed that ChatGPT feedback can be general when it is not guided by clear instructions. Therefore, the effectiveness of ChatGPT in this study was not solely caused by the use of AI technology, but also by its implementation through reflective interaction, which helped students review errors, improve arguments, and revise argumentative texts in a more directed manner.

Pedagogically, ChatGPT feedback is more effective than peer feedback

because ChatGPT can provide fast, specific, personalized, and relatively consistent input on students' writing. In learning to write argumentative texts, students need feedback that helps them identify weaknesses in content, clarify arguments, review the strength of evidence, improve text structure, and refine linguistic aspects. Peer feedback remains useful, but its quality often depends on students' ability to understand assessment criteria and their confidence in giving criticism. In contrast, ChatGPT can serve as a digital reflective partner that helps students evaluate and revise their writing in a more directed manner. This is supported by the findings of the study, which showed that the experimental class obtained a mean score of 92.88, higher than the control class, which obtained a mean score of 85.00, with a significance value of $0.004 < 0.05$.

This study has several limitations. First, the sample size was relatively limited because it involved only two classes, namely class IX.12 as the experimental class and class IX.10 as the control class. Second, the intervention period was relatively short, so the findings do not fully reflect the long-term impact of using ChatGPT in writing instruction. Third, the study was conducted in only one school, namely SMP Negeri 8 Makassar. Therefore, the findings should be understood within the context of that school and cannot yet be broadly generalized to other schools or educational levels.

Future researchers are recommended to involve larger samples, longer intervention periods, and more diverse school contexts. Although reflective interaction with ChatGPT has been proven effective, its use still has limitations. Possible risks include students' dependence on ChatGPT, the possibility of inaccurate feedback, and the decline of writing originality if students merely copy AI-generated responses. Therefore, ChatGPT should be used as a reflective partner, not as a substitute for students' writing ability. Students should still write their own drafts, critically examine the feedback they receive, and revise their writing based on their own understanding.

5. Conclusion

Based on the research findings, reflective interaction with ChatGPT was found to be more effective in learning argumentative text writing among students of SMP Negeri 8 Makassar than reflective interaction with peers. This is indicated by the mean score of the experimental class, which was 92.88 and higher than that of the control class, which was 85.00. In addition, the results of the Mann-Whitney U test showed a significance value of $0.004 < 0.05$, indicating that H1 was accepted and H0 was rejected. ChatGPT functions as a reflective learning tool that is able to provide immediate feedback, help students improve the structure of their writing, develop arguments, enhance coherence, grammar, and vocabulary, and organize ideas more systematically. These findings are in line with several previous studies showing that ChatGPT can support the writing process through formative feedback, error correction, idea development, and improvement in the quality of students' argumentative writing.

Future researchers are recommended to examine the use of ChatGPT in writing instruction across different text types, such as narrative, exposition, descriptive, procedural, or observation report texts. This is important to determine whether reflective interaction with ChatGPT is also effective when applied to other text types besides argumentative texts. In addition, future studies may be conducted at different educational levels, such as elementary school, senior high school, vocational high school, or higher education, so that the findings can provide a broader understanding of the use of ChatGPT in writing instruction. Future researchers are also encouraged to investigate the long-term impact of ChatGPT-assisted writing instruction, particularly on the development of writing skills, critical thinking, learner autonomy, and students' ethical awareness in using artificial intelligence technology.

6. References

- Alexander dkk, (2025). *Exploring generative artificial intelligence to enhance reflective writing in pharmacy education. American Journal of Pharmaceutical Education*, 89(6), 101416.
- Alnemrat, A., Aldamen, H., Almashour, M., Al-Deaibes, M., & AlSharefeen, R. (2025). AI vs. teacher feedback on EFL argumentative writing: A quantitative study. *Frontiers in Education*, 10, Article 1614673. doi: 10.3389/educ.2025.1614673
- Bacı, Ö. (2024). The role of ChatGPT in English as a foreign language (EFL) learning and teaching: A systematic review. *International Journal of Current Educational Studies*, 3(1), 66–82. DOI: 10.46328/ijces.107
- Bašić, Ž., Banovac, A., Kružić, I., & Jerković, I. (2023). ChatGPT-3.5 as writing assistance in students' essays. *Humanities and Social Sciences Communications*, 10, 750. doi: 10.1057/s41599-023-
- Candara, A. (2024). *Peningkatan Kemampuan Menulis Teks Argumentasi dalam Pembelajaran Bahasa Indonesia di Sekolah Menengah Pertama*. *Jurnal Pendidikan Bahasa dan Sastra*, 12(1), 45-58.
- Dharma, M. K., Suryati, N., Laksmi, E. D., & El Khoiri, N. (2024). The effect of peer feedback to improve students' writing ability for lower secondary school students. *Jurnal Pendidikan Humaniora*, 12(2), 54–63.
- Ding, L., Zou, D., & Kohnke, L. (2025). ChatGPT as an automated writing evaluation tool: How students perceive it and how it affects their writing. *Education and Information Technologies*, 30, 26777–26799. doi: 10.1007/s10639-025-13775-3
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. doi: 10.3102/003465430298487
- Hidayati, N. (2023). *Tingkat kesulitan siswa dalam menulis teks argumentasi*. *Jurnal Pembelajaran Bahasa*, 9(4), 78-89.
- Indah, T., Wahyuni, S., Gunawan, D., & Mulyana, H. (2021). *Analisis Kesulitan Siswa*

Dalam Menulis Puisi Naratif Pada Mata Pelajaran Bahasa Indonesia Di Kelas Iv (Studi Deskriptif Kualitatif Di Kelas IV SDN 2 Sukagalih) (Vol. 01, Nomor 02).

- Kafa, N., Ropikoh, S., & Hidayah, N. (2025). Strategi meningkatkan keterampilan menulis siswa melalui metode peer feedback. *EduSpirit: Jurnal Pendidikan Kolaboratif*, 2(1), 241–246.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). *ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences*, Khampusaen, D. (2024). *The impact of ChatGPT on academic writing skills and knowledge: An investigation of its use in argumentative essays. LEARN Journal: Language Education and Acquisition Research Network*, 17(2), 1–25.
- Li, Y., Sha, L., Yan, L., Lin, J., Raković, M., Galbraith, K., Lyons, K., Gašević, D., & Chen, G. (2023). Can large language models write reflectively. *Computers and Education: Artificial Intelligence*, 4(May).
- Mahapatra, S. (2024). *Impact of ChatGPT on ESL students' academic writing skills: a mixed methods intervention study. Smart Learning Environments*, 11(1).
- Marlina, R., Suwono, H., Yuenyong, C., Ibrohim, Mahanal, S., & Hamdani. (2023). *Introducing Reflective Teaching Practice in Remote School: Teacher-Student Interaction Patterns. Jurnal Penelitian Pendidikan IPA*, 9(11), 9720–9728.
- Masruddin, Hartina, S., Arifin, M. A., & Langaji, A. (2024). Flipped learning: facilitating student engagement through repeated instruction and direct feedback. *Cogent Education*, 11(1), 2412500.
- Masruddin, M., Amir, F., Langaji, A., & Rusdiansyah, R. (2023). Conceptualizing linguistic politeness in light of age. *International Journal of Society, Culture & Language*, 11(3), 41-55.
- Mollick, E., & Mollick, L. (2023). *Assigning AI: Seven approaches for students, with prompts. arXiv. arXiv:2306.10052*
- Nurjati, N., Syahria, N., & Khan, A. K. B. S. (2025). ChatGPT-based reflective feedback to improve EFL graduate students' academic writing for publication. *JOLLT Journal of Languages and Language Teaching*, 13(4), 1802–1816.
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 422. doi: 10.3389/fpsyg.2017.00422
- Pancasakti, U., Latif, R., Pancasakti, U., Delresin, B., Pancasakti, U., Pancasakti, U., Welmina, I., Pancasakti, U., & ChatGPT, P. (2025). *Efektivitas Penggunaan ChatGPT Dalam Pembelajaran Menulis Wacana The Effectiveness Of Using ChatGPT In Discourse Writing*. 5(3), 295–301.
- Rahmawati, L., Yuliawati, R., Anisah, A., & Hermansyah, H. (2025). *Analisis Pengaruh Pemanfaatan ChatGPT Terhadap Keterampilan Menyusun Pertanyaan Efektif Pada Mahasiswa Pendidikan Ekonomi. JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(5), 7858–7864.
- Ramadhan, M., Pratama, D., & Sari, L. (2024). *Pemanfaatan ChatGPT sebagai Asisten*

- Pembelajaran dalam Meningkatkan Pemahaman Siswa. Jurnal Teknologi Pendidikan*, 13(3), 145-160.
- Robillos, R. J. (2024). *Synergizing Generative Pre-Trained Transformer (GPT) Chatbots in a Process-Based Writing Paradigm to Enhance University Students' Writing Skill. Journal of Language and Education*, 10(3), 79–94.
- Sari, A., & Rahman, B. (2024). *Pemanfaatan teknologi berbasis kecerdasan buatan dalam menciptakan lingkungan pembelajaran inklusif. Jurnal Teknologi Pendidikan*, 9(1), 45–54.
- Sari, E. P., Trianto, A., & Utomo, P. (2020). *Kesulitan Menulis Teks Eksplanasi Siswa Kelas Viismpnegeri 14 Kota Bengkulu. Jurnal Ilmiah KORPUS*, 4(3), 292–302.
- Septiwi, E., Kuntarto, E., & FKIP Universitas Jambi, P. (n.d.). *Aspek-Aspek Penyebab Kesulitan Menulis Puisi Pada Siswa Kelas V Sekolah Dasar*.
- Sok, S., & Heng, K. (2023). ChatGPT for education and research: A review of benefits and risks. *Cambodian Journal of Educational Research*, 3(1), 110–121
- Suharmawan, W. (2023). Pemanfaatan ChatGPT dalam dunia pendidikan. *Education Journal: Journal Education Research and Development*, 7(2), 158–166. <https://doi.org/10.31537/ej.v7i2.1248>
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. doi: 10.3102/0034654307313795
- Suphari, P., & Chinokul, A. (2023). *Refleksi dalam Proses Pembelajaran: Pengembangan Pemahaman Baru dan Pe nyesuaian Diri. Jurnal Pendidikan dan Pembelajaran*, 22(3), 50-65.
- Teng, L. S., & Zhang, L. J. (2020). Empowering learners in the second/foreign language classroom: Can self-regulated learning strategies-based writing instruction make a difference? *Journal of Second Language Writing*, 48, 100701.
- Utami, A. D., Fadlilah, M., Aini, N. Q., & Syafutri, R. A. (2024). *Struktur dan Ciri-Ciri Teks Argumentasi dalam Bahasa Indonesia: Analisis dan Contoh. Semantik : Jurnal Riset Ilmu Pendidikan, Bahasa dan Budaya*, 2(3), 234–243.
- Werdiningsih, I., Marzuki, Indrawati, I., Rusdin, D., Ivone, F. M., Basthomi, Y., & Zulfahreza. (2024). *Revolutionizing EFL writing: unveiling the strategic use of ChatGPT by Indonesian master's students. Cogent Education*, 11(1).
- Widayati, D., & Ludyanti, L. N. (2022). Metode Maternal Reflektif Dalam Meningkatkan Kemampuan Interaksi Sosial Pada Anak. *Prosiding SPIKesNas: Seminar ...*, 01(02), 354–359.
- Yasmin, M., Fatima, W., & Irshad, I. (2025). *Evaluating ChatGPT ' s effectiveness in enhancing argumentative writing: A quasi-experimental study of EFL learners in Pakistan. Sustainable Futures*, 10(June), 100809.
- Yoon, S.-Y., Miszoglad, E., & Pierce, L. R. (2023). *Evaluation of ChatGPT feedback on ELL writers' coherence and cohesion*. arXiv:2310.06505
- Yulianti, Y., Umar, U., Sudewi, P. W., & Isma, A. (2024). Analysis of Narrative Writing Skills Among Junior High School Students in Indonesia. *Innovations in*

Moh Taufiq Hamsah¹, Syamsudduha², Muhammad Saleh³
The Effectiveness of Reflective Interaction with ChatGPT in Teaching Argumentative Text
Writing at Junior High School
Language Education and Literature, 1(2), 41–47.