The Effectiveness of Ai-Based Race Framework to Enhance EFL Learners Narrative Writing Skills

Muchamad Lorna Maulia | Universitas Islam Majapahit, Mojokerto, Jawa Timur, Indonesia
Ahmad Iklil Saifulloh | Universitas Islam Majapahit, Mojokerto, Jawa Timur, Indonesia
Wiwik Mardiana | Universitas Islam Majapahit, Mojokerto, Jawa Timur, Indonesia
Email: ornareal44@gmail.com

Abstract

Abstract: This study investigates the effectiveness of the AI-based RACE (Role, Action, Context, Execute) Framework in enhancing EFL learners' narrative writing skills. Conducted at SMAN 1 Mojosari, the research employed a quasi-experimental design involving two classes: an experimental group taught using ChatGPT with the RACE Framework and a control group taught using conventional methods. A total of 72 eleventh-grade students participated. Data collection included pre-tests, post-tests, and questionnaires to assess writing performance and motivation. Writing assessment focused on content, organization, vocabulary, language use, and mechanics. The results showed significant improvement in the experimental group compared to the control group. Statistical analysis using SPSS confirmed the effectiveness of the AI-based RACE Framework in improving coherent and well-structured narrative writing. Questionnaire results indicated positive student responses, showing higher motivation and engagement. This study suggests that integrating AI with the RACE Framework creates a structured, supportive environment for enhancing narrative writing in EFL classrooms.
Keywords: RACE Framework, Artificial Intelligence, ChatGPT, Narrative Writing, EFL Learners.

Introduction

Writing skills in Indonesia remain a significant problem, both at the elementary and tertiary levels. Utari and Rambe(Utari & Rambe, 2023) stated that students' low writing ability is due to a lack of practice and a lack of supportive learning environments. This aligns with Tarigan's(Rahmi, 2021) view, which emphasizes that writing is a skill that needs to be developed through continuous practice so that ideas can be expressed effectively. Unfortunately, opportunities for intensive writing practice remain limited, resulting in students' low ability to organize ideas systematically.

At the tertiary level, students face similar challenges in academic writing. Heriyudanta(Heriyudananta, 2021) believes that low academic writing competence is influenced by a lack of effective learning and the absence of early writing habits. Byrne(Dang, Chau, & Tra, 2020) adds that writing is a complex skill that requires an understanding of structure, organization of ideas, and the selection of appropriate diction to convey a message clearly. Therefore, educational institutions need to develop effective writing learning programs and foster a writing culture from an early age.

The author's experience during his two-month teaching internship at SMAN 1 Mojosari (September–October) showed that some eleventh-grade students experienced difficulties developing their writing skills, particularly in narrative text. These difficulties included organizing ideas, constructing a coherent flow, and using correct grammar. This situation aligns with

Harmer's(Salsabila, 2020) view that writing is not simply about putting words together, but also involves critical thinking skills, creativity, and an understanding of language structure.

Poor writing skills are also influenced by a lack of structured practice and a lack of motivation. Byrne emphasized that writing barriers are often caused by limited exposure to quality text examples and limited opportunities for revision and feedback from teachers. Therefore, interactive learning strategies are needed that can train writing skills while deepening students' understanding of narrative text structure.

The use of Artificial Intelligence (AI) is one innovative solution to address this problem. AI-based writing assistants like ChatGPT have the potential to help students structure texts better, improve grammar, and provide instant feedback(Griffiths & Forcier, 2016). AI can provide personalized and adaptive learning tailored to students' needs, thereby increasing their writing confidence. AI can also be utilized to develop story ideas, create writing outlines, and improve sentence coherence. Pinkwart(Pinkwart, 2016) demonstrated that the use of AI can enhance students' critical thinking skills and creativity, particularly in writing.

One approach that can be integrated is the RACE Framework (Role, Action, Context, Execute), which has proven effective in systematically designing AI prompts. Research by Woo et al.(Woo, Wang, Yung, & Guo, 2024) suggests that effective prompt engineering can enhance students' understanding of AI, build confidence, and hone critical thinking skills. Implementing RACE allows teachers to craft clear and contextual prompts, for example, by designating the AI as a "professional story editor" (Role), assigning it the task of "improving the structure and flow of the story" (Action), providing information on themes and characters (Context), and instructing feedback on improvements (Execute).

The RACE Framework aligns with the concepts described in the Trust Insights RACE AI Framework, which emphasizes the importance of structured prompts for AI to produce relevant output. Knoth et al.(Sawalha, Taj, & Shoufan, 2024) added that a deep understanding of prompt engineering increases the effectiveness of using AI in systematic learning. The integration of RACE into AI-based learning not only improves writing skills but also hones students' digital literacy.

Several studies have been conducted on the use of AI in education, particularly in improving the writing skills of EFL (English as a Foreign Language) students. However, most studies still focus on descriptive writing and lack a clear learning framework, resulting in a lack of structure. This research is urgent because it applies the RACE Framework as a systematic approach to improving EFL students' narrative writing skills. This skill is more complex than descriptive writing because it involves story structure, narrative coherence, and character development.

The analysis of over 500 articles (with 304 relevant articles from the Scopus database) shows a significant trend AI research in education is rapidly advancing. While countries like China, the United States, Russia, and the United Kingdom are leading the way, Indonesia lags behind with only six articles related to AI in education.(Maphosa & Maphosa, 2023) Key barriers include limited infrastructure, policies, and expertise. Yet, AI has the potential to reduce teacher burden, personalize learning, increase engagement, and detect student learning difficulties early.

Therefore, this study aims to explore the use of AI based on the RACE Framework in English language learning at the high school level, specifically in teaching narrative text. This approach is expected to effectively improve students' narrative writing skills while equipping them with relevant technological skills in the digital age.

Literature Review

Writing

Writing is the activity of conveying ideas or thoughts, knowledge, and information(White, 1986). The goal is to share the meaning of the article with others. Conversely, writing is the embodiment of linguistic symbols(Tarigan, 1989). This statement is supported by another definition, which states that writing is the same as producing graphic symbols, called linguistic symbols, to represent the language itself(Windarti, Bintoro, & Fahrurrozi, 2018). Someone who understands graphic symbols can ultimately read and understand text. Writing has five component (Heaton, 1988):

1. Content

The content presented must be clear so that the reader can receive the general information in the texts. The content must have one purpose that is always connected to each paragraph. The content section is written easily so it can be expanded into several paragraphs. The content section is also written simply so that it is easy to understand and does not confuse the reader.

2. Organization

Writing organization is concerned with writing, and organizing ideas or messages conveyed by the author to the reader. This organizing section also arranges the order or chronology of the writing from beginning to end.

3. Vocabulary

Vocabulary provides many choices of words for writers to compose their writing. Vocabulary is essential because it affects the writing result. Choosing kind and polite words is the right choice to make an interesting writing. Someone who does not master the vocabulary will have difficulty arranging words, so the writing process will be disrupted.

4. Language Use

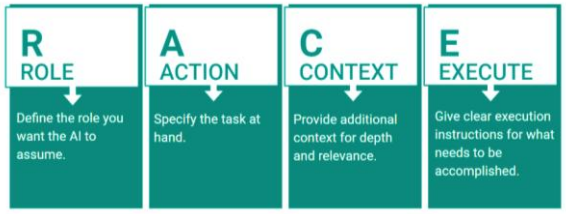
The use of language governs how the language is used and good grammar for writing. In writing, students must be able to use a neat grammatical system. The use of language also regulates the use of formal language. Formal language is considered more polite than informal language.

5. Mechanics

Mechanics discusses two things, namely function and capitalization. The function here regulates the clarity of meaning in writing. Capitalization handles capital letters at the beginning of sentences and after direct quotations. Capital letters are also used to write someone's name, title, place name, month name, etc. Although it looks trivial, the function of capital letters is essential for the clarity of writing.

Race

In the ever-evolving landscape of marketing, generative AI stands at the forefront of the latest technological revolution. Tools like ChatGPT, Claude, and Google's Bard are not just buzzwords; they're reshaping how we approach marketing strategy and content creation. As a marketer in this



AI-driven era, staying ahead means not only understanding these tools but also mastering the art of leveraging them effectively. This is where the concepts of prompt engineering and the RACE framework become crucial(Woo et al., 2024).

1. Role

   The 'Role' aspect of the framework involves assigning a specific identity or function to the AI. This step shapes how the AI frames its responses, aligning with the nature of your content needs. For instance, when creating content for a technical audience, assigning the AI the role of a subject matter expert can lead to more in-depth and appropriate responses. It's about matching the AI's 'character' to the context of your content, ensuring that it resonates well with your intended audience.

2. Action

   'Action' refers to clearly defining the task you want the AI to perform. This could be anything from writing a comprehensive report to generating creative ad copy. The specificity of your action directive guides the AI in focusing its capabilities, leading to outputs that are more aligned with your specific goals. Clear, concise instructions in this step are crucial for the AI to understand and execute your content objectives effectively.

3. Context

   In the 'Context' phase, you enrich your prompt with additional information that guides the AI's response. This might include details about your target audience, the tone of your brand, and specific objectives of your content piece. Providing this context helps the AI tailor its responses, making them more relevant and impactful for your particular marketing situation. It's akin to setting the scene in which the AI operates, ensuring that its contributions are well-suited to your strategic needs.

4. Execute

   Finally, 'Execute' is your command to the AI to begin creating the content based on the preceding elements. This clear instruction is the catalyst that transforms your carefully constructed setup into concrete results. Effective execution commands ensure that the AI's computational abilities are directed towards fulfilling your content requirements in a focused and efficient manner.

Methodology

This study employed a quantitative approach using a quasi-experimental design. The research was conducted at SMAN 1 Mojosari, located in Seduri Village, Mojosari District, Mojokerto Regency. Two eleventh-grade classes were selected as participants: XI-7 as the experimental group and XI-8 as the control group, with a total of 72 students (24 males and 48 females). The experimental group was taught narrative text writing using folktales integrated with the AI-Based RACE Framework through ChatGPT, while the control group received instruction through conventional methods using PowerPoint.

The research instruments included pre-tests, post-tests, and a questionnaire. The pre-test was given before treatment to measure students' initial narrative writing ability, while the post-test was administered after treatment to evaluate improvement. Writing performance was assessed based on five components: content, organization, vocabulary, language use, and mechanics. The questionnaire, using a Likert scale, measured students' motivation and perceptions toward AI-based learning, but was only administered to the experimental group.

The data collection procedure consisted of four stages:

1. **Pre-test** – both groups completed a narrative writing task to establish baseline ability.

2. **Treatment** – the experimental class learned through the AI-Based RACE Framework, where students engaged with ChatGPT prompts structured around Role, Action, Context, and Execute. The control class learned through teacher explanation, sample text analysis, and writing practice without AI assistance.
3. **Post-test** – both groups wrote a new narrative text to measure skill improvement.
4. **Questionnaire** – the experimental group responded to items assessing engagement, motivation, and perception of AI integration.

For data analysis, descriptive statistics (mean and standard deviation) were used to summarize results. The gained score between pre-test and post-test was calculated to determine improvement. An independent sample t-test, conducted using SPSS, was used to identify significant differences between the experimental and control groups. Validity and reliability tests were performed on the instruments to ensure data accuracy and consistency.

This methodology was designed to evaluate not only the effectiveness of the AI-Based RACE Framework in enhancing students' narrative writing skills but also its impact on learner motivation in the EFL classroom context.

Results and Discussion

After administering treatment to both groups, this study analyzed the results of the instrument used, namely tests. The analysis began with normality and mann-whitney tests, followed by paired-sample t-tests to evaluate improvement in each group, and independent-sample t-tests to determine significant differences between the experimental and control groups.

Normality Test

In the normality test section, there are 4 tests to see the distribution data of Pre-Test and Post-Test in the experimental class and control class. The tests included: the experimental class Pre-Test normality test, the control class Pre-Test normality test, the experimental class Post-Test normality test, and the control class Post-Test normality test. The following table shows the normality test for the Pre-Test and Post-Test for the experimental and control classes.

Table 4.8 Normality Test for Experimental and Control Class

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Kelas	Statistic	df	Sig.	Statistic	df	Sig.
Hasil Belajar	Pre-Test Eksperimen(AI)	.140	36	.072	.924	36	.017
	Post-Test Eksperimen(AI)	.126	36	.162	.941	36	.054
	Pre-Test Kontrol(Konvensional)	.078	36	.200*	.966	36	.329
	Post-Test Kontrol(Konvensional)	.226	36	.000	.896	36	.003

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 4.8 shows that the results of the normality test using the Shapiro-Wilk test are known to have a Sig value for each data: 1) Experimental Class Pre-test of 0.017) Experimental Class Post-test of 0.054) Control Class Pre-test of .329) Control Class Post-test of .003. In data analysis, a significance level of 5% or 0.05 was set. Based on the decision that the significance value (p) is greater than 0.05

(p> 0.05). Thus, the distribution of Pre-Test and Post-Test data in the experimental class and control class using the Shapiro-Wilk normality test, the data is declared to be abnormally distributed.

Mann-Whitney

The Mann-Whitney test in this study was used to determine whether there is a difference in the learning outcomes of students who use Chat-GPT compared to those who do not use Chat-GPT. If the data is not normal, the following decision-making process is based on:
a) If Asymp.Sig. (2-tailed) < 0.05, there is a significant difference.
b) If Asymp.Sig. (2-tailed) > 0.05, there is no significant difference.

Table 4.9 Result of Main-Whitney Test

Test Statistics[a]	
	Hasil Belajar
Mann-Whitney U	.000
Wilcoxon W	666.000
Z	-7.315
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Kelas

Accept the hypothesis if Asymp.Sig. < 0.05. Based on the data above, a Sig. of 0.000 is obtained, thus the hypothesis is accepted. It can be concluded that there is a difference in the learning outcomes of students who use Chat-GPT compared to those who do not use Chat-GPT.

Paired Sample T-Test

In this study, the first data was obtained from the control class that was given learning using conventional methods, while the second data came from the experimental class that was given treatment through the use of Chat-GPT as a learning media. The Paired Sample T-Test results for the control class, which were analyzed using IBM SPSS v.25, are presented in the following tables:

Control Class

Table 4.10 Result Paired Sample T-Test Statistic Control Class

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	pretest	58.39	36	4.512	.752
	posttest	71.08	36	5.427	.904

The Paired Sample Statistics table presents the descriptive values of each variable in the sample pair. In the pre-test, the mean value was 58.39 out of 36 data, with a Standard Deviation of 4.512 and a Standard Error Mean of .752. Meanwhile, in the post-test, the average value increased to 71.08, with a standard deviation of 5.427 and a standard error mean of .904.

Based on these descriptive results, it can be concluded that the post-test value is higher than the pre-test value, which shows an increase in learning outcomes in the control class treatment using with conventional method.

Table 4.11 Result of Paired Sample T-Test Control Class

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	pretest - posttest	-12.694	5.371	.895	-14.512	-10.877	-14.181	35	.000

The analysis results show that the Significance value (Sig. 2-tailed) is 0.000, which is smaller than the significance level a = 0.05. Thus, it can be concluded that the alternative hypothesis (Ha) is accepted and the null hypothesis (Ho) is rejected. This indicates that there is a significant improvement between the pre-test and post-test results in the control class.

Experimental Class

Table 4.12 Result of Paired Sample T-Test Statistic Experimental Class

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	pretest	59.92	36	6.026	1.004
	posttetst	88.11	36	2.926	.488

The Paired Sample Statistics table presents the descriptive values of each variable in the sample pair. In the pre-test, the mean value was 59.92 out of 36 data, with a Standard Deviation of 6.026 and a Standard Error Mean of 1.004. Meanwhile, in the post-test, the average value increased to 88.11, with a standard deviation of 2.926 and a standard error mean of .488.

Based on these descriptive results, it can be concluded that the post-test value is higher than the pre-test value, which shows an increase in learning outcomes in the Experiemantal class treatment using with Chat-GPT as a learning media.

Table 4.13 Result of Paired Sample T-Test Experimental Class

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
pretest - posttetst	-28.194	5.879	.980	-30.184	-26.205	-28.775	35	.000

Activate W
Go to Settings

The analysis results show that the Significance value (Sig. 2-tailed) is 0.000, which is smaller than the significance level a = 0.05. Thus, it can be concluded that the alternative hypothesis (Ha) is accepted and the null hypothesis (Ho) is rejected. This indicates that there is a significant improvement between the pre-test and post-test results in the experimental class.

Independent Sample T-Test

This section is used to see if there is a significant difference in student's ability to write narrative texts between the experimental and control classes. Using SPSS version 29, this t-test was computed based on the Post-Test scores for the experimental and control classes.

The basis for decision-making in this independent sample t-test analysis is determined by the significant value of the significance level. Then 2 possibilities can be set, namely: 1) If the value of Sig. (2-tailed) < 0,05, then there is a significant effect of teaching narrative text using folktale through Chat-GPT media to enhance students' ability in writing; 2) If the value of Sig. (2-tailed) > 0.05, then there is no significant effect of teaching narrative text using folktales through Chat-GPT media to enhance students' ability in writing. The following table shows the results of the t-test scores for the experimental and control classes.

Table 4.14 Independent Test Group Statistic

Group Statistics

	kelas	N	Mean	Std. Deviation	Std. Error Mean
hasilbelajar	posttest_con	36	71.08	5.427	.904
	posttest_exper	36	88.11	2.926	.488

Based on the "Group Statistics" output table above, it is known that the number of Experimental Class is 36 students, while for Control Class is 36 students. The average value of student learning outcomes for the Experimental Class is 88.11 with a standard deviation of 2.926 and a mean standard error of .488. Meanwhile, for the Control Class, it is 71.08 with a standard deviation value of 5.427 and a mean standard error value of .904. Thus, statistically descriptively, it is possible to conclude that the results of the post-test scores differ between the experimental and control groups. Furthermore, to determine whether or not there is a significant difference, the researcher must interpret the results of the following Independent Sample T-test.

Table 4.15 Independent Sample T-test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
hasilbelajar	Equal variances assumed	3.818	.055	-16.572	70	.000	17.028	1.028	14.978	19.077
	Equal variances not assumed			-16.572	53.759	.000	17.028	1.028	14,968	19.088

Based on the Independent Sample T-test output table in the Equal variances assumed section, it is known that the value of Sig. (2-tailed) of 0,000 < 0,05 then as the basis for decision making in the Independent Sample T-test, it can be concluded that Ho is rejected and Hα received. Thus it can be concluded that there is a significant effect of teaching narrative text using folktales through Chat-GPT media to strengthen students' ability in writing.

Furthermore, from the output table above, it is known that the "Mean Difference" value is 17,028. This value indicates the difference between the average learning outcomes of the Experimental Class and the average learning outcomes of the Control Class or 88,11 – 71,08 =

17,028 and the gained score between those data is 14,978 to 19,077 (95% Confidence Interval of the Difference Lower Upper).

The following decision basis can be used to guide decision-making in the Independent Sample T-test based on a comparison of the t-count value with the t- t-table.

1. If the t-count > t-table then Ho is rejected and Hα is accepted, it means that there is a significant effect of teaching narrative text using folktale through Chat-GPT media to strengthen students' ability in writing.
2. If the t-count < t-table then Ho is accepted and Hα is rejected, it means that there is no significant effect of teaching narrative text using folktale through Chat-GPT media to strengthen students' ability in writing. It is known that the t-count is 16,572. To compare it, use the t-table value with reference to the formula (α/2). If the entered into the formula it will be (0,05/2 = 0,025). Then the level value of sig. used is 0.025. The value (pdf) is 63 as showed in table 4.8. In the distribution of statistical values, the t-table value obtained based on the calculation of the above formula is 1,999. Thus, the t-count value is 16,572 > t-table value is 1,999, so based on the basis of taking the t-count value with the t-table, it can be concluded that Ho is rejected and Hα is accepted. Based on this decision, the final result of the Independent Sample T-test is that there is a significant effect of teaching narrative text using folktale through Chat-GPT media to strengthen students' ability in writing.

Data were collected from Grades XI-7 and XI-8 at SMAN 1 MOJOSARI. The data collection process involved administering an Initial and Final Exam to students. Initially, students took the Initial Exam. Following the Initial Exam, a series of Treatment sessions were conducted over several meetings. During these sessions, Grade XI-7, which served as the experimental class, utilized Chat-GPT media to study narrative texts in the form of fairy tales. In contrast, Grade XI-8, which served as the control class, used traditional learning methods to study the same material. After the treatment sessions, students took a Final Exam as the final step in data collection. Additionally, for the experimental class, an additional final exam in the form of a questionnaire was administered to assess the media used.

The next step was data analysis based on assessments using five components and SPSS. These five components are Content, Organization, Vocabulary, Language Use, and Mechanics. The Content component encompasses the interpretation of relevant topics, subjects, or character names in a precise, detailed, and accurate manner. The Organization component encompasses the logical, clear, and organized sequence of the story according to the generic structure of narrative text. The Vocabulary component includes effective word/idiom selection, appropriate storyline, and appropriate presentation. The Language Usage component covers prepositions, pronouns, tenses, conjunctions, numbers, and articles. Finally, the Mechanics component covers spelling, punctuation, capitalization, proper paragraph organization, and layout.

The Final Exam results are shown based on the data analysis above. Students taught using Chat-GPT achieved better or superior scores compared to students taught using conventional materials. This is evident in Tables 4.4 and 4.5, where the experimental class had an average score of 59.91 before receiving the Treatment (Pre-Exam) and an average score of 88.11 after receiving the Treatment (Final Exam). Meanwhile, the control class had an average score of 58.38 before receiving the Treatment (Pre-Exam) and an average score of 71.08 after receiving the Treatment (Final Exam). There was a difference in scores on the Final Exam, with the experimental class performing better than the control class. This shows that Chat-GPT media can be used to strengthen students' abilities in writing narrative texts in the form of fairy tales.

Data were collected from Grades XI-7 and XI-8 at SMAN 1 MOJOSARI. The data collection process involved administering an Initial and Final Exam to students. Initially, students

took the Initial Exam. Following the Initial Exam, a series of Treatment sessions were conducted over several meetings. During these sessions, Grade XI-7, which served as the experimental class, utilized Chat-GPT media to study narrative texts in the form of fairy tales. In contrast, Grade XI-8, which served as the control class, used traditional learning methods to study the same material. After the treatment sessions, students took a Final Exam as the final step in data collection. Additionally, for the experimental class, an additional final exam in the form of a questionnaire was administered to assess the media used.

The next step was data analysis based on assessments using five components and SPSS. These five components are Content, Organization, Vocabulary, Language Use, and Mechanics. The Content component encompasses the interpretation of relevant topics, subjects, or character names in a precise, detailed, and accurate manner. The Organization component encompasses the logical, clear, and organized sequence of the story according to the generic structure of narrative text. The Vocabulary component includes effective word/idiom selection, appropriate storyline, and appropriate presentation. The Language Usage component covers prepositions, pronouns, tenses, conjunctions, numbers, and articles. Finally, the Mechanics component covers spelling, punctuation, capitalization, proper paragraph organization, and layout.

The Final Exam results are shown based on the data analysis above. Students taught using Chat-GPT achieved better or superior scores compared to students taught using conventional materials. This is evident in Tables 4.4 and 4.5, where the experimental class had an average score of 59.91 before receiving the Treatment (Pre-Exam) and an average score of 88.11 after receiving the Treatment (Final Exam). Meanwhile, the control class had an average score of 58.38 before receiving the Treatment (Pre-Exam) and an average score of 71.08 after receiving the Treatment (Final Exam). There was a difference in scores on the Final Exam, with the experimental class performing better than the control class. This shows that Chat-GPT media can be used to strengthen students' abilities in writing narrative texts in the form of fairy tales.

Conclusion

Researcher conducted a study at SMAN 1 Mojosari, specifically in Class XI-7 as the experimental class and Class XI-8 as the control class. The research aimed to teach narrative texts using folktale through Chat-GPT. The results demonstrated a higher increase in post-test scores for the experimental class compared to the control class. As discussed in Chapter 4, the experimental class had an average pre-test score of 59.91 and an average post-test score of 88.11 after the treatment. Meanwhile, the control class had an average pre-test score of 58.38 and an average post-test score of 71.08 after the treatment. Therefore, it could be concluded that Chat-GPT could effectively enhance students' ability to write narrative texts using folklore. The results of data analysis were shown from the t-test that t-count 16,572 > t-table 1.999 at a significance level of (0.05/2) was 0.025 and Sig. (2-tailed) was 0.000 < 0.05 with a df of 63. Based on the above calculation, it could be seen that the t-count > t-table or Sig. (2-tailed) < 0.05, which meant that Hα was accepted. So, it could be concluded that there was a significant effect of teaching narrative text in the form of folktale through Chat-GPT to enhance students' ability in writing. Based on the results of the questionnaire validity calculation, it showed that the data used was valid because the value of r-count > r-table as described in table 4.18. The results of the reliability calculation on the questionnaire showed that it was reliable or consistent because the value of r-count > r-table as described in table 4.21. So it could be concluded that teaching narrative text in the form of folktale through Chat-GPT had a good effect on students in enhancing their writing skills.

Limitations and Future Direction

This study will analyze the effectiveness of ChatGPT in improving students' writing proficiency, particularly in narrative texts.. The research will involve a specific group of students, such as those in a particular grade level or language proficiency level, to ensure a coherent and manageable sample size. The study will have a specific duration for the implementation of ChatGPT and the assessment of student writing proficiency. However, it is important to acknowledge that the findings may not be applicable universally, as they will be limited to the selected sample and context.

References

Dang, T. H., Chau, T. H., & Tra, T. Q. (2020). a Study on the Difficulties in Writing Argumentative Essays of English-Majored Sophomores At Tay Do University, Vietnam. *European Journal of English Language Teaching*, *6*(1), 201–240. https://doi.org/10.46827/ejel.v6i1.3389

Griffiths, M., & Forcier, L. B. (2016). Intelligence Unleashed. In *Pearson Educación*.

Heaton, J. B. (1988). Writing English Language Tests. In *Longman*.

Heriyudananta, M. (2021). Analisis Kompetensi Menulis Karya Tulis Ilmiah Mahasiswa di Indonesia. *Ascarya: Journal of Islamic Science, Culture, and Social Studies*, *1*(1), 47–55. https://doi.org/10.53754/iscs.v1i1.5

Maphosa, V., & Maphosa, M. (2023). Artificial intelligence in higher education: a bibliometric analysis and topic modeling approach. *Applied Artificial Intelligence*, *37*(1). https://doi.org/10.1080/08839514.2023.2261730

Pinkwart, N. (2016). Another 25 Years of AIED? Challenges and Opportunities for Intelligent Educational Technologies of the Future. *International Journal of Artificial Intelligence in Education*, *26*(2), 771–783. https://doi.org/10.1007/s40593-016-0099-7

Rahmi. (2021). Bab I Pendahuluanيا حض خ ِ ي. *Galang Tanjung*, (2504), 1–9.

Salsabila, V. A. (2020). *the Teaching of Writing Recount Texts*. *6*(1), 132–134. https://doi.org/10.25134/ieflj.v6i1.2637.Received

Sawalha, G., Taj, I., & Shoufan, A. (2024). Analyzing student prompts and their effect on ChatGPT's performance. *Cogent Education*, *11*. https://doi.org/10.1080/2331186X.2024.2397200

Tarigan, H. G. (1989). *Pengajaran analisis kontrastif bahasa*. Proyek Pengembangan Lembaga Pendidikan Tenaga Kependidikan, Direktorat Jenderal Pendidikan Tinggi, Departemen Pendidikan dan Kebudayaan.

Utari, V., & Rambe, R. N. (2023). Analisis Faktor-Faktor Kesulitan Menulis Pada Siswa Kelas Rendah di SD/MI. *Didaktika: Jurnal Kependidikan*, *12*(3), 362–367.

White, F. D. (1986). *The Writer's Art: A Practical Rhetoric and Handbook*. Wadsworth Publishing Company.

Windarti, W., Bintoro, T., & Fahrurrozi, F. (2018). Improving Descriptive Writing Capabilities in Class Students V School Based Islam Integrated Al Istiqomah Karawang through Scaffolding Approach. *American Journal of Educational Research*, *6*(8), 1093–1097. https://doi.org/10.12691/education-6-8-5

Woo, D. J., Wang, D., Yung, T., & Guo, K. (2024). Effects of a Prompt Engineering Intervention on Undergraduate Students' AI Self-Efficacy, AI Knowledge and Prompt Engineering Ability: A Mixed Methods Study. In *International Journal of Artificial Intelligence in Education*. https://doi.org/10.13140/RG.2.2.32577.36968