

The Effectiveness of ZEP Quiz Game-Based Learning in Improving EFL Students' Descriptive Text Reading Comprehension

Savira Najwa Awwala Arkham¹⁾, Olyvia Revalita Candraloka²⁾

^{1, 2} Universitas: Universitas Islam Nahdlatul Ulama Jepara.

Email Correspondence: 221320000709@unisnu.ac.id, olyviarevalita@unisnu.ac.id.

ARTICLE INFO

Keywords:
EFL reading
comprehension
Descriptive text
Game-based learning
ZEP Quiz
Constructivist learning

Article History:

Received : 25 - 06- 2026
Revised : 08- 07 2026
Accepted : 10 - 08 2026
Published: 20- 08 2026

ABSTRACT

EFL students often encounter difficulties in comprehending descriptive texts, highlighting the need for instructional approaches that actively engage learners in the reading process. Despite the growing adoption of game-based learning platforms in EFL education, empirical evidence regarding the effectiveness of ZEP Quiz in improving descriptive text reading comprehension remains limited. Drawing on constructivist theory, this study investigated whether the use of ZEP Quiz, a game-based learning platform, resulted in a significant difference in students' reading comprehension improvement compared with conventional instruction. A quantitative approach with a quasi-experimental, non-equivalent control group design was employed at SMAN 1 Mayong. The participants consisted of 67 tenth-grade students, including 33 students in the experimental group and 34 students in the control group. Data were collected using a validated 22-item reading comprehension test and analyzed through descriptive statistics, the Shapiro-Wilk normality test, Levene's test of homogeneity, and the Mann-Whitney U test on gain scores. The findings indicated that the control group obtained a higher mean gain score than the experimental group. However, the difference in students' reading comprehension improvement was not statistically significant ($Z = -1.209$, $p = .227$). These findings suggest that the implementation of ZEP Quiz did not result in significantly greater improvement in reading

comprehension than conventional instruction within the context of this study. Nevertheless, ZEP Quiz may still serve as a complementary instructional tool to promote student engagement when supported by appropriate instructional guidance.

1. Introduction

Reading comprehension is a fundamental component of English as a Foreign Language (EFL) learning, enabling students to construct meaning from written texts, acquire contextual vocabulary, and support the development of productive language skills such as writing and speaking (Lutvita & Zuhri, 2025; Salim et al., 2025). This ability becomes more crucial at the senior high school level since students must understand a variety of academic materials from various areas. Beyond the capacity to decipher written symbols, reading comprehension is a dynamic and interactive cognitive process in which readers actively create meaning by fusing textual material with linguistic proficiency, past knowledge, and continuous comprehension monitoring. Measurable reading subskills, such as recognizing key concepts, finding specific information, drawing conclusions, understanding terminology in context, and identifying grammatical allusions to maintain textual coherence, reflect this complicated process in educational situations (William, 2009). Despite its importance, many Indonesian EFL students continue to experience difficulties in mastering these essential reading skills, creating a gap between curriculum expectations and students' actual reading performance.

These persistent reading difficulties are influenced by multiple factors, including limited vocabulary knowledge, insufficient reading strategies, and instructional practices that provide limited opportunities for active learning. Students frequently encounter difficulties in identifying implicit information, making appropriate inferences, and understanding texts because of limited contextual knowledge, strategic awareness, and vocabulary mastery (Lorena et al., 2022; Monalisa et al., 2025). These challenges are often reinforced by teacher-centered instructional approaches that rely heavily on repetitive written exercises, reducing reading activities to passive learning experiences and limiting students' motivation and classroom engagement (Monalisa et al., 2025). These flaws become especially apparent when students encounter specific text genres required by the senior high school curriculum, such as descriptive writings. Such difficulties become particularly evident when students learn descriptive texts, one of the text genres required in the Indonesian senior high school English curriculum. Descriptive texts require students not only to understand the generic structure consisting of identification and description but also to interpret specific linguistic

features, including descriptive adjectives, relational verbs, simple present tense, and expanded noun phrases(Purnamasari et al., 2021). Consequently, students need instructional approaches that support active meaning construction while encouraging them to plan, monitor, and evaluate their own comprehension processes. Such metacognitive engagement is essential for helping learners comprehend increasingly complex texts and develop greater autonomy in reading(Ferdiyanto et al., 2024).

To address these challenges, English language instruction should move beyond teacher-centered practices toward learner-centered approaches grounded in Constructivism. Constructivist theory views learning as an active process in which students construct knowledge by integrating new information with their prior knowledge and experiences(Chand, 2023). Rather than passively receiving information, learners actively interpret, reorganize, and refine their understanding through meaningful interaction with learning tasks. Within this perspective, learning is facilitated through continuous reflection and instructional scaffolding, which provides temporary support to help students bridge the gap between their existing knowledge and potential development(Sudarman et al., 2024). The integration of digital technology further strengthens constructivist learning by creating interactive environments that encourage learner participation, immediate feedback, and collaborative knowledge construction (Olusegun, 2015 ;Setiawan, 2023). Consequently, the incorporation of instructional content into interactive game mechanics has contributed to the development of Game-Based Learning (GBL), an instructional approach that promotes experiential learning through challenges, immediate feedback, and meaningful engagement with learning activities(Nguyen et al., 2024;Shaheen et al., 2023). Previous studies have also shown that game-based learning can enhance students' engagement, collaboration, and long-term knowledge retention when implemented appropriately(Dewi et al., 2025 ; Fatimah, 2025; Mahaswary & Ngafif, 2025).

Among emerging game-based learning platforms, ZEP Quiz has attracted increasing attention because it enables teachers to design interactive learning activities within a virtual environment featuring customizable avatars, timed challenges, and real-time leaderboards(Santoso et al., 2025).These gamification features create immersive learning experiences that encourage active participation while reducing learning anxiety(Hwang et al., 2012;Kapp, 2012). Moreover, the platform provides immediate feedback that functions as digital scaffolding, allowing students to identify misconceptions, revise their understanding, and monitor their learning processes in real time(Kapp, 2012; Zainuddin et al., 2020).ZEP Quiz also supports multiple question formats, enabling teachers to assess students' understanding efficiently while maintaining manageable cognitive demands. Previous studies conducted in various educational contexts have

reported positive pedagogical outcomes associated with ZEP Quiz, including improvements in learning motivation, classroom engagement, and higher-order thinking skills (Widyaningrum et al., 2025; Yulianti et al., 2025; Ratnawati & Wahyuni, 2025). These findings suggest that ZEP Quiz has considerable instructional potential; however, further empirical investigation is required to determine whether its benefits are consistently reflected in students' reading comprehension improvement within EFL classrooms.

Recent empirical evidence generally supports the effectiveness of game-based learning in English language education. Studies employing Quizizz through pre-experimental designs have reported improvements in students' reading comprehension by promoting sustained engagement, immediate feedback, and repeated practice among junior high school and university students (Annisa & Susanti, 2024; Hamzah 2023). Likewise, research on ZEP Quiz has demonstrated positive effects across different learning outcomes. Wahid et al. (2025) using a quasi-experimental pre-test-post-test design, found that ZEP Quiz significantly improved students' vocabulary mastery, while Muslimin (2025) reported enhanced critical thinking skills among elementary school students following its implementation. Santoso et al. (2025) further reported improved narrative reading comprehension among students who learned through ZEP Quiz. Broader evidence also indicates that constructivist digital learning environments support reading development through scaffolded learning strategies, interactive assessment, and active student participation (Kharimah et al., 2025; Salim & Sukarno, 2025; Marsa, 2021). Collectively, these studies consistently report positive educational outcomes; however, they differ considerably in terms of participant characteristics, educational levels, learning objectives, instructional duration, and research design. Consequently, the extent to which these positive findings can be generalized across different EFL instructional contexts remains insufficiently understood. Moreover, previous studies have primarily emphasized successful learning outcomes while providing limited discussion of the contextual and pedagogical conditions that may influence the effectiveness of game-based learning.

This unresolved issue represents an important theoretical gap, as the effectiveness of game-based learning may depend not only on the technological features of the platform but also on how it is implemented within particular instructional settings. Therefore, the present study investigates whether the use of ZEP Quiz results in significantly greater improvement in descriptive text reading comprehension than conventional instruction among Indonesian senior high school EFL students. By addressing this issue, the study contributes additional empirical evidence regarding the consistency of game-based learning outcomes in EFL reading instruction.

Building upon this theoretical gap, the present study examines whether the use of ZEP Quiz results in significantly greater improvement in descriptive text reading comprehension than conventional instruction among tenth-grade Indonesian EFL students. Guided by a constructivist perspective, this study contributes to the growing body of research on digital game-based learning by extending existing evidence beyond the consistently positive findings reported in previous studies. Rather than assuming that game-based learning is universally effective, this study provides empirical evidence regarding the extent to which the effectiveness of ZEP Quiz can be observed within a different instructional context and text genre. The findings are expected to enrich the literature on technology-enhanced language learning while offering practical insights for EFL teachers in selecting and implementing digital game-based learning platforms to support reading comprehension instruction in secondary education.

2. Methods

To carry out this investigation, a quasi-experimental framework specifically employing a non-equivalent control group design was explicitly implemented within a quantitative methodology. The researcher chose this specific approach due to the practical and administrative limitations typical of school environments because school authorities prevent the random assignment of individual students into new groups. Through the strategic utilization of this method the researcher successfully examined the actual impact of ZEP Quiz on students' reading comprehension of descriptive texts while fully preserving both the natural classroom structure and the overall teaching-learning atmosphere. The researcher administered a pre-test to both groups to evaluate their initial reading comprehension abilities as a fundamental step to implement this design. While the researcher gave the experimental group focused reading instruction using the ZEP Quiz platform the control group underwent traditional text-based teaching under identical conditions. At the conclusion of the instructional period the researcher conducted a post-test for both groups to assess the final differences and to gauge any notable enhancements in their reading comprehension abilities.

2.1. Research Design and Data Sources

The population of this research was explicitly constituted by all tenth-grade students at SMAN 1 Mayong in the academic year of 2025/2026 just as the selection of two classes to serve as the research sample was directly facilitated by a purposive sampling technique where the total intended sample size was successfully brought to 72 participants because each class initially contained 36 students. The researcher designated one class as the experimental group to be taught using the ZEP Quiz platform while the other class simultaneously served as the control group instructed through conventional methods. However, several

students were absent during the actual field execution of the pre-test and post-test due to illness or school-related obligations. To ensure strict data integrity and statistical validity, the researcher included only students who completed both testing phases in the final data computation, resulting in a final active sample size of 33 students in the experimental group and 34 students in the control group, for a total of 67 active participants for statistical analysis.

2.2. Research Instruments

This study utilized a reading comprehension test developed from normal EFL materials focusing primarily on descriptive texts as its major data collection tool. A pool of 50 multiple-choice problems initially formed the prototype to effectively assess key reading competencies including identifying primary concepts and locating specific details as well as interpreting language in context and drawing inferences. To guarantee the instrument's quality, the researcher tested the items with a non-participant tenth-grade class of 33 students before statistically analyzing the pilot data for item validity using the Pearson Product-Moment correlation formula where a correlation coefficient (r) higher than the critical table value of 0.344 deemed an item valid. Furthermore, the Cronbach's Alpha formula evaluated the internal consistency and reliability of the test so that the researcher retained only those items demonstrating a reliability coefficient of 0.60 or higher alongside meeting the validity criteria for the final pre-test and post-test instruments.

2.3. Data Collection Procedure

Data collection was conducted systematically using a standardized five-meeting experimental procedure corresponding to each class weekly schedule of three instructional periods (135 minutes). During the first meeting, both the experimental and control groups completed the pre-test to determine their initial reading comprehension levels. The following three meetings were devoted to the instructional treatment. The experimental group learned descriptive texts using ZEP Quiz, whereas the control group received conventional teacher-led instruction covering the same learning objectives and instructional materials. To ensure treatment fidelity, both groups followed comparable lesson plans, received the same instructional duration, and completed equivalent learning objectives. All instructional sessions were conducted by the researcher, who maintained detailed teaching logs and attendance records after each meeting to monitor the consistency of the intervention throughout the study. During the fifth meeting, both groups completed the post-test using the same validated instrument to measure students' reading comprehension after the intervention. Prior to data collection, formal permission to conduct the study was obtained through an official research permit issued by the university and approved by the school administration of SMAN 1 Mayong. The English teacher was also informed about

the study procedures and the implementation schedule. Before the intervention began, students were informed of the purpose and procedures of the study, and all classroom activities were conducted as part of their regular English lessons. Participation involved no additional academic risk beyond normal classroom instruction. All collected data were kept confidential, anonymized during analysis, and used solely for research purposes.

2.4. Data Analysis

Data generated from the pre-test and post-test administrations were processed utilizing IBM SPSS Statistics through both descriptive and inferential analytical methods. Initially, descriptive statistics specifically mean scores and standard deviations were calculated to provide an overview of student reading comprehension achievements across the control and experimental groups. To ensure the data fulfilled prerequisite conditions before conducting hypothesis testing, the Shapiro-Wilk procedure was executed to check for normality, and Levene's test was implemented to evaluate variance homogeneity. Based on the results of these assumption tests, the Mann-Whitney U test was employed to compare the pre-test and post-test scores between the two groups whenever the assumptions for parametric analysis were not fulfilled. To evaluate students' learning improvement, Gain scores from the corresponding post-test score (Gain = Post test - Pretest). The gain scores were first subjected to descriptive statistics and diagnostic testing using the Shapiro-Wilk and Levene's methods. Thereafter, the non-parametric Mann-Whitney U test was deployed to evaluate the difference in reading comprehension gains between the Zep Quiz intervention group and the conventional instruction group. For all conducted analyses, the significance threshold was fixed at $p < .05$.

3. Result

The data collected from the experimental and control groups during the study directly form the basis of the results presented in this section. The researcher conducted several analyses including instrument testing and descriptive statistical analysis as well as prerequisite testing and hypothesis testing to ensure the absolute accuracy and validity of the findings. Furthermore, a systematic arrangement structures these findings so that they provide a comprehensive overview regarding the specific effect of ZEP Quiz on students' reading comprehension achievement.

3.1 Results of the Research Instrument Trial

3.1.1 Validity Test

The validity analysis was conducted to evaluate the appropriateness of the reading comprehension test items before they were administered in the main study. The instrument initially consisted of 50 multiple choice items. Based on the

validity test results, 22 items met the validity criteria and were retained for the research instrument, while the remaining 28 items were excluded. The summary of the validity test results is presented in Table 1.

Table 1. Validity Test Results of the Research Instrument

Category	Item Numbers	Total
Valid	5, 7, 9, 12, 14, 17, 27, 28, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 42, 43, 48, 49	22
Invalid (Dropped)	1, 2, 3, 4, 6, 8, 10, 11, 13, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 29, 40, 41, 44, 45, 46, 47, 50	28
Total		50

Following the validity test, the remaining 22 items were subjected to a reliability analysis using Cronbach's Alpha to examine the internal consistency of the instrument. The analysis produced a Cronbach's Alpha coefficient of .760, indicating good internal consistency. Therefore, the instrument was considered sufficiently reliable for measuring students' reading comprehension of descriptive texts.

3.1.2 Descriptive statistical analysis

Evaluating reading comprehension levels before and after the instruction involved computing descriptive statistics across both student groups. The resulting data encompassed pre- and post-test metrics for each group, specifically detailing the total sample count, lowest and highest obtained scores, mean performances, and standard deviations.

Table 2. Descriptive Statistic for Pre-Test and Post -Test Score

	Descriptive Statistics				
	N	Minimum	Maximum	Mean	Std. Deviation
Experimental Group Pretest	33	55	95	79.03	10.147
Experimental Group Posttest	33	36	100	78.48	13.570
Control Group Pretest	34	55	100	84.74	12.556
Control Group Posttest	34	73	100	91.12	5.902
Valid N (listwise)	33				

The descriptive statistics indicate different patterns of performance between the experimental and control groups. In the experimental group, the mean score

decreased slightly from 79.03 on the pretest to 78.48 on the post-test. The minimum score dropped from 55 to 36, while the maximum score rose from 95 to 100. Furthermore, the standard deviation increased from 10.147 to 13.570, indicating greater variability in students' post-test scores compared to their pre-test performance.

In contrast, overall performance among control group members improved significantly because their mean score rose from 84.74 on the pre-test to 91.12 on the post-test and their minimum score increased from 55 to 73 while the maximum score stayed at 100 in both tests, which ultimately led to a decrease in the standard deviation from 12.556 to 5.902 and thus resulted in far more consistent student outcomes overall.

Gain scores were determined after the intervention by subtracting each student's pre-test score from the post-test score (Gain = Post-test - Pre-test). The gain scores were then employed in inferential analyses to examine the increase in reading comprehension between the experimental and control groups.

To examine students' improvement after the intervention, gain scores were calculated by subtracting the pre-test score from the post-test score for each participant. Table 3 summarizes the descriptive statistics of the gain scores for both groups.

Table 3. Descriptive Statistics of Gain Scores

Descriptives

Class = Experimental

Descriptive Statistics ^a					
	N	Minimum	Maximum	Mean	Std. Deviation
Pretest Score	33	55	95	79.03	10.147
Posttest Score	33	36	100	78.48	13.570
Gain_Score	33	-55.00	14.00	-.5455	13.77745
Valid N (listwise)	33				

a. Class = Experimental

Class = Control

Descriptive Statistics ^a					
	N	Minimum	Maximum	Mean	Std. Deviation
Pretest Score	34	55	100	84.74	12.556
Posttest Score	34	73	100	91.12	5.902
Gain_Score	34	-13.00	27.00	6.3824	11.07007
Valid N (listwise)	34				

a. Class = Control

A mean gain score of -0.55 (SD= 13.78) was recorded for the experimental groups, in contrast to a mean gain score of 6.38 (SD= 11.07) for the control groups. These outcomes reveal that, overall participants in the control class

exhibited a higher degree of reading comprehension growth compared to their peers in the experimental class.

3.2 Assumption Test

Prior to analyzing the initial performance differences between the two classes, data normality was evaluated via the Shapiro-Wilk procedure. This step served to verify whether the pre-test distribution complied with the distributional requirements needed to perform parametric evaluations.

Table 4. Normality Test Pretest Results Using Shapiro-Wilk

Tests of Normality							
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Class	Statistic	df	Sig.	Statistic	df	Sig.
Pretest Score	Experimental	.161	33	.030	.954	33	.172
	Control	.279	34	.000	.869	34	.001

a. Lilliefors Significance Correction

The Shapiro-Wilk test revealed that the experimental cohort's pre-test data had a normal distribution ($p = .172$, $p > .05$), whereas the control cohort's data varied considerably from normality ($p = .001$, $p < .05$). Due to this partial violation of the normality condition, a non-parametric method was chosen as the best analytical framework for further data processing.

Table 5. Homogeneity Pretest

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
Pretest Score	Equal variances assumed	2.389	.127	-2.042	65	.045	-5.705	2.794	-11.285 -1.125
	Equal variances not assumed			-2.048	62.957	.045	-5.705	2.785	-11.271 -.139

The homogeneity of the pre-test scores was investigated using Levene's test for equality of variances. The significance value was .127 ($p > .05$), showing that the experimental and control groups' variances were similar. Although the condition of homogeneity was met, the normalcy assumption was not entirely met. The subsequent analysis used a non-parametric statistical test.

Table 6. Normality Gain Score

Tests of Normality							
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Class	Statistic	df	Sig.	Statistic	df	Sig.
Gain_Score	Experimental	.243	33	.000	.769	33	.000
	Control	.203	34	.001	.907	34	.007

a. Lilliefors Significance Correction

The Shapiro-Wilk test was applied to evaluate whether the gain scores followed a normal distribution. Results revealed a significance value of .000 ($p < .05$) in the treatment group and .007 ($p < .05$) in the control group, confirming that the normality assumption was violated across both participants groups.

Table 7. Homogeneity of Gain Scores

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
Gain_Score	Equal variances assumed	.012	.913	-2.272	65	.026	-6.92781	3.04885	-13.01678 -83883
	Equal variances not assumed			-2.265	61.320	.027	-6.92781	3.05882	-13.04365 -81197

Homogeneity of variance for the gain scores was confirmed through Levene's test, which yielded a significant level of .913 ($p > .05$). while the variance equality assumption was successfully met, the data failed to satisfy the normality criteria. Consequently, a non-parametric statistical method was selected to execute the comparative analysis.

3.3 Hypothesis Test

Based on the results of the assumption tests, the Mann-Whitney U test was used to compare the reading comprehension performance of the experimental and control groups. The test was first conducted on the pre-test scores to examine whether the two groups had comparable initial reading comprehension performance before the intervention. It was then applied to the gain scores to determine whether the improvement in reading comprehension differed significantly between the two groups.

Table 8. Mann-Whitney U Test for the Pretest Scores

Test Statistics^a

	Pretest Score
Mann-Whitney U	373.000
Wilcoxon W	934.000
Z	-2.379
Asymp. Sig. (2-tailed)	.017

a. Grouping Variable: Class

The Mann-Whitney U test revealed that the control group had a higher mean rank (39.53) than the experimental group (28.30). The analysis revealed an Asymp. Sig. (2-tailed) value of .017 ($p < .05$), showing a statistically significant difference between the two groups prior to intervention. These data indicate that the participants did not start the study with equal levels of reading comprehension. After examining the initial differences between the two groups, the analysis focused on the gain scores to evaluate whether the improvement in reading comprehension differed following the implementation of the instructional treatment.

Table 9. Mann-Whitney U Test for the Gain Scores

Test Statistics^a

	Gain_Score
Mann-Whitney U	465.500
Wilcoxon W	1026.500
Z	-1.209
Asymp. Sig. (2-tailed)	.227

a. Grouping Variable: Class

The Mann-Whitney U test showed that the control group had a somewhat higher mean rank (36.81) than the experimental group (31.11). However, the Asymp. Sig. (2-tailed) value was .227 ($p > .05$), indicating that there was no statistically significant difference in gain scores between the two groups. As a result, the data indicate that the progress in reading comprehension was not substantially different between students who learnt via ZEP Quiz and those who got traditional education.

4. Discussion

This study aimed to examine whether there was a significant difference in the improvement of students' reading comprehension between those who learned using ZEP Quiz and those who received conventional instruction. Based on the gain score analysis, the Mann-Whitney U test indicated no statistically significant

difference in students' reading comprehension improvement between the two groups ($p = .227$). Although the control group obtained a higher mean gain score than the experimental group, the difference was not statistically significant. In addition, the findings should be interpreted with consideration of the pre-test results, which showed that the two groups differed significantly in their initial reading comprehension performance before the treatment was implemented.

From a constructivist perspective, these findings suggest that students' active participation in digital learning activities does not necessarily lead to greater improvement in reading comprehension. Constructivist theory emphasizes that learning occurs when students actively construct knowledge by integrating new information with their prior knowledge and experiences. In this study, ZEP Quiz was expected to facilitate this process by promoting active engagement, immediate feedback, and interactive learning activities. However, the absence of a statistically significant difference in gain scores indicates that students' engagement with the platform alone was not sufficient to produce greater improvement in reading comprehension than conventional instruction. As argued by (Le, 2024), the effectiveness of constructivist learning depends not only on students' participation but also on the quality of instructional support and the learning experiences provided. Therefore, while ZEP Quiz encouraged students to participate actively during classroom activities, it may not have fully supported the deeper cognitive processes required for comprehending descriptive texts.

The findings also provide further insight into the implementation of game-based learning in English language teaching. Previous studies have reported that game-based learning can enhance students' motivation, classroom participation, and academic achievement (Wahid et al., 2025; Santoso et al., 2025; Salim & Sukarno, 2025). However, the results of the present study did not demonstrate a significantly greater improvement in reading comprehension among students who learned using ZEP Quiz. One possible explanation is that reading comprehension requires learners to simultaneously process textual information, identify key ideas, make inferences, and integrate new information with their prior knowledge. When game elements are introduced, students may also need to allocate attention to completing game-related tasks. From the perspective of Cognitive Load Theory, this additional demand may increase learners' cognitive load and reduce the cognitive resources available for processing the learning content. As noted by Mojarrad et al. (2022), students' engagement in game-based learning does not necessarily translate into improved learning outcomes. Similarly, Faber et al. (2024) argued that the cognitive demands created by gaming mechanics may limit learners' ability to process instructional content effectively. Unlike previous studies that reported significant improvements following the implementation of ZEP Quiz or other game-based learning platforms, the present study found no statistically

significant difference in students' reading comprehension improvement. This discrepancy suggests that the effectiveness of game-based learning may be influenced by contextual factors such as learners' initial reading ability, instructional implementation, and the complexity of the targeted reading skills rather than by gamification features alone.

Another factor that may have contributed to the findings is students' adaptation to the learning environment. Although digital platforms can create engaging learning experiences, students may require sufficient time to become familiar with new learning technologies and instructional approaches. In addition, reading comprehension is a complex cognitive process that involves identifying relevant information, making inferences, and integrating ideas within a text. These processes may require continuous instructional guidance in addition to interactive learning activities. Consequently, the absence of a statistically significant difference in gain scores may reflect that the implementation of ZEP Quiz alone was not enough to produce greater improvement in reading comprehension within the duration of this study.

The findings of this study contribute to the growing body of research on game-based learning in English language teaching by demonstrating that increased student engagement does not necessarily result in significantly greater improvement in reading comprehension. Rather, the effectiveness of digital learning platforms appears to depend on how well they support students' cognitive processing and comprehension of the learning materials. Therefore, ZEP Quiz may serve as a valuable instructional tool for creating interactive and engaging classroom activities when accompanied by appropriate instructional guidance. Future research is encouraged to examine the implementation of ZEP Quiz over a longer instructional period, across different educational contexts, and with larger participant groups to provide a more comprehensive understanding of its effectiveness in improving students' reading comprehension. These findings imply that the successful implementation of game-based learning depends not only on the incorporation of gamified features but also on instructional designs that effectively manage students' cognitive load while providing sufficient scaffolding for meaning construction. Therefore, digital game-based learning platforms should be integrated with structured reading activities and teacher guidance to maximize meaningful learning rather than relying solely on student engagement.

5. Conclusion

This study examined whether the use of ZEP Quiz resulted in a significantly different improvement in tenth-grade students' descriptive text reading comprehension compared with conventional instruction. The findings showed that, although the control group obtained a higher mean gain score than the experimental group, the Mann-Whitney U test indicated no statistically significant

difference in reading comprehension improvement between the two groups. These results imply that, in the context of this study, the use of ZEP Quiz did not result in greater improvement in students' reading comprehension than conventional instruction. The findings further suggest that meaningful learning requires not only active student engagement but also appropriate instructional guidance and support for deeper cognitive processing. The study also contributes to the current understanding of game-based learning in EFL classrooms by demonstrating that gamification alone does not necessarily lead to improved learning outcomes. Instead, the findings reinforce constructivist perspectives that emphasize the importance of instructional scaffolding while supporting Cognitive Load Theory, which highlights the need to manage learners' cognitive resources during technology-enhanced learning. Therefore, rather than serving as a stand-alone strategy for improving reading comprehension, ZEP Quiz may be better utilized as a complementary instructional tool integrated with structured reading activities and teacher guidance. Future studies are recommended to investigate the implementation of ZEP Quiz over a longer instructional period, across different educational contexts, and with larger samples to provide a more comprehensive understanding of its effectiveness in supporting students' reading comprehension.

References

- Annisa, N., & Susanti, A. (2024). The effect of Quizizz application on the students' reading comprehension. *Wiralodra English Journal*, 8(1), 1–12.
<https://doi.org/10.31943/wej.v8i1.244>
- Chand, S. P. (2023). *Constructivism in Education : Exploring the Contributions of Piaget, Vygotsky, and Bruner*. 12(7), 274–278.
<https://doi.org/10.21275/SR23630021800>
- Dewi, E. P., Tahrur, & Masagus Firdaus. (2025). Improving Reading Comprehension Use Quizizz Application and Cognitive Style. *Esteem Journal of English Education Study Programme*, 8(2), 507–517.
<https://doi.org/10.31851/esteem.v8i2.18196>
- Faber, T. J. E., Dankbaar, M. E. W., Broek, W. W. Van Den, Bruinink, L. J., Hogeveen, M., & Merriënboer, J. J. G. Van. (2024). Effects of adaptive scaffolding on performance, cognitive load and engagement in game - based learning : a randomized controlled trial. *BMC Medical Education*.
<https://doi.org/10.1186/s12909-024-05698-3>
- Fatimah, S. (2025). Enhancing Junior High School Students' Reading Comprehension through Quizizz-Based Gamified Assessment. *ELLINE Journal: English Language and Literature in Education Journal*, 3(1), 38–57.
<https://doi.org/https://doi.org/10.63011/tyn8x842>
- Ferdiyanto, F., Islam, U., & Hasan, Z. (2024). *THE ROLE OF METACOGNITIVE SKILLS IN IMPROVING READING COMPREHENSION AMONG*. 6(2), 149–160.
- Hamzah, I. (2023). *The Effect Of Quizizz On Students Reading Comprehension Universitas Teknokrat Indonesia*. 3(1), 51–58.
- Hwang, G. J., Wu, P. H., & Chen, C. C. (2012). *An Online Game Approach for Improving*

- Students' Learning Performance in Web- based Problem-Solving Activities.*
<https://doi.org/https://doi.org/10.1016/j.compedu.2012.05.009>
- Kapp, K. M. (2012). *The Gamification of Learning and Instruction: Game-based Methods and Strategies for Training and Education.*
<https://books.google.co.id/books?id=M2Rb9ZtFxccC&hl=id>
- Kharimah, I., Nirwanto, R., & Mirza, A. A. (2025). Exploring The Teacher's Strategies In Teaching Reading Comprehension In Senior High School. *Journal of English Language and Education*, 10(3), 751–762.
<https://doi.org/10.31004/jele.v10i3.959>
- Le, H. Van. (2024). *Promoting L2 learners' critical thinking skills : the role of social constructivism in reading class.* June, 1–12.
<https://doi.org/10.3389/feduc.2024.1241973>
- Lorena, N., Budiyanto, D., Tridianti, G., & Palembang, U. T. (2022). *Applying Posse Strategy in Teaching Reading Descriptive.* 3(1), 34–42.
- Lutvita, F. R., & Zuhri, F. (2025). the Effectiveness of Quizizz in Enhancing Junior High School Students' Reading Comprehension of Narrative Texts. *Wiralodra English Journal*, 9(2), 211–222. <https://doi.org/10.31943/wej.v9i2.317>
- Mahaswary, A. N., & Ngafif, A. (2025). *The Effectiveness of Quizizz Application in Teaching Reading Comprehension.* 8.
<http://jiip.stkipyapisdampu.ac.id/index.php/JIIP>
- Marsa, S. S. (2021). The Effect of Kahoot ! Game to Teaching Reading Comprehension Achievement. *Journal of English Teaching*, 7(2021), 133–149.
<https://doi.org/https://doi.org/10.33541/jet.v7i2.2738>
- Mojarrad, H., Hemmati, F., Gohar, M. J., & Sadeghi, A. (2022). *COMPUTER-BASED ASSESSMENT (CBA) VS . PAPER / PENCIL-BASED ASSESSMENT (PPBA) : AN INVESTIGATION INTO THE PERFORMANCE AND ATTITUDE OF IRANIAN EFL LEARNERS ' READING COMPREHENSION.* 4(December), 418–428.
<https://l1nq.com/j9au2hn>
- Monalisa, C. D., Wardah, & Rezeki, Y. S. (2025). Improving Students ' Reading Skill Through. *JEELL (Journal of English Education, Linguistics and Literature)*, 12 No.1. <https://doi.org/https://doi.org/10.32682/jeell.v11i02/14>
- Muslimin. (2025). Integrating Zep Quiz into IPAS Instruction to Enhance Fourth Grade Students ' Critical Thinking Skills. *Journal of Education and Computer Applications*, Vol,2 No.2(ISSN: 3047-8308), 30–37.
<https://doi.org/https://doi.org/10.69693/jeca.v2i2.34>
- Nguyen, T. N., Phan Thanh Tuan, & Nguyen Thi Ha Phuong. (2024). Engaging Students in The Learning Process with Game-Based Learning. *International Journal of English Language Studies*, 6(1), 54–60.
<https://doi.org/10.32996/ijels.2024.6.1.5>
- Olusegun, S. (2015). *Constructivism Learning Theory : A Paradigm for Teaching and Learning.* 5(6), 66–70. <https://doi.org/10.9790/7388-05616670>
- Purnamasari, D., Hidayat, D. N., & Kurniawati, L. (2021). An Analysis of Students' Writing Skill on English Descriptive Text. *English Education: Jurnal Tadris Bahasa Inggris*, 14(1), 101–114. <https://doi.org/10.24042/ee-jtbi.v14i1.7943>
- Ratnawati, D., & Wahyuni, S. (2025). Journal of Educational Sciences. *Journal of Educational Sciences*, 6(3), 444–458.
<https://doi.org/https://doi.org/10.31258/jes.10.1.p.468-479>

- Salim, M., & Sukarno, K. M. (2025). *Using Gamified Quizzes to Enhance Students' English Reading Comprehension Skills*. 33(1), 125–142.
<https://doi.org/doi.org/10.21381/diksi.v33i1.74744>
- Salim, M., Sukarno, K. M., Robbaniyah, I., Fauzan, M., Widyastuti, S. H., Bariyyah, K., & Hidayati, A. N. (2025). *Utilizing Gamified Quizzes to Enhance Students' English Reading Comprehension Skills*. 1, 136–144.
<https://doi.org/https://doi.org/10.36987/jes.v12i1.6555>
- Santoso, D., Tamara, P., Nabhila, C., Sujarwo, G., & Fadlilah, N. (2025). ZepQuiz Game-Based Learning for Improving Students' Narrative Reading Skills: ZepQuiz: Pembelajaran Berbasis Permainan untuk Meningkatkan Keterampilan Membaca Naratif Siswa. *Academia Open*, 10 No.1.
<https://doi.org/https://doi.org/10.21070/acopen.10.2025.12792>
- Setiawan, B. (2023). Game Based Learning Design and Development for Effective Instructional Process at Senior High School Level. *Jurnal Ilmiah Pendidikan Dan Pembelajaran*, 7(3), 425–432. <https://doi.org/10.23887/jipp.v7i3.58391>
- Shaheen, A., Ali, S., & Fotaris, P. (2023). Assessing the Efficacy of Reflective Game Design: A Design-Based Study in Digital Game-Based Learning. *Education Sciences*, 13(12). <https://doi.org/10.3390/educsci13121204>
- Sudarman, R., Damaianti, V. S., & Mulyati, Y. (2024). *Constructivism in Writing Learning in Elementary School ; Systematic Literature Review*. 4(1), 176–192.
- Wahid, R., Putra, P., Karimullah, I. W., & Mustofa, M. (2025). *Zep Quiz in Action : Empowering Students' Vocabulary through Gamified Learning*. 4778, 5461–5473. <https://doi.org/10.24256/ideas>
- Widyaningrum, D., Smaragdina, A. A., & Rizky, D. D. (2025). MENINGKATKAN MOTIVASI BELAJAR SISWA MENGGUNAKAN ZEP QUIZ PADA PEMBELAJARAN INFORMATIKA DI SMP Dhea. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 10(02). <https://doi.org/https://doi.org/10.23969/jp.v10i02.28792>
- William, G. (2009). *Reading in a Second Language: Moving from Theory to Practice*. Cambridge University Press.
https://books.google.co.id/books?hl=id&lr=&id=prvRHZ7DrIcC&oi=fnd&pg=PR11&dq=grabe+2009&ots=DEknQmZU-u&sig=pHK96mGNXL7cDNxCBpqp8QQffLM&redir_esc=y#v=onepage&q=grabe+2009&f=false
- Yulianti, T., Insani, N. H., & Sukoyo, J. (2025). EFEKTIVITAS MEDIA PEMBELAJARAN QUIZ ZEP TERHADAP HASIL BELAJAR KOGNITIF BAHASA JAWA SISWA SMA. *Jurnal Ilmiah Pendidikan Citra Bakti*, 12(November), 1170–1183. <https://doi.org/https://doi.org/10.38048/jipcb.v12i4.5627>
- Zainuddin, Z., Cu, S. K. W., Shujahat, M., & Perera, J. (2020). *The Impact of Gamification on Learning and Instruction: A Systematic Review of Empirical Evidence*. <https://doi.org/https://doi.org/10.1016/j.edurev.2020.100326>

The Role of Drama in Improving Students' Comprehension of Reported Speech.

Rawend F Sitorus ¹⁾, Chairuddin ²⁾, Baihaqi ³⁾

^{1, 2, 3} English Language Education, Teacher Training and Education Faculty, Universitas Samudra

Email Correspondence: rawendfebriansitorus@gmail.com

ARTICLE INFO

Keywords:

[Drama; Role Play; Reported Speech]

Article History:

Received: 11-06-2026

Revised : 08-08-2026

Accepted: 10-08-2026

Published: 20-08-2026

ABSTRACT

This study aimed to examine the role of drama scripts implemented through role-play activities in improving students' comprehension of reported speech at SMK Negeri 2 Langsa. The study employed a quantitative approach using a one-group pre-test and post-test quasi-experimental design. The population consisted of all eleventh-grade students at SMK Negeri 2 Langsa, while the sample comprised 25 students from class XI RPL selected through purposive sampling. The instrument used to collect the data was a grammar test consisting of 25 multiple-choice questions focusing on reported speech. The research procedure included administering a pre-test to measure students' initial comprehension, implementing drama scripts through role-play activities as the treatment, and administering a post-test to measure students' comprehension after the treatment. The results revealed a significant improvement in students' comprehension of reported speech. The mean score increased from 37.52 in the pre-test to 82.24 in the post-test. Furthermore, the paired sample *t*-test showed that the significance value (Sig. 2-tailed) was 0.000, which was lower than 0.05. These findings indicated that the null hypothesis (H_0) was rejected and the alternative hypothesis (H_1) was accepted. Therefore, it was concluded that the implementation of drama scripts through role-play activities improved students' comprehension of reported speech. The findings suggested that drama-based learning activities could serve as an engaging and effective strategy for teaching grammar in EFL classrooms.

1. Introduction

Grammar is crucial for learning English, as it establishes the basis for creating clear and precise communication. By utilizing grammatical knowledge, learners can arrange ideas, convey relationships, and communicate efficiently in both verbal and written formats. Within the almost English as a Foreign Language (EFL), mastering grammar is viewed as equally essential as communicative competence since it allows learners to generate language that is grammatically correct and suitable for the context. Nevertheless, grammar teaching in numerous classrooms remains largely influenced by conventional methods, including rote learning, repetitive tasks, and teacher-focused instruction. These methods frequently offer few chances for students to utilize grammatical knowledge in significant communication. Consequently, numerous students can articulate grammatical rules yet struggle to apply them correctly in actual communication contexts (Normawati & Nugrahaeni, 2024).

One grammatical structure that remains challenging for many EFL learners is reported speech. Reported speech is used to convey what someone has said without quoting the speaker's exact words. This grammatical structure is widely used in spoken and written communication, particularly in academic and professional contexts where information is often summarized or paraphrased (Nst et al., 2025). However, transforming direct speech into reported speech requires several grammatical adjustments, including changes in tense, pronouns, reporting verbs, and expressions of time and place. These structural changes make reported speech one of the grammar topics that students frequently find difficult to understand and apply correctly.

Previous studies have consistently shown that students experience various difficulties in learning reported speech. (Miranti, 2020) found that university students frequently made errors when transforming direct speech into reported speech, particularly in verb tense changes, with many of the errors influenced by intralingual interference. (Wahyudi, 2021) also reported that high school students often encountered problems in changing tenses when converting direct speech into indirect speech, indicating that tense transformation remains one of the most challenging aspects of reported speech. In addition, (Rizki, 2019) found that limited grammatical awareness and insufficient contextual practice contributed to students' difficulties in transforming interrogative sentences into reported speech. These findings indicate that many students still rely heavily on memorizing grammatical rules without fully understanding how those rules function in authentic communication. Consequently, students may understand the concept of reported speech theoretically but continue to experience difficulties when applying it accurately in actual language use. This condition suggests that conventional grammar instruction does not always provide sufficient opportunities for students to practice grammar meaningfully in communicative contexts. As a result, grammar learning often becomes abstract, reducing students' motivation and active participation in the learning process (Altun, 2019).

One instructional approach that has gained increasing attention in English language teaching is drama-based learning. Drama encourages students to learn through active participation, collaboration, and meaningful interaction rather than through

memorization alone. Through activities such as dialogues, role-play, and drama scripts, students are provided with opportunities to use language in situations that resemble real communication. (Anderson & Dunn, 2013) explain that drama supports language learning by engaging students in narrative experiences and collaborative activities that strengthen both language comprehension and retention. Likewise, (Davies, 1990) emphasizes that drama creates a supportive classroom environment in which learners can experiment with the target language without fear of making mistakes. Role-play also enables students to express ideas, negotiate meaning, and practice language structures in authentic communicative situations (Ladousse, 1987). Similarly, (Livingstone, 1983) argues that role-play allows learners to practice grammatical structures in meaningful contexts rather than through isolated drills. Collectively, these studies suggest that drama-based learning can create an interactive learning environment that encourages students to use language meaningfully while developing both their grammatical understanding and communicative competence.

Although the educational benefits of drama have been widely discussed, previous studies have primarily examined its contribution to improving speaking skills, learner motivation, and communicative competence. Comparatively, fewer studies have investigated the use of drama as an instructional strategy for teaching specific grammatical structures. Likewise, previous studies on reported speech have mainly focused on identifying students' grammatical difficulties rather than exploring instructional approaches designed to address those difficulties through interactive learning (Kolar & Sándorová, 2025). Consequently, there is still limited evidence regarding how drama scripts implemented through role-play activities can support students' comprehension of reported speech in EFL classrooms. This gap indicates the need for further investigation into the use of drama-based learning as a contextual approach to grammar instruction. Based on this gap, the present study examines the role of drama scripts implemented through role-play activities in improving students' comprehension of reported speech among eleventh-grade students at SMK Negeri 2 Langsa. Unlike previous studies that mainly investigated drama as a means of improving speaking ability, learner motivation, or general communicative competence, this study focuses specifically on its application in grammar instruction. The novelty of this study lies in integrating structured drama scripts with role-play activities to facilitate students' understanding of reported speech through contextualized language practice. By providing empirical evidence on the role of drama-based learning in teaching a specific grammatical structure, this study contributes to the development of more interactive and meaningful approaches to grammar instruction in EFL classrooms and offers practical insights for English teachers seeking alternative strategies for teaching reported speech.

2. Method

This study employed a quantitative approach using a quasi-experimental design with a one-group pre-test and post-test design. The design was selected to examine students' comprehension of reported speech before and after the implementation of drama scripts through role-play activities. In this design, students completed a pre-test before receiving

the treatment and a post-test after the treatment. The research design can be illustrated as $O_1 - X - O_2$, where O_1 represents the pre-test, X represents the treatment, and O_2 represents the post-test (Creswell, 2017). The study was conducted at SMK Negeri 2 Langsa during the 2025/2026 academic year. The population consisted of all eleventh-grade students, while the sample comprised 25 students from class XI RPL selected through purposive sampling. This class was chosen because the students had previously studied reported speech but still experienced difficulties in understanding and applying the grammatical structure correctly. Therefore, the class was considered appropriate for examining the implementation of drama scripts through role-play activities.

The research instrument was a grammar test consisting of 25 multiple-choice questions designed to measure students' comprehension of reported speech. The test covered essential aspects of reported speech, including tense changes, pronoun changes, adverbial changes of time and place, and the transformation of statements, questions, and commands. The instrument was adopted from a previously validated study on students' mastery of reported speech conducted by Rahayu (2012). The instrument had fulfilled content validity in the original study and was used without modification in the present research. Therefore, additional statistical item validation was not conducted. Likewise, no additional expert judgment, pilot testing, or reliability analysis, such as Cronbach's Alpha, was performed because the instrument had already been adopted in its original validated form. Each correct answer received one point, while incorrect answers received zero points. The same test was administered as both the pre-test and the post-test to ensure consistency in measuring students' comprehension before and after the treatment.

Data collection was carried out in three stages. The first stage was administering the pre-test to identify students' initial comprehension of reported speech. The second stage involved implementing the treatment using drama scripts through role-play activities, there are four meetings. During the treatment, students worked collaboratively in groups using drama scripts containing dialogues that required them to identify and practice reported speech in meaningful communicative contexts. Students first read and discussed the scripts, identified the reported speech expressions, practiced transforming direct speech into reported speech, rehearsed the dialogues with their group members, and finally performed the drama in front of the class. Throughout these activities, the teacher guided students in recognizing grammatical features such as tense shifts, pronoun changes, and adverbial changes while providing feedback on the accuracy of their language use. The final stage was administering the post-test using the same instrument to measure students' comprehension after completing the treatment. The collected data were analyzed using IBM SPSS Statistics version 25. Descriptive statistics were first used to summarize students' performance by presenting the minimum score, maximum score, mean, and standard deviation of both the pre-test and post-test. Before testing the hypothesis, a Shapiro-Wilk normality test was conducted because the sample consisted of fewer than 50 participants. A significance value greater than 0.05 indicated that the data were normally distributed. Since the data met the assumption of normality, a paired sample *t*-test was employed to compare students' pre-test and post-test scores.

The null hypothesis (H_0) was rejected and the alternative hypothesis (H_1) was accepted if the significance value (Sig. 2-tailed) was less than 0.05.

3. Result

The following table shows the frequency distribution of students' pre-test and post-test scores before and after using drama scripts in role-play activities.

Table 1. Distribution Frequency of Students' Pre-Test Score

Class Interval	Category	Frequency	Percent
86 - 100	Very Good	0	0%
70 - 85	Good	0	0%
56 - 69	Fair	4	16%
36 - 55	Poor	10	40%
0 - 35	Very Poor	11	44%
Σ		25	100%

Table 2. Distribution Frequency of Students' Post-Test Score

Class Interval	Category	Frequency	Percent
86 - 100	Very Good	9	36%
70 - 85	Good	14	56%
56 - 69	Fair	2	8%
36 - 55	Poor	0	0%
0 - 35	Very Poor	0	0%
Σ		25	100%

After the students' pre-test and post-test scores were collected, the scores/data were analyzed using IBM SPSS 25. The following are the steps of the analysis:

3.1 Descriptive Statistics

Descriptive statistics were used to summarize students' performance before and after the treatment. The analysis includes the number of participants, minimum score, maximum score, mean score, and standard deviation.

Table 2. Descriptive Statistics

Descriptive Statistics						
	N	Minimum	Maximum	Sum	Mean	Std. Deviation
Pre-Test	25	8	56	938	37.52	13.800
Post-Test	25	68	100	2056	82.24	8.570
Valid N (listwise)	25					

The pre-test results indicate that students' initial comprehension of reported speech was relatively low, with a mean score of 37.52. After the implementation of drama scripts through role-play activities, the mean score increased to 82.24. The decrease in standard deviation from 13.800 to 8.570 also indicates that students' scores became more consistent after the treatment. These results suggest that students demonstrated a significant improvement in understanding reported speech after participating in the drama-based learning activities.

3.2 Normality Test

Before conducting hypothesis testing, a normality test was performed to determine whether the pre-test and post-test data were normally distributed. Since the number of participants was fewer than 50, the Shapiro–Wilk test was used.

Table3. Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pre-Test	.136	25	.200*	.948	25	.227
Post-Test	.127	25	.200*	.967	25	.565

The results show that the significance value for the pre-test is 0.227 and for the post-test is 0.565, both of which are greater than 0.05. This indicates that the data are normally distributed. Therefore, the assumption for parametric statistical analysis was met, allowing further analysis using a paired sample t-test.

3.3 Paired Sample T-test

A paired sample t-test was conducted to examine whether there was a significant difference between students' pre-test and post-test scores.

Table.4 Paired Test

		Paired Differences							
				Std. Error	95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Pre-Test - Post-Test	-44.720	13.303	2.661	-50.211	-39.229	-16.809	24	.000

The results show that the mean difference between the pre-test and post-test scores is -44.720, indicating that the post-test scores were significantly higher than the pre-test scores. The significance value (0.000) is lower than the significance level of 0.05, which means that the difference between the two scores is statistically significant. These findings indicate that the implementation of drama scripts through role-play activities significantly improved students' comprehension of reported speech.

3.4 Hypothesis Test

Hypothesis testing was conducted to determine whether the alternative hypothesis (H1) was accepted or rejected based on the statistical analysis. The test was carried out using the results of the paired sample t-test in SPSS version 25. The purpose of this test was to examine whether the use of drama scripts through role-play activities had a significant effect on students' comprehension of reported speech.

The hypotheses of this study are formulated as follows:

- a. H0 (Null Hypothesis): There is no significant effect of using drama scripts through role-play activities on students' comprehension of reported speech.
- b. H1 (Alternative Hypothesis): There is a significant effect of using drama scripts through role-play activities on students' comprehension of reported speech.

The decision criteria are:

- a. If the significance value (Sig. 2-tailed) is less than 0.05 ($p < 0.05$), then H0 is rejected and H1 is accepted.
- b. If the significance value (Sig. 2-tailed) is greater than 0.05 ($p > 0.05$), then H0 is accepted and H1 is rejected.

Based on the paired sample t-test results, the significance value (Sig. 2-tailed) is 0.000, which is lower than the significance level of 0.05 ($0.000 < 0.05$). Therefore, the null hypothesis (H0) is rejected and the alternative hypothesis (H1) is accepted. This result indicates that there is a statistically significant difference between students' pre-test and post-test scores. In other words, the implementation of drama scripts through role-play activities significantly improves students' comprehension of reported speech in class XI RPL at SMK Negeri 2 Langsa.

4. Discussion

The findings of this study demonstrate that the implementation of drama scripts through role-play activities improved students' comprehension of reported speech. The increase in students' post-test scores indicates that learning grammar through contextualized and interactive activities enabled students to understand and apply the rules of reported speech more effectively than before the treatment. Rather than memorizing grammatical rules in isolation, students actively practiced transforming direct speech into reported speech while participating in meaningful communication.

This learning experience allowed them to connect grammatical knowledge with authentic language use, making reported speech easier to understand and apply in classroom situations. One possible explanation for these findings is that drama-based learning provides students with repeated opportunities to use grammatical structures in meaningful contexts. Throughout the learning process, students worked collaboratively to read drama scripts, identify reported speech expressions, transform direct speech into reported speech, rehearse dialogues, and perform role-play activities. These activities encouraged students to practice grammatical structures repeatedly while communicating with their peers instead of completing isolated grammar exercises. As explained by (Davies, 1990), drama creates a supportive learning environment in which learners can experiment with language without fear of making mistakes. This learning atmosphere encourages active participation, increases learners' confidence, and helps students develop a deeper understanding of grammar through authentic communication.

The findings are consistent with constructivist learning theory, which emphasizes that knowledge is actively constructed through experience. According to (Piaget, 1970), learners develop understanding by integrating new knowledge with their existing cognitive structures. In the present study, students were not limited to memorizing the rules of reported speech. Instead, they actively constructed their understanding by applying grammatical rules while interacting with drama scripts and participating in role-play activities. As a result, students developed a more meaningful understanding of reported speech because grammar learning was closely connected to communicative situations rather than presented as isolated rules. The present findings can also be explained through Vygotsky's Social Constructivism, which emphasizes that learning occurs through social interaction and collaboration. During group discussions, rehearsals, and classroom performances, students exchanged ideas, solved grammatical problems together, and received guidance from both their peers and the teacher. These collaborative activities created opportunities for students to strengthen their understanding through interaction before applying the target grammatical structure independently (Santhi & Selvam, n.d.). In addition, Long's Interaction Hypothesis suggests that language acquisition develops through meaningful interaction and negotiation of meaning. Throughout the role-play activities, students continuously interpreted dialogues, discussed language forms, and practiced transforming direct speech into reported speech while communicating with others (Zulianingrum, 2025). This interaction enabled students to notice grammatical forms within meaningful communication, supporting both grammatical development and language use simultaneously.

The findings of this study are generally consistent with previous studies on drama-based language learning. (Anderson & Dunn, 2013) argue that drama promotes language development by engaging learners in collaborative experiences that strengthen language comprehension and retention. Similarly, the improvement observed in this study suggests that collaborative drama activities helped students develop a better understanding of reported speech. Likewise, (Ladousse, 1987) explains that role-play encourages learners to express ideas naturally while practicing language structures in communicative situations. Moreover (Livingstone, 1983) also emphasizes that role-play

enables learners to practice grammar in meaningful contexts instead of relying solely on mechanical drills. These similarities indicate that drama-based learning provides opportunities for students to use language meaningfully while reinforcing grammatical knowledge through authentic interaction. However, the present study also extends previous research in several important ways. Earlier studies have predominantly investigated the contribution of drama to improving speaking ability, learner motivation, and communicative competence. In contrast, this study specifically examined the role of drama scripts in supporting students' comprehension of reported speech as a particular grammatical structure. Rather than focusing on oral fluency, the learning activities were designed to help students recognize, transform, and apply reported speech accurately during classroom interaction. This difference may explain why students demonstrated considerable improvement in grammatical comprehension after participating in the role-play activities. The findings therefore suggest that drama-based learning can contribute not only to communicative competence but also to grammar instruction when learning activities are carefully designed to provide contextualized grammatical practice.

Based on these findings, this study proposes a conceptual pedagogical framework for teaching grammar through drama scripts and role-play activities. The learning process begins with contextualized practice, in which students are introduced to reported speech through drama scripts that reflect meaningful communication (Topal, 2024). It continues with collaborative interaction, where students discuss the scripts, identify grammatical forms, rehearse dialogues, and receive feedback from their peers and teacher. These activities promote learner engagement by encouraging active participation, confidence, and meaningful language use throughout the learning process. As students repeatedly apply reported speech in authentic communicative situations, they gradually develop stronger grammatical understanding, enabling them to comprehend and use reported speech more accurately. This framework explains how drama-based learning supports grammar instruction while providing practical guidance for English teachers who wish to integrate drama scripts and role-play activities into grammar lessons. Overall, the findings suggest that drama scripts implemented through role-play activities provide an alternative approach to grammar instruction by integrating grammatical knowledge with meaningful communication. Instead of viewing grammar as a collection of rules to be memorized, students learn to understand and apply grammatical structures through interaction, collaboration, and authentic language use. Therefore, this study contributes to the growing body of literature on drama-based learning by demonstrating its potential not only to support communicative competence but also to facilitate students' comprehension of specific grammatical structures, particularly reported speech.

5. Conclusion

This study concludes that the implementation of drama scripts through role-play activities contributed to improving students' comprehension of reported speech among eleventh-grade students at SMK Negeri 2 Langsa. The findings demonstrated a significant improvement in students' post-test scores compared with their pre-test performance, indicating that integrating grammar instruction with contextualized and interactive

learning activities enabled students to understand and apply reported speech more effectively. By engaging students in meaningful communication through drama, grammar learning became more active, collaborative, and relevant to authentic language use. These findings also contribute to English language teaching by demonstrating that drama scripts can function not only as speaking activities but also as instructional tools that support students' comprehension of grammatical structures through contextualized practice and collaborative learning. The findings offer several practical implications for English language teaching. English teachers are encouraged to integrate drama scripts into grammar instruction by selecting scripts that contain the target grammatical structure, organizing students into small groups, guiding them to identify and transform direct speech into reported speech, allowing them to rehearse and perform the dialogues through role-play activities, and providing corrective feedback after each performance. These learning activities enable students to practice grammar in meaningful communicative contexts rather than relying primarily on rule memorization and mechanical exercises. Consequently, students have greater opportunities to understand grammatical concepts while simultaneously developing their confidence and participation in classroom interaction.

The results also have implications for curriculum development and educational institutions. Curriculum designers are encouraged to incorporate drama-based learning activities into English language programs, particularly as complementary strategies for teaching grammar through meaningful communication. Drama scripts may be integrated into classroom modules, supplementary learning materials, or project-based learning activities that encourage students to apply grammatical structures in authentic situations. To support the successful implementation of this approach, schools should provide appropriate instructional resources, allocate sufficient classroom time for collaborative learning activities, and facilitate professional development programs that equip English teachers with practical strategies for integrating drama into grammar instruction. Despite these positive findings, this study has several limitations. The research employed a one-group pre-test and post-test quasi-experimental design without a control group, making it difficult to attribute the observed improvement solely to the treatment. In addition, the study involved a relatively small number of participants from a single school, which may limit the generalizability of the findings to other educational settings. Future research is recommended to employ more rigorous research designs by involving larger and more diverse samples as well as comparison or control groups to strengthen the evidence regarding the implementation of drama scripts in grammar instruction. In addition, future studies are encouraged to adopt mixed-methods approaches by combining quantitative findings with qualitative data obtained through classroom observations, interviews, or students' reflective responses to gain a more comprehensive understanding of students' learning experiences. Longitudinal studies are also recommended to investigate the long-term retention of reported speech after drama-based instruction and to examine whether this approach remains effective over time. Furthermore, future research may explore the implementation of drama scripts in teaching other grammatical structures and across different educational contexts to

broaden the applicability of drama-based grammar instruction in English language learning.

References

- Altun, M. (2019). Drama: a neglected source in language teaching to improve communication. *International Journal of English Linguistics*.
https://www.researchgate.net/profile/Mustafa-Altun/publication/335443086_Drama_A_Neglected_Source_in_Language_Teaching_to_Improve_Communication/links/5d662dd5299bf1f70b124f8d/Drama-A-Neglected-Source-in-Language-Teaching-to-Improve-Communication.pdf
- Anderson, M., & Dunn, J. (2013). *How drama activates learning: Contemporary research and practice*. books.google.com.
https://books.google.com/books?hl=en%5C&lr=%5C&id=wQFMAQAAQBAJ%5C&oi=fnd%5C&pg=PR3%5C&dq=anderson+m+dunn+j+2022++how+drama+activates+learning+contemporary+research+and+practice%5C&ots=nZOGdt_RYJ%5C&sig=h79pEPAVmkEI00OarxyfP1jTOY
- Davies, P. (1990). The use of drama in English language teaching. *TESL Canada Journal*.
<https://teslcanadajournal.ca/index.php/tesl/article/view/581>
- Kolar, M., & Sándorová, Z. (2025). Experiment as a research method for investigating the effectiveness of drama in EFL Higher Education: A systematic review. *Glottodidactica*. https://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.ojs-doi-10_14746_gl_2025_52_2_7
- Ladousse, G. P. (1987). *Role play*. books.google.com.
https://books.google.com/books?hl=en%5C&lr=%5C&id=oS4STyscmpYC%5C&oi=fnd%5C&pg=IA4%5C&dq=role+play%5C&ots=bmNnbnUQH%5C&sig=q_tF8346ie2R0pXL6sC8GMZrSi4
- Livingstone, C. (1983). *Role Play in Language Learning*. ERIC.
<https://eric.ed.gov/?id=ED240866>
- Miranti, C. D. (2020). *Students' Difficulties In Learning Reported Speech (A Study at the Third Semester Students of English Department)*. repository.ar-raniry.ac.id.
<https://repository.ar-raniry.ac.id/id/eprint/11056/>
- Normawati, A., & Nugrahaeni, D. A. (2024). Grammar teaching and learning in English language class: Students' view. *English Education: Journal of*
<https://ojs.unpkediri.ac.id/index.php/inggris/article/view/21536>
- Nst, S. A., Pasaribu, N. S., & ... (2025). The Use of Gerunds and Reported Speech in Academic and Informal Texts. *PEMA: Jurnal* <http://jurnal.permapendis-sumut.org/index.php/pema/article/view/1386>
- Piaget, J. (1970). *Science of education and the psychology of the child*. Trans. D. Colman.

psycnet.apa.org. <https://psycnet.apa.org/record/1970-19308-000>

Rizki, H. (2019). *Difficulties In Learning Reported Speech In Form Of Question Of The Second Graders Of English Language Education Of* Eprints.Hamzanwadi.Ac.Id. <https://Eprints.Hamzanwadi.Ac.Id/3081/>

Santhi, V. J., & Selvam, R. (n.d.). Teaching Grammatical Structures Using Interaction. *Researchgate.Net*. https://Www.Researchgate.Net/Profile/Jeya-Santhiv/Publication/343675009_Teaching_Grammatical_Structures_Using_Interaction/Links/5f37de69a6fdcccc43ca3427/Teaching-Grammatical-Structures-Using-Interaction.Pdf

Topal, I. H. (2024). Communicative Activities for Written and Spoken Production of Narrative Tenses in English. *EIKI Journal of Effective Teaching Methods*. https://www.researchgate.net/profile/Ibrahim-Topal-3/publication/386581827_Communicative_Activities_for_Written_and_Spoken_Production_of_Narrative_Tenses_in_English/links/6757e07cdf681a5748f8fc0a/Communicative-Activities-for-Written-and-Spoken-Production-of-Narrative-Tenses-in-English.pdf

Wahyudi, E. (2021). *Grammatical Error Analysis in Students' Reported Speech (The Case of the Eleventh Grade Students of SMA Negeri Gondangrejo Karanganyar in 2010/2011* digilib.uns.ac.id. <https://digilib.uns.ac.id/dokumen/detail/27014>

Zulianingrum, A. (2025). Empowering EFL Learners' Communicative Competence through Role-Play in Task-Based Language Teaching (TBLT). *Mozaic: Islam Nusantara*. <https://journal.unusia.ac.id/mozaic/article/view/1495>

The Effect of Using Song Vocabulary to Students' Retention

Rifka Nurin Nafisah ¹⁾, Albadri ²⁾, Musdalifah ³⁾

^{1,2,3}English Department, Social and Humanity Science Faculty, University of Ibrahim

Email Correspondence : lifamudalifa.sit@gmail.com.

ARTICLE INFO

Keywords:

*Song Vocabulary
Students' Retention*

Article History:

Received : 08-05-2026
Revised : 08-08-2026
Accepted : 10-08-2026
Published: 20-08-2026

ABSTRACT

This study determines the effect of using song vocabulary to students' retention. This research was conducted at the grade Intermediate in Syu'batul Lughah Al-Khairiyah, Ma'hadul Qur'an, Sukorejo, Situbondo. This research used a quantitative method by using Pre-Experimental research type one group pretest posttest design. This research used saturated sampling technique, which a whole of population which consists 19 students are sample. The instrument data collection used test. The questionnaire also took in this research to approve the students' retention of vocabulary through a song. Data analysis technique using analysis of t-test by paired sample t-test. It was gotten $t_0 = (-19.396) > t_t = (2,861)$ with sig. (2-tailed) value was 0,000. So, it meant that H_0 is refused and H_a is accepted.

1. Introduction

English as an international language used as tool communication in many countries. In Indonesia, English is a foreign language, people use English for tourism, education and international affairs. Not only in Indonesia but also in some countries, English is a foreign language. It is important for us to learn English, because we need it to communicate each other without misunderstanding.

According to (Gushendra, 2017) the first step to study language especially English skill is mastering some vocabularies. Vocabulary is one of the important factors in all language teaching. Student must continually be learning words as they learn structure of vocabulary. Vocabulary is the first key to communicate each other. In the communication we need some vocabularies. Because when someone try to speak English, he must master some vocabularies as much as he can.

Vocabulary can be interpreted as all the words a person's understands that are likely to be used to make a new sentence. The richness of a person-owned vocabulary can be considered to indicate the level of intelligence or education. Vocabulary is a knowledge about the meaning of words and it is a central part of language. This means that

vocabulary is an important part of someone's knowledge which is used to communicate in their daily life (Kansızoğlu & Bekiroğlu, 2025).

Vocabulary is the main component of language. People use vocabulary to construct a sentence to express the idea. By having a lot of vocabulary, people can express more idea. In second language acquisition, vocabulary achievement is an important element. Based on (Sánchez-Gutiérrez et al., 2025), "Without grammar little things can be conveyed, without vocabularies, nothing can be conveyed". The mastery of vocabulary is essential to construct sentences that contain some messages for social interaction.

One of the things that taught for each language learner in the classroom is vocabulary. Vocabulary also has a significant influence on language comprehension. Where, vocabulary can enrich the language skills of students, such as listening, reading, speaking and writing. Student may be more fluent in doing those skills. Nevertheless, there has been some students who have difficulty learning it. Students can't get the sense of words, they can't get the text details and they can't get the information from other people. Those problems happen because the students are not given a chance to improve their knowledge about the new words. Therefore, teacher uses inappropriate technique in teaching.

According to (Yaşar & Koçoğlu, 2025) state that retention in other word is memory. Retention is ability to recognize or recall what has been experienced or studied. It also uses to call as self-regulation especially on vocabulary learning. The model of human memory is the process of information that involves encodes, stores, and retrieves information.

Human retention cannot intercept all what they have memorized in a long time. Therefore, it is important to make a new innovation to make students easier to memorize in a longer time by using song vocabulary and easier to memorize. Because, in the fact, human retention will be easier to be developed with some stimulations which has been given.

Many methods which are able to be used by educator in the teaching process, but to ward saturation off and students boredom, singing is the exact solution. Because, generally someone who feels bored, he will sing a song or listen to the music. Music is one of things can sharpen brain until brain will do accordance with the function after refreshing. By singing music can be happier atmosphere. Until realize happy feeling and new spirit, then makes students faster in the memorizing some vocabularies (Muhammad, 2012).

Song vocabulary is a technique which is used for making students be easier in the memorizing vocabulary. They can choose what song they want to be used, then they change the lyric with their vocabularies that they have. It is the different way that apply as a song than the other research. By song vocabulary, they are able to repeat it every time easily.

There are some subjects in the English class that they have to join. Such as: grammar, speaking, pronunciation and vocabulary class. For grammar, speaking and pronunciation class are hold at the evening. For vocabulary class is hold at the morning.

In the vocabulary class, they may get about five vocabularies in every meeting. And in the week end they have to memorize some vocabularies that they got to their tutor. Not

only memorize but also they have to make some sentences from those vocabularies to improve their writing skill well.

In the fact, some students have difficulties in the memorizing vocabularies. The quantity of their memorizing, it makes them difficult to memorize some vocabularies. Their difficulties which is faced by them, they need new innovation/ learning method to makes them easier to memorize it in the long term memory.

This research is limited to the teaching vocabulary of intermediate grade in Syu'batul Lughah Al-Khairiyah, Ma'hadul Qur'an, Sukorejo, Situbondo. Based on the problem above, researcher decides to conduct an action research to overcome the problem of the students in the learning English, especially in memorizing some vocabularies that they have. The researcher wants to apply song vocabulary in the memorizing vocabulary to improve students' retention.

So, the writer wants to conduct a research about the students' retention of vocabulary by using song vocabulary with the entitled. "The effect of Using Song Vocabulary to Students' Retention in Syu'batul Lughah Al-Khairiyah, Ma'hadul Dormitory Female, Sukorejo, Situbondo". While the hypothesis of the research H_0 there is no a significant difference between students' retention in memorizing vocabulary by using song. On the other hand H_a there is a significant difference between students' retention in memorizing vocabulary by using song.

2. Methods

This research was held at the grade Intermediate in Syu'batul Lughah Al-Khairiyah, Ma'hadul Qur'an, Sukorejo, Situbondo. The population of this research is 19 student of intermediate class. The research used saturated sampling which a whole of the population is sample. According to (Sugiono, 2016), saturated sampling is technique of sample determining, if of population is used as sample, it is often used if it is relative small, less than 30 people or the researcher wants to make generalization with smallest error. In other name saturated sample is census where all of population is sample.

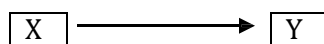
Researcher used a quantitative method. It is designed by using Pre-Experimental research type One Group Pretest Posttest design. The research instrument used questionnaire and test. The test consisted pretest and posttest. The question of the test in pre-test and post-test are 40 question. 20 questions in pre-test and 20 questions in post-test. Therefore a questioner indicates of some aspects in retention. A validity of the test and questionnaire were applied by content validity. It defines content validity as the extent to which the test is representative of a defined body of content consisting of topics, material and processes (Sulistyo, 2018). The material of the topic on the research through vocabulary are about noun, verb, adjective, synonym and antonym. While the questionnaire consist of some aspects of retention. Those are about the students' focus on the vocabulary and meaning, quickly in remembering vocabulary, easier in getting a meaning of vocabulary, and getting a long retention on their students' main on the vocabulary through a song. Data analysis techniques using analysis of normality, homogeneity and t-test by paired sample t-test (Gunawan, 2015).

Additionally, (Setyosari, 2015) stated in this One Group Pretest-Posttest Design, the researcher gives pretest in the beginning before giving treatment. This research includes 3 steps, there are: giving pretest to measure dependent variable, giving treatment or experiment in four meetings and giving posttest to measure the result or effect of the treatment to dependent variable. So, the effect of the treatment is determined by comparing the pretest and posttest value. Therefore a formula of the pre experiment research in the research can be shown on the table below.

Table 1. One Group Pretest-Posttest Design Information:

	O1	X	O2
O1	: Pretest		
X	: Treatment		
O2	: Posttest		

Table 2. Theoretical Framework



Information:	X	: Song Vocabulary
	Y	: Students' Retention

From the theoretical framework, the researcher estimated that song vocabulary had an effect on students' retention in the memorizing vocabulary.

The researcher used test in a multiple choice test related with noun, verb and adjectives through the song vocabulary. There are twenty question on pretest and twenty questions on posttest. While a questioner was using a Likert Scale. Those components are completely disagree (score 1), Disagree (score 2), Neutral (score 3), Agree (score 4), completely agree (score 5). It was conducted in the grade of Intermediate in Syu'batul Lughah Al-Khairiyah, Ma'hadul Qur'an, Sukorejo, Situbondo before and after treatment.

In this research the normality test was used to analyze data that were normally distributed or not. This test used SPSS statistic program v. 26 with the *Shapiro-Wilk formulation* because the sample used was less than 50 respondents (Susetyo, 2014). If the value of sig. (2-tailed) is smaller than the alpha level 5% ($\alpha = 0,05$). So, the data was normally distributed. The homogeneity test used to determine whether the data obtained had a homogeneous variance or not (Susetyo, 2014)

To find out the influence of media for the sample which can be used to population, so it will give influence or not. the data analysis technique uses SPSS v.26 which is paired samples t-test (Gunawan, 2015).

3. Result

The. The findings of this research based on the data description, data analysis and test of hypothesis. Based on the research result that has been done in Syu'batul Lughah Al-Khairiyah, Ma'hadul Qur'an Dormitory Female, Sukorejo, Situbondo, so, it's been gotten

the result of value distribution before giving treatment as pre-test and after giving treatment as post-test to 19 respondents as follows:

Table 3. Value of Pretest and Posttest

Value (pre-test)	Frequency	Value (post- test)	Frequency
45	1	45	-
50	2	50	-
55	2	55	-
60	5	60	-
65	3	65	-
70	5	70	4
75	1	75	1
80	-	80	6
85	-	85	3
90	-	90	4
95	-	95	1
Total	19	Total	19

In this research the normality test was used to analyze data that were normally distributed or not. After the normality test was carried out, a homogeneity test was also carried out to determine whether the data obtained had a homogeneous variance or not. The normality test in this research are using the kolmogorov- smirnov test. The result can be seen in the following table:

Table 4. Normality test

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
Pretest	,156	19	,200*	,941	19	,275
Posttest	,170	19	,152	,914	19	,086

Based on Shapiro-Wilk test, the normality test outcome above shows that the significance level was 0,086. It means that the probability value (p) of group was above (>) the significance level of 5% ($\alpha = 0,05$). So, data above were normally distributed.

Based on normality calculation, the researcher got the result that all data in pre-test and post-test have normally been distributed. The next step of the calculation was to find out the homogeneity test of pre-test and post-test of the data by using SPSS v.26. The result of the test of homogeneity of variances are shown as follows:

Table 5. Homogeneity Test

Test of Homogeneity of Variances					
		Levene Statistic	df1	df2	Sig.
Song vocabulary	Based on Mean	,083	1	36	,775
	Based on Median	,107	1	36	,745
	Based on median and with adjusted df	,107	1	35,974	,745

Based on trimmed mean	,096	1	36	,759
-----------------------	------	---	----	------

Based on the test of Homogeneity of Variances above, it got the calculation of the test value based on the mean obtained a significant value of 0,775. And the significant level = 5% / 0, 05. Because it was higher than 0, 05. So, it shows that all of the tests are homogeneous.

To find out the influence of media for the sample which can be used to population, so it will give influence or not. The data analysis technique uses SPSS v.26 which is paired samples t-test:

Table 6. Paired Sample Test

Paired Samples Test									
Paired Differences									
		Mean	Std. Deviation	Std. error mean	95% confidence interval of the difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Pretest-posttest	-19,47368	4,37631	1,00399	-21,58300	-17,36437	-19,396	18	,000

From the table above, we know that $t_{count} = -19,369$ with Sig. (2-tailed) is 0,000. Meanwhile the $t_{table} = 2,861$ with sig (2-tailed) value was 0,000. It aimed that $t_{count} \geq t_{table}$. So, it can be concluded that H_0 is refused and H_a is accepted.

While the questionnaire of the students' through song in learning vocabularies on the retention of the students also applied on the research to approved the students' retention on the research. The result of the students' questionnaire of the research can be seen on the table below.

Table 7. Students' score on questionnaire

Student	y1	y2	y3	y4	total
1	4	3	3	2	12
2	4	4	4	3	15
3	4	4	3	3	14
4	4	4	4	4	16
5	4	3	3	2	12
6	4	4	4	3	15
7	3	3	2	2	10
8	5	3	3	5	16
9	2	2	2	2	8

10	4	4	4	2	14
11	4	5	5	5	19
12	4	4	4	4	16
13	4	4	5	4	17
14	3	3	4	3	13
15	5	4	4	4	17
16	4	4	4	4	16
17	4	3	4	4	15
18	5	4	4	4	17
19	4	4	3	3	14

The data of questionnaire consist of some aspects of retention. Those are about y1 is the students' focus on the vocabulary and meaning, y2 is quickly in remembering vocabulary, y3 is easier in getting a meaning of vocabulary, and y4 is getting a long retention on their students' main on the vocabulary through a song.

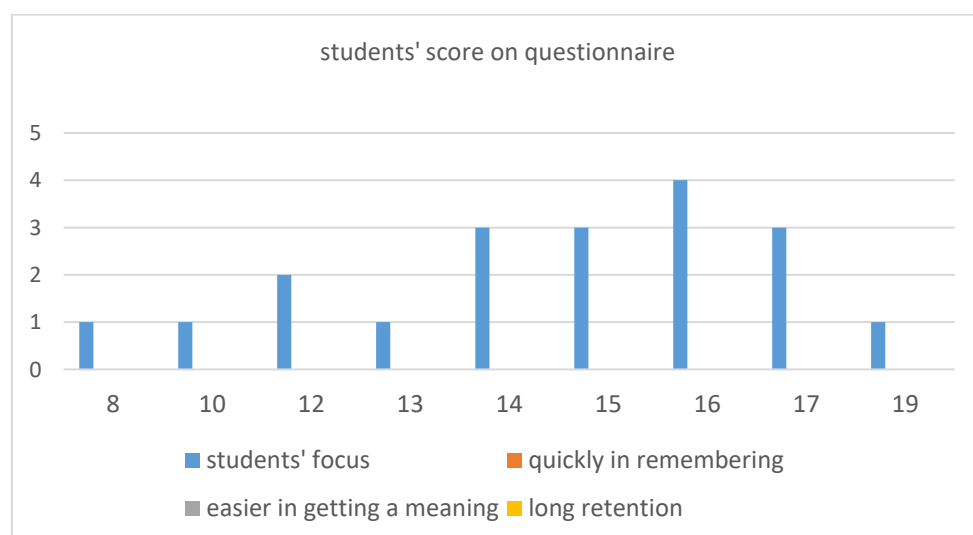


Fig. 1. Students' score on questionnaire

Based on the distribution diagram of the total scores for students' retention on learning vocabulary through a song, the majority of respondents scored within the range of 14–17. This distribution indicates that the vast majority of respondents expressed a high level of

agreement with the indicators of students' retention. There were only two respondents with low scores (8 and 10), while the highest score (19) was achieved by a single respondent. Thus, in general, respondents' perceptions of students' retention can be categorized as positive, as the majority of scores fell within the 'Agree' and 'Strongly Agree' categories.

4. Discussion

Based on the results of the paired samples test, a mean difference on the research was 19.47 it was indicating an increase in students' retention on vocabulary through song. It suggests that the intervention or learning model that was implemented by a song in vocabularies are able to significantly effect in students' retention. Similar findings were reported by (Zaharani, 2023), who stated that learning vocabulary through song input positively affect vocabulary learning and retention. Statistically, this effect shows that the post-test score was better than the pre-test score. These results are consistent with various studies which state that the pre-test-post-test design is an effective for evaluating the success of a learning vocabularies, as it is able to measure changes in participants' score before and after the treatment.

The results of the paired-sample t-test showed a t-value of 19.39 with $df = 18$ and a significance value two-tailed of 0.000 ($p < 0.05$). Consequently, the null hypothesis (H_0) is rejected and the alternative hypothesis (H_a) is accepted. It can be concluded that there is a significant difference between the pre-test and post-test scores. The very high t-value indicates that the research has positive effect in applying a song on vocabulary through students' retention. This finding is supported by research showing that a p-value of < 0.001 in a paired t-test provides strong evidence that a learning on vocabularies through song has a significant effect on in students' retention. It was supported in the (Melinda et al., 2025) indicated that in their research through students' vocabulary retention through song.

Furthermore, the 95% confidence interval ranges from 17.36 to 21.58. It means increase in scores due to the treatment of using a song in students' vocabulary retention. Similarly, (Chang, 2024) reported that using a song has an effect on students' vocabulary retention through applying three mode those are listening+singing+reading lyric. On the other hand, the standard deviation of 4.37 indicates that the variation in score increases amongst respondents was relatively small, suggesting that the effect in learning vocabulary on students' retention through song fairly consistently. It also in line (Nie et al., 2022) that vocabulary retention has a positive effect through listening a song.

The improvement in learning on students' retention observed in this research can also be explained by the characteristics of songs as a learning medium capable of creating a more engaging and meaningful learning environment. Songs provide a natural context for the use of vocabulary, so that students are not only memories the meanings of words but also understand how to use them in sentences. This process facilitates deeper, meaningful learning, which ultimately enhances students' retention of the material studied. In addition to strengthening cognitive aspects, the use of songs also enhances students' retention, attention and engagement during the learning process. These findings support

the results of a systematic review by (Yakoub-Agha et al., 2026) using a song in learning vocabulary has a positive impact on the students' visual memory.

Additionally, (Maulana & Putraidi, 2024) which concluded that the use of songs. It was consistently has a positive impact on vocabulary acquisition, memory retention and learning motivation. The findings of (Melinda et al., 2025) also show that learning using adapted songs can significantly improve vocabulary retention while creating a more interactive classroom atmosphere.

5. Conclusion

Provide Based on the results of the Paired Samples Test in the research, a significance value (Sig. 2-tailed) of 0.000 was obtained, which is less than 0.05. It can be concluded that there is a highly significant effect or difference in means between on the pre-test and post-test scores following the retention on students' learning vocabulary through a song. Although the intervention showed clear positive results with a mean difference of -19.47368, this study still has several limitations, including a very small sample size of only 19 respondents (as indicated by $df = 18$) and a fairly wide confidence interval, which indicates high inter-individual variation in the data. Therefore, for future research, it is recommended to increase the sample size to ensure more accurate generalization of the results, to use an experimental design involving a control group for comparison, and to include a follow-up test on post-test to measure the long-term consistency and effectiveness of the intervention administered.

References

- Chang, A. C. S. (2024). The Effects of Varying Practice Modes on L2 Learners' Vocabulary Retention Through Songs: Listening, singing, and oral reading. *Tesl-Ej*, 28(3), 1–18. <https://doi.org/10.55593/ej.28111a9>
- Gunawan, M. A. (2015). *Statistik Penelitian Bidang Pendidikan, Psikologi dan Sosial*. Yogyakarta: Parama Publishing.
- Gushendra, R. (2017). An experimental Study: Improving Students' Vocabulary mastery by using English song. *IJIELT (Indonesian Journal Of Integreted English Language Teaching)*, 3(1). <https://doi.org/https://doi.org/10.24014/ijielt.v3i1.3968>
- Kansızoğlu, N., & Bekiroğlu, N. (2025). The effect of vocabulary development interventions on cognitive outcomes of vocabulary: A meta-analysis study. *Journal of Pedagogical Research*, 9(1), 123–148. <https://doi.org/10.33902/JPR.202532083>
- Maulana, E. S., & Putraidi, K. (2024). *The Effect of Using Modern English Song Lyrics to The Vocabulary Mastery of Grade VIII*. 7, 5–9. <https://doi.org/https://doi.org/10.37824/0cfjgh69>
- Melinda, Sunra, L., & Farahdiba, S. (2025). Integrating modified english songs to improve vocabulary retention: a pre-experimental study with grade-7 EFL learners in indonesia. *Celebes Journal of Language Studies*, 5(2), 2776–7493. <https://doi.org/https://doi.org/10.51629/cjls.v5i2.265>
- Muhammad, F. (2012). *Desain Pembelajaran PAUD*. Ar-Ruzz Media.

- Nie, K., Fu, J., Rehman, H., & Zaigham, G. H. K. (2022). An Empirical Study of the Effects of Incidental Vocabulary Learning Through Listening to Songs. *Frontiers in Psychology*, 13(May), 1–12. <https://doi.org/10.3389/fpsyg.2022.891146>
- Sánchez-Gutiérrez, C. H., Robles-García, P., & Serrano, M. P. (2025). L2 Spanish vocabulary teaching in US universities: Instructors' beliefs and reported practices. *Language Teaching Research*, 29(2), 677–699. <https://doi.org/10.1177/13621688221074443>
- Setyosari, P. (2015). *Metode Penelitian Pendidikan & Pengembangan*. Jakarta: PRENAMEDIA GROUP.
- Sugiono. (2016). *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D)*. Bandung: Alfabeta.
- Sulistyo, G. H. (2018). *EFL Learning ASSESSMENT AT SCHOOLS An Introduction to Its Basic Concepts and Principles*. CV. Bintang Sejahtera.
- Susetyo, B. (2014). *Statistika Untuk Analisis Data Penelitian*. Bandung: PT Refika Aditama, 2014.
- Yakoub-Agha, L., Soqia, J., Nahas, L., Alameer, M. B., Mohamad, L., Yakoub-Agha, M., Almohsen, C., & Mohsen, S. (2026). Effects of five-minutes music intervention on college students' memory. *Scientific Reports*, 16(1), 1–13. <https://doi.org/10.1038/s41598-025-32240-2>
- Yaşar, S., & Koçoğlu, Z. B. (2025). Enhancing vocabulary mastery: The impact of learner-created digital flashcards on L2 vocabulary learning and self-regulation. *Australasian Journal of Educational Technology*, 41(4), 17–31. <https://doi.org/10.14742/ajet.10402>
- Zaharani, H. (2023). The Effect of Using Songs to Elevate Students' Vocabulary Skill at SMP Yabes School. *Journal of Classroom Action Research*, 2(2), 52–65. <https://doi.org/10.52622/jcar.v2i2.164>

The Effect of Project-Based Method on Improving the English-Speaking Skills

Akhmad Nuris Hidayat ¹⁾ Dwi Nur Hardiansyah²⁾

¹⁾English Department, STIT Miftahul Midad Lumajang, ²⁾ English Department, Ibrahimy University Situbondo

Email Correspondence: akhmadnuris1993@gmail.com

ARTICLE INFO

Keywords:

Project-Based Method
Speaking Skill
English Language

Article History:

Received : 15 - 03 - 2026
Revised : 24 - 06 - 2026
Accepted : 08 - 08 - 2026
Published: 20 - 15 - 2026

ABSTRACT

In this era, Madrasah Aliyah students must have skills to compete with others. One of them is English speaking skill. Madrasah Aliyah students must have the English-speaking skills necessary to face challenges worldwide. However, due to the lack of learning strategies that support activeness and communication skills, these skills are still underdeveloped. This study aims to evaluate the effect of implementing a project-based method on the English-speaking skills of Madrasah Aliyah students in Lumajang Regency. A quantitative approach with a quasi-experimental design was used in this study with a pretest-post-test and control group design. An English-speaking ability test was conducted to collect data, and the pretest and post-test results of the experimental and control groups were the data sources. English speaking ability between the two groups differed on average, and a t-test at a significance level of 0.05 was conducted for data analysis. Madrasah Aliyah students throughout Lumajang Regency were the subjects of this study, who were selected purposively based on the characteristics of the study. In this study, the tentative assumption is that, compared to conventional learning approaches, the project-based method has a significant effect on students' English-speaking ability. The findings are used for students, teachers, researchers and others to use and improve the method in the process of teaching and learning, & for further research.

1. Introduction

In many Madrasah Aliyah in Lumajang Regency, English speaking ability remains a major challenge for students despite years of formal instruction. Classroom observations and preliminary data indicated that many students were able to understand written texts and complete grammar exercises, yet they struggled to express ideas orally with fluency and confidence. Students often hesitated when asked to speak in front of the class, relied heavily on memorized sentences, and demonstrated limited vocabulary during spontaneous communication. The learning process was generally dominated by teacher explanation and textbook-based activities, providing minimal opportunities for authentic interaction. As a result, students' speaking performance tended to remain at a moderate level with slow progress over time. This phenomenon highlights the need for a more interactive and student-centered instructional approach that can create meaningful speaking opportunities and foster communicative competence in EFL classrooms. To improve the quality of flexible human resources worldwide, Madrasah Aliyah (MA) students must master English speaking skills. Speaking skills require not only grammatical skills, but also courage, fluency, and the ability to communicate correctly in English. However, facts on the ground show that MA students have low English-speaking skills, including at Madrasah Aliyah in Lumajang Regency. Empirical experience gathered by researchers from the English Coaching Clinic (ECC) activity held in November 2025 and involving 102 Madrasah Aliyah students in Lumajang Regency reinforces this phenomenon. The collected data shows that only about 18% of students have English speaking skills in the good category (fluent and confident), 27% have sufficient skills, and 55% are still in the low category. Most students experience difficulties when expressing opinions orally in English; they experience obstacles in speaking fluency, limited vocabulary, incorrect pronunciation, and lack of confidence. This quantitative data shows that the majority of MA students throughout Lumajang Regency have not yet achieved the expected English-speaking skills. In fact, there is a flagship program from the Ministry of Religious Affairs of Lumajang Regency, namely GEMA *Berbaris* (Arabic and English Madrasah Movement), which requires students to be able to speak English well. The low achievement figures indicate a fundamental problem in the English language learning process at Madrasah Aliyah. The conventional teacher-centred approach still dominates learning, with the main focus on grammar mastery and written exercises. Students still have little opportunity to practice speaking actively and contextually. As a result, their oral communication skills are not well developed.

This condition triggers research questions about better learning strategies that might improve MA students' English-speaking skills. Project-Based Learning, or PBL, is considered an alternative solution to this problem. PBL theoretically emphasizes that students must be actively involved in the learning process

through important activities in planning, implementing, and presenting projects (Thomas, 2000; Bell, 2010). With this method, students can actually speak, work in groups, and present in English. Previous studies have shown that project-based learning improves students' speaking ability, motivation, and confidence in English learning (Fragoulis, 2009; Guo *et al.*, 2020). However, there is limited empirical research specifically examining the effect of the Project-Based Method on the English-speaking ability of Madrasah Aliyah students, particularly in Lumajang Regency. Therefore, this study aims to systematically investigate the effect of the Project-Based Method on the English-speaking ability of Madrasah Aliyah students in Lumajang Regency using a quasi-experimental approach. Clear benefits and advantages are expected from this research. Theoretically, this study has the potential to enhance research on English language learning, particularly on the effectiveness of the project-based method in improving speaking skills. Practically, this study is expected to serve as a reference for Madrasah Aliyah English teachers to implement a more creative and communicative approach to teaching English to students. Furthermore, the findings of this study will serve as a basis for policy-making related to the GEMA Berbaris program of the Ministry of Religious Affairs of Lumajang Regency.

Based on the previous information, this study aims to determine: 1) Is the English-speaking ability of MA students in Lumajang Regency different from students taught through a project-based approach compared to students taught through conventional methods? 2) Is the English-speaking ability of MA students in Lumajang Regency significantly improved by project-based learning? The purpose of this study is to 1) find out how the English-speaking ability of MA students in Lumajang Regency is different from students taught through a project-based approach compared to students taught through conventional methods; 2) To determine how much influence project-based learning has on the English-speaking ability of MA students in Lumajang Regency. It is hoped that this study will have theoretical and practical impacts on English learning in Madrasah Aliyah. Theoretically, the results of this study will help increase knowledge about English language education, particularly regarding the effectiveness of project-based learning in improving students' speaking skills. Practically, this study is expected to provide teachers with examples of how to use innovative and informed learning strategies. Furthermore, the results can be used as a basis for developing policies and learning models that support 21st-century skills for madrasahs. It is hoped that this study will serve as a reference for further research involving various contexts, variables, and methods.

A literature review, also called a literature review, is a comprehensive description of previous findings related to the current study. The purpose of this section is to describe the research's position, identify similarities and differences with previous research, and identify research gaps (gap analysis) that can be used

as the basis for innovative arguments. According to research conducted by Thomas in 2019, this approach significantly improves students' English-speaking skills compared to conventional learning approaches. During project activities, such as group work, presentations, and discussions, students become more active and confident in using English. In their study of eleventh-grade high school students, Lestari and Suryani (2022) produced similar results. Their study showed that project-based learning improved students' speaking skills and increased their participation and verbal interaction during learning. Nurhayati and Azmi (2021) investigated the use of project-based learning in English language education at Islamic Senior High Schools (Madrasah Aliyah). This study found that project-based learning can improve students' confidence and fluency in speaking English, but there are several obstacles, such as limited time and learning resources. Furthermore, Putra (2020) used a quasi-experimental approach to show that project-based learning benefits high school students' oral communication skills, helping them speak fluently and convey ideas systematically. Previous studies have shown similarities: a focus on project-based learning and measuring students' English-speaking skills. However, there are fundamental differences, particularly in terms of the context of the educational institution, the scope of the study, and the number of research subjects. Previous studies were mostly conducted in public schools (SMA) or in specific educational units with limited scope. Consequently, these findings do not fully reflect the learning conditions at Islamic Senior High Schools (Madrasah Aliyah), especially on a regional scale. Quantitative research, according to a gap analysis, specifically investigating how project-based methods affect the English-speaking skills of Madrasah Aliyah students within a single district is still limited. Furthermore, few studies have investigated the effectiveness of project-based learning when considering the unique characteristics of madrasahs, including the curriculum, academic culture, and student backgrounds. Consequently, this study is novel in its focus, the context of the research location, and its aim to increase knowledge about the effectiveness of project-based methods in English language learning at Madrasah Aliyah in Lumajang Regency.

Project-Based Learning (PjBL) is a learning approach that emphasizes active student involvement in the learning process through the completion of authentic, contextual, and product-oriented projects. According to Krajcik and Blumenfeld (2020), PjBL is designed to help students build deep conceptual understanding through the exploration of real-life problems that require critical thinking, collaboration, and communication. In the context of 21st-century education, PjBL is seen as a relevant learning model because it simultaneously integrates the 4C skills (critical thinking, creativity, collaboration, and communication). Theoretically, PjBL is firmly rooted in social constructivism theory, which asserts that learning is an active process built through social interaction and direct experience. Vygotsky emphasized the importance of interaction and scaffolding,

which in PjBL are realized through group work, project discussions, and teacher guidance. Recent empirical studies have shown that PjBL can increase student engagement and intrinsic motivation because they feel responsible for the learning process and outcomes (Guo et al., 2020). This situation makes Project-Based Learning (PjBL) an effective approach to language learning, particularly in developing productive skills such as speaking. In English learning, Project-Based Learning (PjBL) provides space for students to use the target language actively and meaningfully. Kokotsaki, Menzies, and Wiggins (2021) state that Project-Based Learning (PjBL) is highly effective in language learning because it encourages language use in real-life contexts, rather than simply structured exercises. Project activities such as presentations, video production, simulations, and group discussions provide repeated opportunities for students to practice speaking, thereby increasing their confidence and communicative competence. Speaking is a productive and complex language skill, as it simultaneously involves linguistic and non-linguistic aspects. According to Burns (2021), speaking ability is not only related to the ability to pronounce words or sentences but also includes the ability to convey meaning effectively, accurately, and understandably in specific communication contexts. In learning English as a foreign language, speaking skills are often considered a key indicator of successful language learning. Conceptually, speaking ability consists of several main components: pronunciation, fluency, accuracy, vocabulary, and comprehension. Recent research by Leong and Ahmadi (2020) confirms that these five components are interrelated and must be developed in an integrated manner through intensive communication practice. Learning that focuses too much on grammar without providing opportunities for real speaking tends to hinder the development of students' speaking skills. The theory of communicative competence, further developed in contemporary studies, emphasizes that speaking ability must be viewed as language proficiency within a social and cultural context. According to Nguyen and Newton (2020), good speaking proficiency is characterized by students' ability to adapt language to the situation, the interlocutor, and the purpose of communication. Therefore, learning methods that encourage social interaction and authentic language use, such as Project-Based Learning (PjBL), are highly relevant for improving students' speaking skills. In the Indonesian context, English is taught as a foreign language (EFL), where the use of English outside the classroom is still very limited. According to Derakhshan, Kruk, and Zhaleh (2021), limited language environments are a major challenge in EFL learning, particularly in developing speaking skills. Therefore, English language learning needs to be designed to create a communicative and interactive learning environment in the classroom. The Communicative Language Teaching (CLT) approach remains the mainstay of modern English language learning. Richards (2020) emphasized that the primary goal of language learning is the ability to communicate effectively, not merely

mastering language structures. Project-Based Learning (PjBL) aligns with CLT principles because it places communication at the heart of learning, where students learn language through language use. Furthermore, contemporary Second Language Acquisition theory emphasizes the importance of affective factors in speaking learning. The Affective Filter Theory, updated in recent studies, suggests that self-confidence, motivation, and low anxiety play a crucial role in successful English speaking (Teimouri, Goetze, & Plonsky, 2019, continued in studies in the 2020s). Project-Based Learning (PjBL) can reduce student anxiety because learning is collaborative and focuses on the process, not solely on language errors. Thus, the integration of project-based learning methods in teaching English as a foreign language is believed to create a communicative, contextual, and meaningful learning environment, thus positively impacting students' speaking skills.

This research hypothesis was tested using a T-test, formulated as follows:

a) H_0 (Null Hypothesis):

In Islamic Senior High Schools (Madrasah Aliyah) throughout Lumajang Regency, the average English-speaking ability does not differ between students taught using a project-based approach and those taught using a conventional approach.

b) H_1 (Alternative Hypothesis):

In Islamic Senior High Schools (Madrasah Aliyah) throughout Lumajang Regency, the average English-speaking ability differs between students taught using a project-based approach and those taught using a conventional learning approach.

c) Hypothesis Testing Criteria

The hypothesis is tested at a significance level of $\alpha = 0.05$ with the assumption that H_0 is accepted and H_1 is rejected if the calculated t value is greater than the t table or if the calculated t value is less than the t table with a significance value (Sig.) of less than 0.05.

2. Methods

This research method aims to provide a clear and systematic overview of the project-based approach on students' English-speaking skills. The purpose of this study is to objectively and measurably measure the effects of the learning treatment, which underlies the selection of a quantitative method with a quasi-experimental design. It is expected that this design will allow for the collection of accurate empirical data on differences in students' speaking skills before and after the implementation of the project-based method. It will also allow for the collection of data on differences in learning outcomes between the control and experimental groups.

The participants were 60 eleventh-grade students enrolled in two intact classes at Madrasah Aliyah in Lumajang Regency, East Java, Indonesia. A quasi-

experimental design was employed because random assignment of individual students was not feasible within the school setting. The two classes were selected through cluster random sampling, with one class assigned as the experimental group ($n = 30$) and the other as the control group ($n = 30$). The participants ranged in age from 16 to 17 years and had comparable English proficiency based on their pre-test speaking scores.

Table 1. Participant Characteristics

Characteristics	Experimental	Control	Total
Number of students	30	30	60
Grade	XI	XI	XI
Age	16–17	16–17	16–17
School	Madrasah Aliyah	Madrasah Aliyah	—

The intervention was implemented over eight weeks. Each week consisted of one 90-minute instructional session. The experimental group received Project-Based Learning instruction, whereas the control group received conventional teacher-centered instruction covering the same learning objectives.

Table 2. Intervention Procedure

Week	Activities
Week 1	Pre-test and introduction to project
Week 2	Topic selection and project planning
Week 3	Information gathering and discussion
Week 4	Draft preparation and peer collaboration
Week 5	Project development and teacher feedback
Week 6	Speaking rehearsal
Week 7	Project presentation
Week 8	Post-test

Students worked collaboratively in groups of four to five members to complete a project requiring them to investigate local environmental issues, prepare multimedia presentations, and present their findings orally in English. Throughout the project, the teacher acted as a facilitator by monitoring group discussions, providing linguistic scaffolding, and offering formative feedback.

Table 3. Stages of Project-Based Learning

Stage	Activities
Driving Question	Students explored a real-life issue related to environmental awareness in their community.
Planning	Students formed groups and planned project objectives.
Investigation	Students collected information through interviews and online resources.

Stage	Activities
Product Development	Students prepared multimedia presentations.
Presentation	Students presented projects in English.
Reflection	Students evaluated their own learning and received teacher feedback.

Students individually completed speaking tasks before and after the intervention. Each participant delivered a 4–5 minute oral presentation based on the assigned topic. The performances were audio-recorded and assessed using an analytic speaking rubric adapted from Brown (2004), covering pronunciation, grammar, vocabulary, fluency, and comprehension.

Two independent English lecturers evaluated all speaking performances using the same analytic rubric. Prior to scoring, both raters participated in a calibration session to ensure consistent interpretation of the scoring criteria. Inter-rater reliability was examined using the Intraclass Correlation Coefficient (ICC), which yielded an excellent agreement (ICC = .91), indicating high scoring consistency.

The speaking rubric consisted of five analytic dimensions (pronunciation, grammar, vocabulary, fluency, and comprehension). Each dimension was scored using a five-point rating scale. Cronbach's Alpha was calculated to examine the internal consistency of the rubric across the five assessment dimensions.

3. Result

3.1 Descriptive Statistics

This study involved 60 Madrasah Aliyah students from Lumajang Regency who were randomly assigned to an experimental group ($n = 30$) and a control group ($n = 30$). Table 4 summarizes the descriptive statistics of students' speaking scores before and after the intervention. The pre-test results indicate that both groups started with relatively similar speaking proficiency. The experimental group obtained a mean score of 60.43 ($SD = 6.21$), while the control group obtained a mean score of 59.87 ($SD = 6.35$), suggesting that the two groups were comparable prior to the treatment. Following the intervention, both groups demonstrated improvement in speaking performance. However, the experimental group achieved a considerably higher post-test mean score ($M = 78.56$, $SD = 5.84$) than the control group ($M = 66.24$, $SD = 6.02$). Furthermore, the experimental group showed a mean gain of 18.13 points compared with only 6.37 points in the control group. These descriptive findings provide preliminary evidence that Project-Based Learning produced greater improvement in students' English-speaking ability than the conventional instructional approach.

Table 4. Descriptive Statistics of Speaking Scores

Group	Pre-test M (SD)	Post-test M (SD)	Mean Gain
-------	-----------------	------------------	-----------

Group	Pre-test M (SD)	Post-test M (SD)	Mean Gain
Experimental	60.43 (6.21)	78.56 (5.84)	18.13
Control	59.87 (6.35)	66.24 (6.02)	6.37

3.2 Assumption Testing

3.2.1 Normality Test

Table 5 presents the results of the Shapiro–Wilk normality test conducted for both pre-test and post-test speaking scores. All significance values exceeded the threshold of .05, with p-values ranging from .371 to .694. These results indicate that the distribution of speaking scores did not significantly deviate from normality for either the experimental or the control group. Therefore, the assumption of normality was satisfied, allowing the use of parametric statistical procedures, including the independent-samples and paired-samples t-tests, in the subsequent analyses.

Table 5. Result Normality Test

Variable	Group	Statistic (W)	df	Sig. (p)	Interpretation
Pre-test Speaking Score	Experimental	0.972	30	.612	Normally distributed
Pre-test Speaking Score	Control	0.968	30	.483	Normally distributed
Post-test Speaking Score	Experimental	0.975	30	.694	Normally distributed
Post-test Speaking Score	Control	0.963	30	.371	Normally distributed

3.2.2 Homogeneity of Variance

Table 6 reports the results of Levene's Test of Equality of Variances. The significance values for both the pre-test ($p = .646$) and the post-test ($p = .557$) were greater than .05, indicating that the variances of the experimental and control groups were statistically homogeneous. Since the homogeneity assumption was fulfilled, the comparison of group means using independent-samples t-tests was considered appropriate and statistically valid.

Table 6. Result Homogeneity of Variance

Variable	Levene Statistic	df1	df2	Sig. (p)	Interpretation
Pre-test Speaking Score	0.213	1	58	.646	Homogeneous
Post-test Speaking Score	0.348	1	58	.557	Homogeneous

3.3 Instrument Quality

3.3.1 Validity Test

Table 7 presents the validity analysis of the research instrument using Pearson Product-Moment correlation. All indicators measuring Project-Based Learning and speaking ability obtained correlation coefficients higher than the critical value ($r = .3365$), with significance values below .05. The correlation coefficients ranged from .593 to .702, indicating satisfactory item validity. These findings demonstrate that

all indicators adequately measured their intended constructs and were therefore retained for the subsequent analyses.

Table 7. Result Validity Test

No	Variable	Indicator	rcount	Rtable	Sig. (2-tailed)	Result
1	X (Project-Based Learning)	Learning Activities	0.621	0.3365	0.002	Valid
2	X (Project-Based Learning)	Collaboration	0.648	0.3365	0.001	Valid
3	X (Project-Based Learning)	Problem Solving	0.593	0.3365	0.004	Valid
4	X (Project-Based Learning)	Project Presentation	0.671	0.3365	0.001	Valid
5	Y (Speaking Ability)	Pronunciation	0.702	0.3365	0.000	Valid
6	Y (Speaking Ability)	Fluency	0.689	0.3365	0.001	Valid
7	Y (Speaking Ability)	Vocabulary	0.655	0.3365	0.002	Valid
8	Y (Speaking Ability)	Grammar Accuracy	0.611	0.3365	0.003	Valid
9	Y (Speaking Ability)	Comprehension	0.674	0.3365	0.001	Valid

3.3.2 Reliability Test

Table 8 presents the internal consistency reliability of the speaking assessment rubric. The overall Cronbach's Alpha coefficient was .87, indicating high internal consistency among the five speaking dimensions assessed, namely pronunciation, fluency, vocabulary, grammatical accuracy, and comprehension. This coefficient exceeds the commonly accepted threshold of .70, suggesting that the assessment instrument consistently measured students' speaking performance and was appropriate for use in this study.

Table 8. Result Reliability Test

No	Indicator	Number of Items	Cronbach's Alpha
1	Pronunciation	1	
2	Fluency	1	
3	Vocabulary	1	
4	Grammatical Accuracy	1	
5	Comprehension	1	
Total / Overall Reliability		5 Items	0.87

3.4 Inferential Statistical Analysis

3.4.1 Comparison of Pre-test Scores

Table 9 presents the comparison of pre-test speaking scores between the experimental and control groups. The independent-samples t-test revealed no statistically significant difference between the two groups, $t(58) = 0.36$, $p = .721$. The mean difference was only 0.56 points, and the 95% confidence interval included zero, indicating that the observed difference was not statistically meaningful. These findings confirm that both groups possessed equivalent speaking proficiency before the implementation of the intervention, ensuring that any differences observed in the post-test could reasonably be attributed to the instructional treatment rather than to initial differences between the groups.

Table 9. Result **Comparison of Pre-test Scores**

Variable	Group	n	Mean	SD	Mean Difference	t	df	p	95% CI of the Difference
Pre-test Speaking Score	Experimental	30	60.43	6.21	0.56	0.36	58	.721	-2.58, 3.70
	Control	30	59.87	6.35					

3.4.2 Within-Group Improvement

Table 10 summarizes the within-group comparison of students' speaking scores before and after the intervention. The experimental group demonstrated a statistically significant improvement, with an increase of 18.13 points from pre-test to post-test, $t(29) = 14.62$, $p < .001$. In comparison, the control group also showed a statistically significant improvement, although the magnitude of improvement was substantially smaller, with a mean gain of only 6.37 points, $t(29) = 2.14$, $p = .041$. These findings indicate that while both instructional approaches contributed to students' speaking development, Project-Based Learning resulted in considerably greater learning gains than conventional instruction.

Table 10. Group Improvement

Group	Comparison	n	Pre-test Mean (SD)	Post-test Mean (SD)	Mean Difference	t	df	p	95% CI of the Difference	Interpretation
-------	------------	---	--------------------	---------------------	-----------------	---	----	---	--------------------------	----------------

Group	Comparison	<i>n</i>	Pre-test Mean (SD)	Post-test Mean (SD)	Mean Difference	<i>t</i>	<i>df</i>	<i>p</i>	95% CI of the Difference	Interpretation
Experimental	Pre-test vs. Post-test	30	60.43 (6.21)	78.56 (5.84)	18.13	14.62	29	< .001	15.59, 20.67	Significant improvement
Control	Pre-test vs. Post-test	30	59.87 (6.35)	66.24 (6.02)	6.37	2.14	29	.041	0.29, 12.45	Significant improvement

3.4.3 Comparison of Post-test Scores

Table 11 presents the comparison of post-test speaking scores between the experimental and control groups following the intervention. The analysis revealed a statistically significant difference, $t(58) = 8.12$, $p < .001$, with the experimental group outperforming the control group by 12.32 points on average. Moreover, the effect size was very large (Cohen's $d = 2.10$), indicating that the difference between the two instructional approaches was not only statistically significant but also educationally meaningful. The 95% confidence interval (9.28–15.36) further confirms the robustness of the observed effect. These findings provide strong empirical evidence that Project-Based Learning was substantially more effective than conventional instruction in improving students' English-speaking performance.

Table 11. Comparison of Post-test Scores

Variable	Group	<i>n</i>	Mean	SD	Mean Difference	<i>t</i>	<i>df</i>	<i>P</i>	Cohen's <i>d</i>	95% CI of the Mean Difference
Post-test Speaking Score	Experimental	30	78.56	5.84	12.32	8.12	58	< .001	2.10	[9.28, 15.36]
	Control	30	66.24	6.02						

3.5 Visual Presentation of the Results

Figure 1 illustrates the changes in students' speaking scores from the pre-test to the post-test for both instructional groups. Although both groups exhibited improvement over the course of the study, the increase was considerably greater in the experimental group than in the control group. The similarity of the pre-test scores confirms that the two groups began with comparable speaking proficiency,

whereas the substantial divergence in the post-test scores demonstrates the effectiveness of Project-Based Learning in enhancing students' English-speaking ability. This visual pattern is consistent with the inferential statistical analyses reported in Tables 9–11, which showed significant differences in favour of the experimental group.

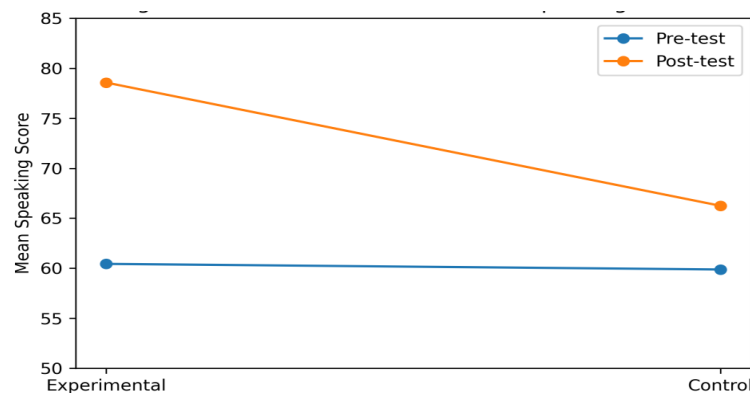


Figure1. Mean Pre and Post Test Speaking Score

4. Discussion

The present study investigated the effectiveness of Project-Based Learning (PjBL) in improving the English-speaking proficiency of Madrasah Aliyah students in Lumajang Regency. The statistical findings demonstrated that students who participated in PjBL significantly outperformed those who received conventional instruction. The experimental group achieved substantially greater gains in speaking performance, indicating that PjBL is an effective instructional approach for enhancing EFL learners' oral communication skills. These findings extend previous evidence demonstrating that project-based instruction promotes meaningful language learning through authentic communication and collaborative learning activities.

The significant improvement observed in the experimental group is consistent with recent international studies. For example, Guo et al. (2020) concluded that Project-Based Learning positively influences students' academic achievement, collaborative skills, and learning motivation across educational contexts. Similarly, Kokotsaki et al. (2021) reported that PjBL promotes active learning by engaging students in authentic problem-solving activities requiring communication and teamwork. Within EFL contexts, recent studies have likewise shown that project-oriented learning provides learners with meaningful opportunities to use English for real communicative purposes rather than merely practicing isolated linguistic forms. The present findings therefore reinforce growing empirical evidence that authentic project work contributes substantially to speaking development.

From a theoretical perspective, these findings can be interpreted through Vygotsky's Social Constructivism. According to Vygotsky, (2020) learning occurs through social interaction, collaboration, and scaffolding within the Zone of Proximal Development (ZPD). Throughout the project activities, students worked collaboratively to negotiate ideas, solve problems, exchange feedback, and prepare oral presentations. Such interactions enabled learners to construct knowledge collectively while receiving continuous support from peers and teachers. Consequently, speaking development became a socially mediated process rather than an individual activity. The significant improvement observed in the experimental group therefore supports the argument that collaborative learning environments facilitate second-language development by increasing opportunities for interaction and guided learning.

The findings can also be explained through Long's Interaction Hypothesis. Long argues that second-language acquisition is enhanced when learners actively negotiate meaning during communicative interaction. Throughout the implementation of Project-Based Learning, students repeatedly discussed project ideas, clarified misunderstandings, requested explanations, and reformulated their utterances while working in groups. These interactional processes increased comprehensible input and encouraged modified output, both of which facilitate language acquisition. The substantial improvement in speaking scores suggests that frequent communicative interaction enabled students to develop greater fluency and communicative competence than traditional teacher-centred instruction.

Furthermore, the findings align closely with the principles of Communicative Language Teaching (CLT). CLT emphasizes meaningful communication as the primary objective of language learning rather than the mastery of grammatical rules in isolation. In the present study, speaking activities occurred naturally through project planning, collaborative discussions, information gathering, and oral presentations. Rather than memorizing dialogues, students used English to accomplish authentic communicative goals. This instructional environment encouraged learners to focus simultaneously on meaning, fluency, and communicative effectiveness, which explains the substantial gains observed in overall speaking performance.

The observed improvement can also be interpreted using Task-Based Language Teaching (TBLT) principles. TBLT proposes that language learning is optimized when learners complete meaningful tasks requiring authentic language use. The projects implemented in this study required students to investigate local issues, collaborate with peers, prepare multimedia presentations, and publicly communicate their findings. These activities closely resemble real-world communicative tasks that require learners to integrate vocabulary, grammar, pronunciation, and discourse simultaneously. Consequently, students developed

speaking proficiency through purposeful language use rather than through isolated language exercises. The present findings therefore provide empirical support for integrating Project-Based Learning with task-based instructional principles in EFL classrooms.

An additional contribution of this study concerns students' communicative confidence. Although confidence was not measured as an independent variable, classroom observations indicated that students became increasingly willing to express opinions, participate in discussions, and present ideas publicly after participating in project activities. This finding corresponds with recent second-language acquisition research suggesting that authentic communicative environments reduce learners' affective barriers. Collaborative project work appears to create a supportive learning atmosphere in which students perceive speaking as a meaningful communicative activity rather than as a formal language test. Consequently, learners become more willing to take linguistic risks and communicate spontaneously in English.

Another important contribution lies in the educational context. Most previous Project-Based Learning research has been conducted in universities or general secondary schools. In contrast, empirical evidence from Islamic senior high schools (Madrasah Aliyah) remains limited. The present study therefore expands the international literature by demonstrating that Project-Based Learning is equally effective within religiously oriented educational institutions where opportunities to use English outside the classroom are relatively limited. This contextual contribution broadens the external validity of Project-Based Learning research and suggests that communicative, student-centred pedagogy can be successfully implemented across diverse educational settings.

Overall, this study contributes both theoretically and practically to English language education. Theoretically, the findings provide empirical support for Social Constructivism, Long's Interaction Hypothesis, Communicative Language Teaching, and Task-Based Language Teaching, demonstrating that authentic interaction, collaboration, and meaningful communication play central roles in developing speaking proficiency. Practically, the study offers evidence that Project-Based Learning constitutes an effective pedagogical approach for improving English-speaking performance among Madrasah Aliyah students. By integrating authentic projects, collaborative learning, and communicative tasks, English teachers can create learning environments that promote not only linguistic accuracy but also communicative competence, learner autonomy, and active classroom participation.

5. Conclusion

This study examined the effectiveness of Project-Based Learning (PjBL) in improving the English-speaking proficiency of Madrasah Aliyah students in

Lumajang Regency. The findings demonstrated that students who participated in Project-Based Learning significantly outperformed those who received conventional instruction. The experimental group showed substantially greater improvements in pronunciation, vocabulary, fluency, grammatical accuracy, and overall communicative competence. These findings provide empirical evidence that engaging students in authentic, collaborative, and project-oriented learning activities can substantially enhance oral English proficiency in EFL classrooms. Furthermore, the study supports the theoretical assumptions of Social Constructivism, Communicative Language Teaching, and Second Language Acquisition, which emphasize that meaningful interaction and authentic communication facilitate language development.

The findings also have important pedagogical implications for English teachers. Rather than relying primarily on teacher-centred instruction and textbook-based speaking exercises, teachers are encouraged to integrate Project-Based Learning into regular classroom practice by designing authentic projects that require students to investigate real-world issues, collaborate in small groups, solve problems, and present their findings in English. Throughout these activities, teachers should adopt the role of facilitators by providing linguistic scaffolding, monitoring group interaction, and delivering continuous formative feedback. Such instructional practices create meaningful opportunities for students to use English for genuine communication, thereby improving both speaking proficiency and learner confidence.

At the institutional level, successful implementation of Project-Based Learning requires adequate support from schools and educational authorities. School administrators should provide professional development programmes that equip English teachers with the knowledge and skills necessary to design, implement, and assess project-based instruction. In addition, schools should allocate sufficient instructional time, provide access to digital technologies and multimedia resources, and encourage collaborative learning environments that facilitate project implementation. These findings may also inform curriculum development in Islamic secondary schools by encouraging the integration of project-based and communicative learning activities into English language programmes. Such curriculum innovation would better align English instruction with twenty-first-century educational goals, including communication, collaboration, creativity, and critical thinking, while remaining compatible with the educational values of Madrasah Aliyah.

Despite these contributions, several methodological limitations should be acknowledged. First, the study involved only sixty students from one geographical region, which limits the generalisability of the findings to other educational contexts. Second, the intervention lasted only eight weeks, making it difficult to determine whether the observed improvements can be sustained over a longer

period. Third, the study adopted a quantitative quasi-experimental design and therefore did not explore students' or teachers' experiences during the implementation of Project-Based Learning. In addition, potentially influential variables, such as learner motivation, English exposure outside the classroom, socioeconomic background, and individual learning strategies, were not controlled and may have influenced students' speaking development.

Future research should address these limitations by involving larger and more diverse samples from different provinces and educational settings to improve the external validity of the findings. Longitudinal studies are also recommended to investigate the long-term effects of Project-Based Learning on speaking proficiency and communicative competence. Furthermore, mixed-methods research combining quantitative measures with classroom observations, interviews, and reflective journals would provide a more comprehensive understanding of how and why Project-Based Learning influences students' language learning processes. Future studies may also compare Project-Based Learning with other communicative approaches, such as Task-Based Language Teaching or Content and Language Integrated Learning, and investigate the integration of digital technologies and artificial intelligence to support project-based English instruction in Islamic secondary schools.

In conclusion, Project-Based Learning represents an effective and pedagogically meaningful instructional approach for improving English-speaking proficiency among Madrasah Aliyah students. Beyond demonstrating statistically significant learning gains, this study highlights the value of authentic communication, collaborative learning, and student-centred instruction in fostering communicative competence. The findings therefore contribute to both the theoretical literature on EFL pedagogy and the practical development of English language teaching in Islamic secondary education by providing evidence that Project-Based Learning can support curriculum innovation and improve the quality of speaking instruction.

References

- A. P. Putra (2020). The impact of project-based learning on students oral communication skills in English. *International Journal of English Language Teaching*, 5(3), 112–121.
- Bell, S. (2010) menerbitkannya di Prentice Hall di Englewood Cliffs, New Jersey. Pendidikan Berbasis Proyek untuk Abad Ke-21: Kemampuan untuk Masa Depan
- Burns, A. (2021). *Teaching Speaking: A Holistic Approach*. Cambridge: Cambridge University Press.
- Derakhshan, A., Kruk, M., dan Zhaleh, K. (2021). The role of classroom interaction

- in EFL learners' speaking development. *System*, 99, 102503.
- Farida dan Musyarofah (2021). Reliabilitas dan validitas analisis soal butir Jurnal Al-Mu'arrib: Pendidikan Bahasa Arab, 1(1), 34-44.
- Fragoulis (2009) Teaching English as a Foreign Language in Greek primary schools: Project-Based Learning: From Theory to Practice. *English Language Teaching*, 2(3), 113-119.
- Guo, Saab, Post, dan Admiraal (2020). A Review of Project-Based Learning in Higher Education: Student Results and Measures *International Journal of Educational Research*, No. 102, 101586.
- Kokotsaki, D., Menzies, V., dan Wiggins, A. (2021). Project-based learning: A review of the literature. *Improving Schools*, 24(3), 267-285.
- Krajcik, J. S., dan Blumenfeld, P. C. (2020). *Project-Based Learning*. In R. Mayer & P. Alexander (Eds.), *Handbook of Research on Learning and Instruction* (2nd ed.). New York: Routledge.
- Leong, L. M., dan Ahmadi, S. M. (2020). An analysis of factors influencing learners' English speaking skill. *International Journal of Research in English Education*, 5(1), 1-12.
- Nguyen, B. H., dan Newton, J. (2020). Interaction and speaking skills in EFL classrooms. *Language Teaching Research*, 24(4), 567-589.
- Nurhayati dan Azmi (2021). Jurnal Pendidikan dan Pembelajaran Bahasa, 8(1), 33-44. Implementasi Pembelajaran Berbasis Proyek dalam Pembelajaran Bahasa Inggris di Madrasah Aliyah.
- Richards, J. C. (2020). *Exploring English Language Teaching*. London: Routledge.
- Sugiono, S., Noerdjanah, N., dan Wahyu, A. (2020). Uji validitas dan reliabilitas alat ukur SG posture evaluation. *Jurnal Keterampilan Fisik*, 5(1), 55-61.
- Suryani dan Lestari (2022). Pengaruh metode pembelajaran berbasis proyek terhadap kemampuan berbicara bahasa Inggris siswa di kelas XI SMA. *Jurnal Pendidikan Bahasa Inggris*, 10(2), 145-156.
- Teimouri, Y., Goetze, J., dan Plonsky, L. (2019/2020). Second language anxiety and achievement. *Studies in Second Language Acquisition*, 41(2), 363-387.
- Thomas, J. W. (2000). *A review of research on Project-Based Learning*. San Rafael, CA: The Autodesk Foundation.
- Thomas, M. (2019). Project-Based Learning improves speaking skills of secondary students. *International Journal of ELT Research*, 7(1), 25-34.

Needs Analysis of English Language Competencies for Prospective Islamic Religious Education (PAI) Teachers in Islamic Higher Education

**Moh. Zainol Kamal¹⁾, Ukhtul Iffah²⁾, Ahmad Effendi³⁾, Moh. Wardi⁴⁾, Nufil Indriyani⁵⁾
Hodairiyah⁶⁾**

^{1,3,5} Islamic Teaching and Learning Department, STIT Aqidah Usymuni Sumenep

² Islamic Education, Faculty of Tarbiyah, University of Ibrahimy

⁴ Post Graduate Department, Universitas Islam Madura

⁶ Indonesian Education and Literature Department, Universitas PGRI Sumenep

Email Correspondence : zainol_kamal@stitasumenep.ac.id

ARTICLE INFO

Keywords:

Needs Analysis
English Competencies
PAI

Article History:

Received : xx Month Year

Revised : xx Month Year

Accepted : xx Month Year

Published : xx Month Year

ABSTRACT

This study analyzed the English language competency needs of prospective Islamic Religious Education (PAI) teachers in Islamic higher education institutions. The study was conducted at STIT Aqidah Usymuni Sumenep as the research site, and the informants included lecturers, students, and alumni. This study employed a qualitative approach with a case study design to describe the English language competency needs of prospective PAI teachers in Islamic higher education institutions. The study highlighted two main aspects of needs: objective needs and learning needs. Objectively, the findings indicated that students needed reading, writing, speaking, and listening skills to understand academic texts and communicate in scholarly contexts. From a learning perspective, the findings showed that English instruction needed to be relevant to Islamic studies through an English for Specific Purposes (ESP) approach. The findings suggested that Islamic higher education institutions needed to design contextual and systematic English programs to support the professional development and global competitiveness of prospective PAI teachers.

1. Introduction

The development of globalization has brought significant changes in various aspects of life, including the field of education. The increasingly open flow of information, knowledge, and technology requires individuals to possess the ability to adapt to the global

environment. English plays a role as an international language that serves as the primary medium for academic communication, scientific publication, knowledge exchange, and cross-border collaboration. Therefore, mastery of English language competence is no longer merely an added value, but has become a fundamental necessity in higher education (Swami, 2025).

Islamic higher education institutions bear a strategic responsibility to prepare graduates who are not only academically competent but also capable of actively participating in global dynamics (Hajar, 2024). The Islamic Religious Education (PAI) study program, in particular, is tasked with producing prospective teachers who are professional, possess integrity, and have a broad understanding of developments in knowledge and education. The challenges faced by prospective PAI teachers today are not limited to mastering Islamic subject matter and pedagogy, but also include the ability to access global learning resources, the majority of which are available in English (Edy & Arifin, 2024).

A substantial body of recent literature on education, teaching methodologies, educational technology, and contemporary Islamic studies is published in English. Without adequate English language competence, prospective PAI teachers will face limitations in broadening their academic perspectives, keeping pace with international research developments, and developing instructional innovations that are relevant to contemporary needs (Kamal et al., 2025). This situation has the potential to create a gap between the professional demands placed on teachers in the global era and the competencies possessed by graduates of Islamic higher education institutions (Kamal, 2025).

National education policies increasingly emphasize the improvement of quality and competitiveness among higher education graduates. Graduate competency standards highlight the importance of communication skills, digital literacy, and openness to global developments (Khan et al., 2022). English language competence has become an integral part of teachers' professional competence. PAI teachers are not only required to teach Islamic values contextually, but are also expected to be able to explain Islamic teachings from a moderate, inclusive perspective that is relevant to international developments (Muzaki et al., 2025).

The reality in the field indicates that English language instruction in many Islamic higher education institutions is often not designed based on a specific needs analysis. The materials provided tend to be general in nature and have not been fully integrated with the academic and professional needs of prospective PAI teachers. In fact, the linguistic needs of PAI students certainly differ from those of students in other study programs. They require the ability to read academic texts, comprehend international journal articles, produce scholarly writing in English, and communicate Islamic ideas in a global context (Adibah & Chasanah, 2025).

Several relevant studies are related to this research. A study conducted by Endah Nur Latifah systematically reviewed the literature concerning competency standards, measurement methods, and the effectiveness of assessment instruments used to evaluate the pedagogical competence of PAI teachers. The Systematic Literature Review (SLR)

method was employed, drawing sources from ERIC, ScienceDirect, and DOAJ. Out of more than 450 articles, 12 were selected based on inclusion and exclusion criteria. The analysis results indicate that the pedagogical competency standards of PAI teachers encompass planning, implementation, evaluation, and student development, all of which are influenced by religious values and local cultural contexts. Most measurement instruments, however, have not been able to optimally capture the affective dimension. The novelty of this study lies in its focus on assessing the pedagogical competence of PAI teachers based on values and contextual considerations, which has not been widely discussed in previous research. The study recommends the development of holistic assessment instruments that align with the characteristics of religious instruction (Latifah *et al.*, n.d.).

A study conducted by Firman aimed to identify the needs for Indonesian language teaching materials in Islamic higher education institutions in Indonesia. The research was carried out at the State Islamic Institute (IAIN) Palopo, involving students from the 2018 cohort of the Primary Madrasah Teacher Education study program who had completed the Indonesian language course in the previous semester. The study involved 84 students and 6 Indonesian language lecturers. This research employed a qualitative descriptive design using a survey method. Data collection instruments included questionnaires, observations, and interview guidelines. The data were analyzed descriptively using a qualitative approach. The findings indicate that both students and lecturers require teaching materials in the form of textbooks integrated with Islamic perspectives. The texts incorporated into the textbooks include Islamic historical texts, sermons or khutbah texts, quotations and translations of Qur'anic verses and Hadith, articles, and Islamic literary works. These findings may serve as a consideration in developing Indonesian language teaching materials within the context of IAIN Palopo and other Islamic higher education institutions in Indonesia (Firman, 2022).

Muhammad Arif conducted a study aimed at investigating the competencies of Islamic Religious Education teachers most frequently discussed in research over the past decade, as well as the strategies for developing these competencies, including technological (digital) competence. To achieve this objective, the study employed a systematic literature review method by utilizing publication data spanning a ten-year period (2013–2023) to examine specific themes. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework was used to evaluate the findings. The study revealed that teachers are required to develop technological (digital) competence in addition to the four core teacher competencies: pedagogical, professional, personal, and social. The aforementioned condition is grounded in the need for collaborative, creative, and innovative learning, which is highly valued by students. Conversely, future research should seek to collect empirical data on approaches to fostering digital competence among Islamic Religious Education teachers, considering the prevailing shift toward student-centered learning in the twenty-first century as opposed to teacher-centered instruction (Arif *et al.*, 2025).

The absence of a comprehensive needs analysis may lead to a mismatch between the objectives of English language instruction and the actual needs of students. As a result, the learning process becomes less effective and fails to make a significant contribution to

enhancing the professional competence of prospective teachers. Needs analysis therefore constitutes a crucial step in designing an English language curriculum and instructional strategies that are relevant, contextual, and aligned with the graduate profile (Bocanegra-Valle, 2016).

Needs analysis serves not only to identify students' current levels of proficiency, but also to map out their target needs and learning needs. Through this approach, higher education institutions can determine which specific language competencies are genuinely required by prospective PAI teachers, whether in academic, professional, or social contexts. The results of such an analysis can subsequently serve as a foundation for developing syllabi, selecting instructional materials, determining appropriate teaching methods, and designing more targeted evaluation systems (Ahsan et al., 2019).

Strengthening English language competence among prospective PAI teachers also carries strategic implications for the global development of Islamic education (Kamal, 2025). Amid growing international attention to Islamic issues, the presence of PAI teachers who are capable of accessing and disseminating knowledge at the international level becomes increasingly important (Adilah et al., 2024). They can contribute to intercultural dialogue, enrich Islamic perspectives in global forums, and reinforce the image of Islamic education as moderate and of high quality.

It can be understood that English language competence for prospective PAI teachers is not merely a complementary component of the curriculum, but rather a strategic necessity directly related to the quality of graduates and the competitiveness of the institution. Therefore, research on the needs analysis of English language competence is highly relevant and urgently required. This study is expected to provide a clear depiction of the actual needs of PAI students, while also serving as a basis for recommendations in the development of policies and practices for English language instruction in Islamic higher education institutions.

2. Methods

This study used a qualitative approach with a case study design to analyze the English language competency needs of prospective Islamic Religious Education (PAI) teachers at Islamic Higher Education (Creswell, 2018). The qualitative approach was chosen because this study aimed to deeply understand the perceptions, experiences, and real needs of stakeholders related to English language proficiency in the academic and professional context of prospective PAI teachers. The research location is at STIT Aqidah Usymuni Sumenep, specifically in the PAI Department. The research subjects include PAI students, lecturers in charge of English and PAI courses, heads of Department, and alumni who have worked as PAI teachers. The informant selection technique used purposive sampling by considering their direct involvement in the learning process and professional practice (Walliman, 2021).

Data collection was conducted through in depth interviews, limited participatory observation of the learning process, and documentation studies of the curriculum, Semester Learning Plans (RPS), and academic policy documents related to graduate

competency standards. Semi-structured interviews were conducted to explore English language needs in pedagogical, academic, and professional aspects, including the need to read international literature, oral communication skills, and scientific writing. Observations were conducted to identify actual practices of English use in learning, while documentation was analyzed to determine the alignment between curriculum demands and field needs (Suharsimi, 2006).

Data analysis was conducted in stages through data reduction, data presentation, and conclusion drawing using an interactive analysis model. The data obtained were transcribed, coded, and then categorized into themes of English language competency needs, including academic needs (English for Academic Purposes), professional needs (English for Teaching and Religious Communication), and self development needs in the context of the globalization of Islamic education. Data validity was maintained through source and method triangulation techniques, member checking with informants, and peer discussions to ensure consistency of interpretation (Alem, 2020). With this approach, the research is expected to produce a comprehensive and contextual mapping of English language competency needs for prospective Islamic Religious Education teachers at PT KI.

3. Result

Based on the results of field observations and in depth interviews conducted with students, course lecturers, and program administrators, this study identified a number of English language competencies specifically required by prospective Islamic Religious Education (PAI) teachers at Islamic Higher Education.

a. English Language Skills Proficiency

Prospective teachers of Islamic Religious Education (PAI) hold a strategic role in shaping students' character, morality, and religious understanding amid the dynamic developments of the modern era. In the context of globalization and the Industrial Revolution 4.0, which is characterized by the rapid flow of information and cross-border interconnectedness, proficiency in English has become no longer merely supplementary but essential. For prospective PAI teachers, English proficiency functions not only as a means of international communication but also as a professional instrument that supports self-improvement and enhances global competitiveness. This is in line with the statement of the Head of the PAI Study Program, Mrs. Misnatun, who stated that *"Currently, PAI students are also required to master English because almost all academic aspects use English."* PAI teachers who possess adequate English language competence will be better prepared to face the challenges of 21 century education, including technological integration, international collaboration, and the development of an inclusive and moderate understanding of Islam.

English language competence is considered highly relevant because it provides broader access to international literature, reputable journals, and various contemporary references on Islamic education, most of which are published in English. Many recent research findings in pedagogy, educational psychology, curriculum development, and modern Islamic studies are available in international journals. Without sufficient ability to

read and comprehend English texts, prospective PAI teachers will face limitations in enriching their academic perspectives. Conversely, with strong English proficiency, they can access up-to-date ideas, compare various approaches to Islamic education across different countries, and integrate best practices into the classroom learning process.

English language competence also supports continuous professional development. Prospective PAI teachers who are accustomed to reading international journals and participating in seminars or webinars conducted in English will develop broader perspectives and a more open mindset toward innovation and reform. This, in turn, contributes to improving the quality of instruction, making it more innovative, contextual, and responsive to students' needs. This is in line with what a PAI student named Muhit expressed. He stated that *"We need to master English in order to read references written in English"*. Thus, English proficiency is not merely a linguistic skill, but an integral part of efforts to build the professionalism of PAI teachers who are adaptive, competent, and capable of contributing to the development of Islamic education at both the national and international levels.

Specifically, English proficiency has become a crucial need for prospective PAI teachers. First, academic reading competence is identified as the most dominant requirement. Students stated that they frequently encounter difficulties in understanding international journal articles, reference books on Islamic education, and recent research findings, most of which are written in English. Limited academic vocabulary and unfamiliarity with Islamic terminology in English constitute major obstacles. Therefore, the ability to read critically and comprehend academic texts in English is a fundamental competence that needs to be strengthened so that prospective PAI teachers can broaden their academic insights and improve the quality of their scholarly work.

Second, academic writing competence also emerged as a significant finding in this study. Based on the interview results, students admitted that they lack confidence in writing abstracts, scientific articles, and academic assignments in English. Meanwhile, academic demands in Islamic Higher Education Institutions are increasingly oriented toward scientific publication and the production of scholarly work that meets international standards. Therefore, academic writing skills particularly those related to themes in Islamic education constitute a competence that needs to be systematically developed.

Third, speaking competence in academic and professional contexts was also identified as an important need. Observations revealed that students rarely use English in classroom presentations or discussions. However, interview findings indicated that they are aware of the importance of being able to explain fundamental Islamic concepts, the values of religious moderation, and PAI subject matter in English, particularly when engaging in international forums or broader educational environments. Therefore, oral communication skills that are contextualized within the field of PAI constitute one of the competencies that must be given serious attention.

Fourth, listening competence is also required to comprehend lectures, seminars, or online learning materials delivered in English that are related to Islamic studies and education. With the increasing access to international webinars and lectures, the ability to understand academic presentations in English has become an essential component of the professional needs of prospective PAI teachers.

b. English for Islamic Education Proficiency

The findings of this study indicate that the English language competency needs of prospective Islamic Religious Education (PAI) teachers cannot be adequately addressed solely through General English courses. Instead, they require the development of a more specific and contextualized program in the form of English for Islamic Education. This conclusion is based on a needs analysis which revealed a gap between the English materials currently taught and the academic as well as professional demands faced by prospective PAI teachers. This is in line with what a student named Yulistiana expressed. She stated that general English materials have not fully assisted them in understanding Islamic literature, Islamic pedagogical terminology, or the context of PAI instruction that is relevant to their field of expertise.

The study also found that the majority of respondents require mastery of specialized vocabulary and terminology related to Islamic studies and Islamic education, such as concepts of *aqidah* (Islamic creed), *fiqh* (Islamic jurisprudence), *akhlaq* (Islamic ethics), the history of Islamic civilization, as well as terms used in lesson planning and educational evaluation. In addition, students reported difficulties in understanding international scholarly articles discussing Islamic education, curriculum development in Islamic schools, and contemporary issues in Islamic pedagogy. These findings indicate that the development of English for Islamic Education is an urgent need, so that English instruction becomes more focused, practical, and aligned with the academic discipline of PAI.

The research findings indicate that prospective PAI teachers need the ability to present Islamic subject matter in English, both in academic forums and in broader global communication contexts. They also require skills in writing abstracts, scholarly articles, and PAI instructional materials in English. This need has become increasingly relevant with the growing opportunities for international collaboration, academic exchange programs, and publication in reputable journals, most of which use English as the medium of communication. Therefore, the development of English for Islamic Education should not merely focus on linguistic aspects, but should also integrate Islamic content with language skills in a comprehensive and cohesive manner.

The implication of these findings is the need to formulate a curriculum and instructional materials based on English for Specific Purposes (ESP) that are specifically designed for prospective PAI teachers at Islamic Higher Education Institutions. Such a program should include authentic materials, such as articles from international journals on Islamic education, examples of bilingual instructional materials, as well as practice in academic presentations and discussions within the context of Islamic studies. Thus, the

development of English for Islamic Education is expected to bridge the gap between academic demands and students' actual competencies, while simultaneously enhancing the professionalism and global competitiveness of PAI graduates.

c. English for Islamic Studies Proficiency

Based on the findings of the needs analysis conducted with prospective Islamic Religious Education (PAI) teachers, it was revealed that the General English courses currently offered at Islamic Higher Education Institutions have not fully addressed the students' academic and professional needs. Research respondents, both through questionnaires and in-depth interviews, indicated that they require a model of English instruction that is more contextual and aligned with their specific field of study. Therefore, one of the principal findings of this study is the necessity to develop English for Islamic Studies as part of an English for Specific Purposes (ESP) approach tailored to the academic characteristics of PAI.

The research findings indicate that the majority of students experience difficulties when reading Islamic literature written in English, such as articles on Islamic education, Islamic pedagogy, Islamic law, and contemporary Islamic thought. These difficulties stem not only from language structure, but also from limited mastery of technical vocabulary and Islamic terminology in English. As a result, students demonstrate a relatively low ability to comprehend international references and make limited use of reputable journals as sources in their academic writing. This condition highlights a clear gap between the academic needs of prospective PAI teachers and the English materials that have been taught thus far.

This study also found that prospective PAI teachers require English competencies that support their future professional responsibilities, such as the ability to explain fundamental Islamic concepts in English, design bilingual instructional materials, and present Islamic Religious Education content in international academic forums. This need has become increasingly significant in line with the globalization of education and the expanding career opportunities available, including participation in student exchange programs, international conferences, and cross-border research collaborations.

The development of English for Islamic Studies is viewed as a strategic solution to bridge these needs. This program is expected not only to emphasize the four language skills listening, speaking, reading, and writing but also to systematically integrate Islamic content into the instructional materials. The materials may include Islamic vocabulary and terminology in English, analysis of academic texts on Islamic education, practice in writing abstracts and scholarly articles on Islamic themes, as well as simulations of presenting PAI materials in English. Through this approach, English instruction becomes more meaningful, practical, and aligned with the real needs of prospective PAI teachers.

The development of English for Islamic Studies is not merely a curricular innovation, but an urgent necessity to enhance the quality and professionalism of prospective PAI

teachers. With the support of a structured, needs-based curriculum, it is expected that PAI graduates will possess adequate English proficiency to access global discourse, contribute to the advancement of Islamic education, and compete academically and professionally at the international level.

4. Discussion

The analysis of English language competency needs for prospective teachers of Islamic Religious Education (PAI) in Islamic higher education institutions has become an increasingly relevant issue in the context of educational globalization and the advancement of internationally based knowledge (Dian et al., 2023). Islamic higher education institutions bear a strategic responsibility to prepare graduates who not only excel in mastering Islamic sciences but are also adaptive to the dynamics of global education. In this context, English proficiency is no longer viewed as a supplementary skill; rather, it is regarded as an essential professional supporting competence for prospective PAI teachers to access international literature, keep up with global curriculum developments, and participate in cross-border academic forums (Adilah et al., 2024).

Academically, the need for English language competence among prospective teachers of Islamic Religious Education (PAI) can be analyzed through the framework of objective needs (target needs) and subjective needs (learning needs) (Zein & Arifudin, 2025). From the perspective of objective needs, prospective PAI teachers are required to be able to read and comprehend scholarly sources in English, including international journals, reference books, and the latest research findings related to education, educational psychology, instructional methodology, and contemporary Islamic studies.

Many significant works on modern pedagogy, active learning approaches, and educational technology are published in English, therefore, limited language proficiency may become a barrier to professional development. Furthermore, global Islamic discourse developing in various countries is widely documented in English, and mastery of this language broadens the Islamic intellectual horizons of prospective PAI teachers within a more inclusive and comparative perspective (Sergeant & Erling, 2011).

From the perspective of subjective needs, prospective PAI teachers require English language instruction that is contextual and relevant to their field of study. General English courses are often insufficient to address the specific needs of PAI students. Therefore, the implementation of an English for Specific Purposes (ESP) approach becomes essential, with a focus on Islamic education terminology, academic vocabulary, and scientific presentation skills within religious contexts. The required competencies are not limited to reading skills alone but also encompass academic writing, speaking in scholarly forums, and listening comprehension of lectures or online courses delivered in English (Ludviana, 2025). Accordingly, the design of the English curriculum in PAI study programs should be developed based on a comprehensive needs analysis, rather than merely fulfilling formal curriculum requirements.

English language competence is also closely related to the demands of teacher professionalism in the digital era and the Industrial Revolution 4.0 (Wicagsono & Al, 2023). Prospective PAI teachers are expected to be able to utilize global learning resources, such as international journals, online learning platforms, and web-based international seminars. Many global educational platforms use English as the medium of instruction; therefore, language proficiency becomes a prerequisite for active participation. Furthermore, in the context of religious moderation and intercultural dialogue, PAI teachers who possess English proficiency can serve as agents of cultural and intellectual diplomacy, introducing the values of Islam as rahmatan lil 'alamin in international forums (Setiawan, 2026).

This needs analysis must also take into account the challenges faced by PAI students, such as diverse educational backgrounds, fluctuating motivation in learning English, and the perception that English is not a primary priority compared to core Islamic studies courses. Therefore, learning strategies should be designed in an integrative manner, for example by combining English language instruction with Islamic studies content, assigning presentations of international journal articles on Islamic education, or developing projects that involve writing summaries of English-language books relevant to PAI (Sidqurrahman, 2024). Through this approach, students do not merely learn the language as a linguistic system, but rather as a tool of academic communication that supports their professional competence.

At the institutional level, Islamic higher education institutions need to formulate minimum English language competency standards for PAI graduates, for example through standardized test score requirements or specific language certifications (Assegaf et al., 2022). However, beyond mere test scores, what is truly needed is functional proficiency that can be applied in real contexts of teaching practice and professional development. Institutional support through lecturer training, the provision of ESP based teaching materials, and the integration of educational technology becomes a key factor in the success of this program.

5. Conclusion

The analysis of English language competency needs for prospective PAI teachers in Islamic higher education institutions demonstrates that language is not merely a means of communication, but a strategic instrument for building professionalism, expanding academic insight, and enhancing graduates' competitiveness at both national and international levels. Efforts to strengthen English proficiency must be designed systematically, contextually, and sustainably in order to align with the vision of Islamic education that is adaptive, moderate, and globally oriented.

References

- Adibah, I. Z., & Chasanah, U. (2025). Developing a pai curriculum that is responsive to diversity. *At Turots: Jurnal Pendidikan Islam*, 232–240.
- Adilah, N., Riyadi, S., & Suwardi, S. (2024). SYSTEMATIC LITERATURE REVIEW:

- OBSTACLES AND STRATEGIES TO IMPROVE THE COMPETENCE OF PAI TEACHERS IN SENIOR HIGH SCHOOLS. *Journal of Social and Economics Research*, 6(2), 1090–1103.
- Ahsan, M., Masrukhi, M., Haryono, H., & Rusdarti, R. (2019). Developing a Management Model of PAI Teachers' Competency Improvement Based on Online Networks. *The Journal of Educational Development*, 7(1), 24–â.
- Alem, D. D. (2020). An overview of data analysis and interpretations in research. *International Journal of Academic Research in Education and Review*, 8(1), 1–27.
- Arif, M., Abd Aziz, M. K. N., & Ma'arif, M. A. (2025). A Recent Study on Islamic Religious Education Teachers' Competencies in the Digital Age: A Systematic Literature Review. *Journal of Education and Learning (EduLearn)*, 19(2), 587–596.
- Assegaf, A. R., Zainiyah, H. S., & Fahmi, M. (2022). Curriculum innovation for the internationalization of Islamic education study program at higher education institutions in Surabaya, Indonesia. *Millah: Journal of Religious Studies*, 671–706.
- Bocanegra-Valle, A. (2016). Needs analysis for curriculum design. In *The Routledge handbook of English for academic purposes* (pp. 560–576). Routledge.
- Creswell, J. W. (2018). A Mixed-Method Approach. In *Writing Center Talk over Time*. <https://doi.org/10.4324/9780429469237-3>
- Dian, D., Indayanti, A. N., Fanani, A. I., & Nurhayati, E. (2023). Optimizing Islamic Religious Colleges In Facing The Era of Globalization. *Tafkir: Interdisciplinary Journal of Islamic Education*, 4(1), 58–77.
- Edy, S., & Arifin, B. S. (2024). Development Of The Islamic Religious Education Curriculum Considering Academic And Practical Balance. *Journal of Language and Literature Education*, 1(3), 148–167.
- Firman. (2022). Need Analysis of Indonesian Teaching Materials at Islamic Religious Universities. *Edukasi Islami: Jurnal Pendidikan Islam*, 11(1), 1–9. <https://doi.org/10.30868/ei.v11i01.3781>
- Hajar, A. (2024). Navigating Globalization: Reforming Islamic Education for the 21st Century. *Sinergi International Journal of Islamic Studies*, 2(1), 53–65.
- Kamal, M. Z. (2025). ANALYSIS OF STUDENTS'DIFFICULTIES FOR ENGLISH SUBJECT IN THE ISLAMIC EDUCATION DEPARTMENT AT STIT AQIDAH USYUNI SUMENEP. *An-Nas*, 9(2), 184–195.
- Kamal, M. Z., Albadri, A., & Effendi, A. (2025). Analysis of Speaking Problems for English Subject in Islamic Education Department at STIT Aqidah Usumuni Sumenep: English. *JOEY: Journal of English Ibrahimy*, 4(2), 28–33.
- Khan, N., Sarwar, A., Chen, T. B., & Khan, S. (2022). Connecting digital literacy in higher education to the 21st century workforce. *Knowledge Management & E-Learning*, 14(1), 46–61.
- Latifah, E. N., Malihah, N., & Irwansyah, F. S. (n.d.). Assessment of Pedagogical Competence of Islamic Religious Education (PAI) Teachers: A Literature Review. *ASEAN Journal of Religion, Education, and Society; Vol 4, No 1 (2025): AJORES: Volume 4, Issue 1, April 2025*. <https://ejournal.bumipublikasinusantara.id/index.php/ajores/article/view/683>
- Ludviana, V. (2025). A Contextual and Adaptive Approach to PAI Learning for 21st Century Education at SMPN 1 Kras, Kediri. *Ensiklopedia: Jurnal Pendidikan Dan Inovasi Pembelajaran Saburai*, 5(02), 270–279.
- Muzaki, M., Firdaus, S., & Mustajab, D. (2025). Revitalizing the Value of Tolerance in Islamic Religious Education (PAI) Learning to Face the Challenges of Multiculturalism. *Journal of Educational Sciences*, 9(6), 5804–5813.

- Seargeant, P., & Erling, E. J. (2011). The discourse of 'English as a language for international development': Policy assumptions and practical challenges. *Dreams and Realities*, 248–268.
- Setiawan, R. A. (2026). Rahmatan lil 'Alamin as a Pedagogical Paradigm in Islamic Education: A Literature Review. *Jurnal Harmoni Pendidikan*, 2(1), 81–96.
- Sidqurrahman, Z. (2024). The integration of Islamic perspectives in teaching English as a foreign language: A systematic literature review. *Journal of Social and Economics Research*, 6(2), 223–245.
- Suharsimi, A. (2006). Prosedur penelitian suatu pendekatan praktik. *Jakarta: Rineka Cipta*, 134, 252.
- Swami, N. (2025). THE ROLE OF ENGLISH IN GLOBAL COMMUNICATION: CHALLENGES AND OPPORTUNITIES. *Journal of Advanced Interdisciplinary Studies-p-ISSN 3050-8924 En e-ISSN 3050-8932*, 1(1), 1–9.
- Walliman, N. (2021). *Research methods: The basics*. Routledge.
- Wicagsono, M. A., & Al, B. M. A. M. H. (2023). Strategies for improving teacher pedagogic competence industrial revolution era 4.0. *Multicultural Islamic Education Review*, 1(1), 15–25.
- Zein, K. N., & Arifudin, I. (2025). Analysis of Instructional Needs and General Instructional Objectives in Islamic Religious Education (IRE) Learning: An Integrated Literature Review. *Cristal International Journal of Interdisciplinary Research*, 1(2), 82–90.

Relationship between Vocabulary Knowledge and Inferential Reading Comprehension in Junior High Students

Tarisyia Amelia Indana Zulfa ¹⁾, Eva Rahmawati ²⁾

^{1,2} English Education, Faculty of Languages and Arts, Universitas Negeri Surabaya, Indonesia

Email Correspondence: evarahmawati@unesa.ac.id

ARTICLE INFO

Keywords:

[Vocabulary Knowledge;
Inferential Reading
Comprehension]

Article History:

Received : 21-07-2026

Revised : 13-08-2026

Accepted : 13-08-2026

Published : 20-08-2026

ABSTRACT

Vocabulary knowledge is critical component of reading comprehension, particularly inferential comprehension, which requires readers to interpret implicit meaning beyond explicitly stated information. Although previous studies have confirmed the importance of vocabulary for reading comprehension, limited research had examined how vocabulary knowledge across different word-frequency levels related to inferential comprehension among junior high school students. This study aimed to identify students' level of vocabulary knowledge, describe their inferential reading comprehension, and determine whether a significant relationship existed between the two. A quantitative correlational design was employed with 53 eighth-grade students from a private junior high school in Sidoarjo. Vocabulary knowledge was measured using the Updated Vocabulary Levels Test (1000-5000 levels), and inferential reading comprehension was assessed using an inference-based test adapted from TOEFL Junior materials. Data were analyzed using descriptive statistics and Spearman's rho. Results showed strong vocabulary mastery at the 1000-word level (92.5%), which declined and became more varied from 2000- through 5000-word levels. Inferential reading comprehension yielded a mean score of 57.83, falling into the sufficient category. A significant moderate positive correlation was found between vocabulary knowledge and inferential reading comprehension ($r = .488$, $p = .000$), strongest at the 1000-word ($r = .513$) and 2000-word ($r = .541$) levels. These findings indicated that vocabulary mastery at the basic and intermediate levels served as a primary predictor of inferential reading performance, suggesting that teachers should prioritize structured, frequency-informed vocabulary

instruction paired with explicit inference practice to support students' transitions toward higher-order reading skills.

1. Introduction

Reading comprehension has long been regarded as one of the central competencies in second and foreign language learning, given its direct contribution to learners' ability to acquire information, build knowledge, and participate meaningfully in academic life. For learners of English as a Foreign Language (EFL), reading frequently functions as the primary channel of language input, particularly in settings where opportunities for face-to-face interaction with native speakers are scarce but access to written English is comparatively abundant (Saeedi et al., 2016). Through sustained engagement with texts, learners are exposed not only to vocabulary and grammatical patterns but also to higher-order cognitive operations such as analyzing, synthesizing, and evaluating information (Castles et al., 2018). For this reason, reading comprehension is widely treated as the ultimate goal of reading instruction and is considered one of the strongest indicators of a student's broader academic trajectory.

Reading comprehension itself is not a monolithic skill. Building on Barrett's three-level taxonomy of reading comprehension (as cited in Brassell & Rasinski, 2008), it is commonly described as unfolding across three hierarchical levels: literal comprehension, in which the reader identifies information that is explicitly stated in a text, such as names, places, or events mentioned directly; inferential comprehension, in which the reader draws on background knowledge to uncover meanings and implications that are not directly expressed; and critical comprehension, the most demanding level, in which the reader evaluates a text's validity, quality, and relevance against personal experience and external standards. This three-level structure is echoed in Koda's (2005) account of the components underlying comprehension, which identifies decoding knowledge, vocabulary knowledge, and syntactic knowledge as the building blocks readers rely on to move from surface-level word recognition toward deeper meaning-making. Among the three levels of comprehension, inferential comprehension occupies a particularly important position, since it requires readers to merge textual cues with prior knowledge in order to reconstruct meanings the author has left implicit (Hall et al., 2020). This process draws on several distinct cognitive operations, including predictive inference, in which readers anticipate what will happen next (Graesser et al., 1994); bridging (or connective) inference, in which readers link separate pieces of explicit information into a coherent whole (Singer et al., 1992); elaborative inference, in which readers add detail or context drawn from their own experience (Clinton et al., 2020); and evaluative inference, the most demanding of the four, in which readers form judgments about an author's attitudes, motivations, or unstated message (Cooter et al., 1986). Together, these operations allow readers to build a coherent and complete mental representation of a text rather than a fragmented list of isolated facts (Kispaal, 2008; Noordman & Vonk, 2015). Crucially, each of these inferential operations is constrained by the quality of the reader's underlying lexical knowledge. The Lexical Quality Hypothesis proposed by Perfetti (2007) asserts that the high-quality representation in lexical aspects such as orthography,

phonology, semantics, and syntax allow for easier identification of words and also support higher level comprehension processes such as making inferences. This mechanism directly underlies the inferential operations described earlier. Readers can hold multiple explicit propositions active long enough to connect them into bridging inferences, can access the fuller network of semantic associations a word affords in order to supply elaborative detail, can anticipate upcoming content in predictive inference, and can register the subtler connotative and pragmatic cues that signal an author's attitude in evaluative inference. When lexical representations are low-quality or only partially specified, by contrast, word identification itself becomes effortful, leaving few resources available for the inferential comprehension but as a plausible cognitive mechanism through which vocabulary knowledge enables it.

Of the three levels, inferential comprehension is consistently identified as the most difficult EFL learners to master, largely because it demands more advanced cognitive processing and reasoning than literal comprehension (Hall et al., 2020). A learner may be entirely capable of locating explicit facts in a text yet still struggle to determine what those facts imply. In the Indonesian context, this issue carries particular weight, as the *Kurikulum Merdeka* explicitly foregrounds the development of deep textual understanding and the ability to infer implicit meaning as part of its English learning outcomes for junior high school students (Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi, 2022). However, empirical evidence shows a significant gap between literal and inferential comprehension abilities. Previous research by Rizki Asani (2022) revealed that Indonesian students tend to succeed in literal comprehension tasks, most of them still struggle when the same text requires inferential reasoning, with up to 63% of students failing at the inferential level. This pattern indicates that the transition from literal to inferential comprehension is not an automatic byproduct of general reading ability, but instead depends on learners possessing sufficient linguistic and cognitive resources to support the inferential process.

Among the linguistic factors thought to underlie this transition, vocabulary knowledge has received sustained attention as one of the most influential (Shi, 2023). Vocabulary knowledge is not just about knowing the meaning of words. According to Nation (2022), this also involves understanding the form, meaning, and use of words in various situations. These aspects are generally described under two major dimensions which are vocabulary breadth, referring to the total amount of vocabulary knowledge a learner possesses, and vocabulary depth, referring to the degree of knowledge a learner has about a particular word, including semantics, relationships, and appropriate use (McGregor et al., 2013). This breadth-depth distinction reflects the same lexical-quality mechanism proposed by Perfetti (2007). Breadth determines how many word a reader can bring a well-specified representations to bear on, while depth determines how fully specified any single representation is across its orthographic, phonological, semantic, and syntactic components. Learners with well-developed vocabulary, in both breadth and depth, are therefore better positioned to connect the meanings of words to textual context and prior knowledge, improving their ability to draw inferences from what they read.

However, learning vocabulary is seen as a continuing challenge for EFL students in Indonesia. The *Kurikulum Merdeka* situates the expected English proficiency of junior high school students (Phase D) at the B1 level of the Common European Framework of Reference (CEFR) (Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi, 2022). It is worth noting that the Learning Outcomes document itself defines the B1 target primarily in terms of communicative competencies (e.g., sustaining interaction, conveying main ideas, and making basic inferences from implicit information) rather than a fixed vocabulary count. Independent vocabulary-size research nonetheless offers an approximate benchmark for what a B1-level learner might be expected to know, with estimates in the literature ranging from around 2750 to over 3250-word families depending on the estimation method used (Milton & Alexiou, 2009). This range is used here only as an approximate point of reference and should not be read as an official numerical target embedded in the curriculum document. This goal has not been fully achieved by most students therefore with this limited vocabulary, students' understanding of implied meaning and reading comprehension of information requiring inferential comprehension is limited.

Consequently, vocabulary knowledge and inferential reading comprehension are closely interconnected. Studies by Binder et al. (2017), Ma & Lin (2015), and Gallagher et al. (2019) provide evidence that learners who have a wider and deeper vocabulary acquire and demonstrate greater reading comprehension. In addition, each study has shown that the contribution of vocabulary knowledge to the learning of inferential comprehension needs further research, particularly when applied to learning English in a foreign country, such as Indonesia.

A growing body of research has established that vocabulary knowledge is a strong predictor of reading comprehension in EFL contexts. Studies consistently demonstrate that both vocabulary breadth and depth are significantly associated with learners' reading performance, indicating that students who know more words and understand them more deeply tend to comprehend texts more successfully (Al Qunayeer, 2021; Alsowat, 2022; Dagnaw, 2023; Rabadi, 2023). In addition, vocabulary size has been found to correlate positively with learners' reading and writing achievement, further highlighting the central role of lexical knowledge in language development (Quines, 2023). These findings suggest that strengthening students' vocabulary knowledge is an effective approach to improving reading comprehension.

However, several limitations can be identified in the existing literature. Most previous studies have been conducted at the university level and international settings, with limited attention given to junior high school students, particularly in Indonesia. Moreover, reading comprehension has commonly been treated as a general construct, while specific sub-skills such as inferential reading comprehension remain underexplored. Critically, none of the studies reviewed above (Al Qunayeer, 2021; Alsowat, 2022; Dagnaw, 2023; Quines, 2023; Rabadi, 2023) examined vocabulary knowledge as a graded, multi-level construct in relation to inferential comprehension specifically. Vocabulary was instead treated as a single composite score correlated with general reading comprehension. There have also

been limited studies employing age-appropriate instruments that specifically measure inferential comprehension.

Therefore, this study intends to examine the relationship between vocabulary knowledge and inferential reading comprehension among junior high school students in Indonesia. Unlike prior research, which has generally correlated a single overall vocabulary score with general reading comprehension among university-level or adult EFL learners, this study distinguishes itself by (a) disaggregating vocabulary knowledge into five word-frequency levels (1000-5000) to examine whether the strength of its relationship with inferential comprehension varies across levels, (b) focusing specifically on inferential comprehension rather than reading comprehension as a whole, and (c) situating the investigation within the Indonesian junior high school context under the *Kurikulum Merdeka*, a population and policy setting that remains largely absent from existing literature. This research aims to (1) identify the level of vocabulary knowledge of junior high school students, (2) describe the performance of junior high school students in inferential reading comprehension, (3) determine whether there is a significant relationship between the level of vocabulary knowledge and inferential reading comprehension of junior high school students. The findings of this study are expected to contribute theoretically to a more specific understanding of how vocabulary knowledge, and its distribution across frequency levels, supports inferential reading and pedagogically to inform English teachers about the importance of integrating frequency-informed systematic vocabulary instruction to enhance students' higher-level reading comprehension.

2. Methods

This study employed a quantitative descriptive-correlational design to describe junior high school students' vocabulary knowledge and inferential reading comprehension, and to examine the relationship between the two variables without any manipulation (Ary et al., 2009). The descriptive component was intended to portray students' performance on vocabulary knowledge and inferential reading comprehension, while the correlational component was employed to determine whether a statistically significant relationship existed between the two variables.

The population of this study consisted of all eighth-grade students of a private junior high school in Sidoarjo, distributed across seven classes (three International Class Orientation/ICO classes, one Tahfidz class, and three Excellent classes), totaling 201 students. The ICO classes were prioritized over the Tahfidz and Excellent classes because, within this school's tracking system, ICO students receive a comparatively greater number of weekly English contact hours and are taught using English-medium instructional materials, making them the subgroup most likely to have developed the vocabulary and reading skills targeted by this study. A purposive sampling technique was applied to select participants considered representative of the research objectives (Ary et al., 2009). Since the three ICO classes follow the same intensive English curriculum, they were considered relatively homogenous in terms of English learning experience. Two ICO classes (8A and 8C) were selected based on the English teacher's recommendation, considering relatively

balanced academic performance and intensive exposure to English, as well as class schedule feasibility, since 8A and 8C were the two classes whose schedules aligned with the timeframe available for data collection, so that the remaining ICO class was not included for this reason. The final sample comprised 53 eighth-grade students, which exceeds the minimum sample size recommended for correlational research (Cohen et al., 2007). This sample size, drawn from a relatively homogenous subpopulation and exceeding the recommend minimum, further supports the adequacy of the sample for the statistical aims of this study.

Two instruments were used for data collection, such as the Updated Vocabulary Levels Test (VLT) by Webb et al. (2017) and Inferential Reading Comprehension Test adapted from The Official Guide to the TOEFL Junior® Standard Test (ETS, 2014) and Master TOEFL Junior Intermediate (CEFR Level B1): Reading Comprehension by Hahn (2014). The VLT measures receptive vocabulary knowledge across five frequency levels (1000–5000-word families) using a matching format. Each level consisted of 10 clusters of three definitions and six words (three keys and three distractors). There were 30 definitions at each level. The test was distributed to students through Google Forms.

The Inferential Reading Comprehension Test was adapted from the two sources through the following procedure. From The Official Guide to the TOEFL Junior® Standard Test (ETS, 2014), the researcher screened the 42 multiple-choice items in the reading section and retained only items classified as inference questions, since most items measured literal comprehension, vocabulary, or reference skills rather than inference. This screening yielded five inference items drawn from three reading passages (two items from the first passage, two from the second, and one from the third). To supplement this limited number, 15 additional items were adopted from Unit 6: Inference Questions in Master TOEFL Junior intermediate (CEFR Level B1): Reading Comprehension Hahn (2014). This unit is specifically designed to assess inferential reading, so all items were already inference-based and did not require further selection. Combining the two sources produced a 20-item instrument that pairs standardized TOEFL Junior items with a sufficient number of items specifically targeting inference.

The 20 items were classified according to four established typed of inference. Table 1 presents the blueprint of the Inferential Reading Comprehension instrument, including the descriptions of each inference type and its corresponding item numbers

Table 1. Blueprint of the Inferential Reading Comprehension Instrument

Inference Type	Description	Item Numbers	N
Bridging	Connecting two or more explicitly stated details within the text to draw a local conclusion (Singer et al., 1992)	1, 2, 3, 4, 5, 8, 10, 13, 14, 17, 20	11
Elaborative	Drawing on background knowledge or schema to build a situation model that goes beyond information directly linked in the text (Clinton et al., 2020)	6, 7, 9, 12, 15, 18	6
Evaluative	Judging characters or authors' motivations, emotions, attitudes, or reputations that are not explicitly stated (Cooter et al., 1986)	11, 16, 19	3

Predictive	Anticipating forthcoming events based on cues from an unfolding, sequential narrative (Graesser et al., 1994). Not represented in this instrument (see note below).	–	0
Total			20

As shown in Table 1, none of the 20 items specifically targeted predictive inference. Predictive inference relies on cues from an unfolding, sequential narrative that allow readers to anticipate forthcoming events, a text structure that is largely absent from the short, self-contained expository and narrative passages characteristic of the TOEFL Junior reading format used as the basis for this instrument. In addition, the inferential reading test consisted of 20 multiple-choice items and distributed to students in a paper-based format.

2.1 Validity Test

Valid means that the instrument is suitable or can be used to measure what should be measured (Ary et al., 2009). A valid instrument is one of the requirements to get a valid result of research (Sugiyono, 2021). Both instruments were reviewed by an expert validator, namely a lecturer in the English Education study program at State University of Surabaya, with expertise in language testing and assessment, and was therefore qualified to judge the content relevance and linguistic appropriateness of both instruments for junior high school test-takers. The validity of both instruments was measured using a validation sheet with a five-point Likert scale (1 = poor to 5 = excellent). The validation sheet covered nine indicators grouped under six broader aspects, such as clarity (of individual items and instructions), content accuracy (appropriateness of the item format for measuring the intended construct), relevance (alignment of items with the research objectives), content validity (accuracy of the information the items convey), and language accuracy (comprehensibility and correctness of the language used). The Updated VLT obtained a validity score of 34 out of a maximum of 45, while the inferential reading comprehension test obtained a score of 35 out of 45. Based on these results, both instruments were judged suitable for use without revision, thereby establishing their content validity.

2.2 Reliability Test

Reliability refers to the score test consistency or the stability from the instrument whether stable and consistent or not (Ary et al., 2009). The reliability of both instruments was estimated using Cronbach's Alpha, with a coefficient of .70 regarded as an acceptable threshold for internal consistency (Taber, 2018). As presented in Table 2, all VLT levels and the inferential reading comprehension test yielded Cronbach's Alpha coefficients above .70, indicating that both instruments were reliable for use in this study.

Table 2. Reliability of the Instruments

Instrument	N of Items	Cronbach's Alpha
VLT Level 1000	10	.925
VLT Level 2000	10	.895
VLT Level 3000	10	.888

VLT Level 4000	10	.913
VLT Level 5000	10	.903
Inferential Reading Comprehension Test	20	.807

The collected data were analyzed using SPSS version 26. Both instruments were conducted offline in the classroom during regular English lessons on two separate occasions under the supervision of the researcher and the classroom teacher.

2.3 Descriptive Statistics

Descriptive statistics (minimum, maximum, mean, and standard deviation) were used to summarize students' scores on both variables.

Students' VLT score were analyzed using the raw scores obtained, with each correct answer given a score of 1 and a maximum score of 30 at each level. In addition, students' vocabulary mastery was descriptively classified based on the cut-off scores proposed by Webb et al. (2017), as presented in Table 3. Students were considered to have mastered the 1000, 2000, and 3000 levels if they scored between 29 and 30, whereas mastery of the 4000 and 5000 levels was determined by a score between 24 and 30. This classification was used only to describe the extent of students' vocabulary mastery at each level and did not alter the numerical scores used in the subsequent statistical analysis.

Table 3. Criteria for Vocabulary Level Mastery Based on the Vocabulary Levels Test

Vocabulary Level	Score Range	Category
1000 Level	29-30	Mastered
	0-28	Not Mastered
2000 Level	29-30	Mastered
	0-28	Not Mastered
3000 Level	29-30	Mastered
	0-28	Not Mastered
4000 Level	24-30	Mastered
	0-23	Not Mastered
5000 Level	24-30	Mastered
	0-23	Not Mastered

For the inferential reading comprehension test, each correct answer was given a score of 1 and each incorrect answer a score of 0. To facilitate data interpretation, the raw scores obtained by the students are converted into a 100-point scale using the following formula:

$$Final\ Score = \frac{\sum Raw\ Score}{20} \times 100$$

To determine students' ability level, the mean score was consulted against the five achievement categories proposed by Arikunto (2018), as shown in Table 4.

Table 4. Classification of Students' Inferential Reading Ability Levels

Score Range	Criteria
0 – 20	Very Bad

21 – 40	Bad
41 – 60	Sufficient
61 – 80	Good
81 – 100	Very Good

Prior to correlation analysis, normality and linearity tests were conducted to determine the appropriate statistical technique.

2.4 Normality Test

Normality test was examined using the Kolmogorov-Smirnov test, which is appropriate for a sample size greater than 50 (Mishra et al., 2019). Data are considered normally distributed when the significance value (p) is greater than 0.05 (Pallant, 2020). As shown in Table 5, all variables yielded significance values below 0.05, indicating that the data were not normally distributed. Consequently, a non-parametric statistical technique was required for the correlation analysis.

Table 5. Normality Test Results

	Kolmogorov-Smirnov		
	Statistic	df	Sig.
VLT_1000	.253	53	.000
VLT_2000	.233	53	.000
VLT_3000	.123	53	.043
VLT_4000	.249	53	.000
VLT_5000	.215	53	.000
VLT Total Score	.205	53	.000
Inferential Reading Score	.169	53	.001

a. Lilliefors Significance Correction

2.5 Linearity Test

A formal linearity test was not required for the correlation analysis in this study, as Spearman's rho assesses monotonic rather than strictly linear association (Field, 2018). Nevertheless, a supplementary Test of Linearity through one-way ANOVA was conducted to provide additional descriptive information on the shape of the relationship between each VLT subtest and inferential reading comprehension scores. The linearity assumption is met when the significance value of the Linearity term is below 0.05 and the significance value of Deviation from linearity is above 0.05 (Ghozali, 2018). As shown in Table 6, most VLT levels (1000, 2000, 3000, 4000, 5000, and the total score) showed a significant linearity term with a non-significant deviation from linearity, indicating a predominantly linear pattern, while VLT 3000 showed a significant deviation from linearity ($p = .005$), suggesting a less linear relationship at this level. As Spearman's rho does not require linearity, this finding does not affect the validity of the correlation analysis reported in this study.

Table 6. Linearity Test Results

	Sig. Linear Term	Sig. Deviation from Linearity	Linearity Pattern
VLT 1000	.000	.938	Linear
VLT 2000	.000	.377	Linear
VLT 3000	.000	.005	Not Linear
VLT 4000	.011	.247	Linear
VLT 5000	.007	.111	Linear
VLT Total Score	.001	.295	Linear

2.6 Hypothesis Testing

Since the data were not normally distributed, non-parametric correlation analysis was employed using Spearman's rho. Data are classified as significantly correlated if the correlation coefficient is less than 0.05. on the other hand, the data can be classified as not significantly correlated if the coefficient correlation is greater than 0.05. The strength of correlation was interpreted using the criteria proposed by Cohen et al. (2007), as shown in Table 7.

Table 7. Correlation coefficient interpretation

Correlation Value Ranging	Degree of Correlation Coefficient
0.20 – 0.35	Low
0.35 – 0.65	Moderate
0.65 – 0.85	High
>0.85	Very High

The hypothesis tested in this study are formulated as follows:

1. H_0 (Null Hypothesis): There is no significant relationship between vocabulary knowledge and inferential reading comprehension.
2. H_1 (Alternative Hypothesis): There is a significant relationship between vocabulary knowledge and inferential reading comprehension.

If the significance value (p) obtained from Spearman's rho is less than 0.05, H_0 is rejected and H_1 is accepted. In contrast, if p is greater than 0.05, H_0 is accepted and H_1 is rejected.

3. Result

3.1 Students' Vocabulary Knowledge

This section attempts to answer the first research question i.e. What is the level of vocabulary knowledge of junior high school students. Once the data had been collected, the researcher examined it through descriptive statistical analysis to investigate the mean, minimum, maximum, and standard deviation of students' VLT score. The results are presented in Table 8.

Table 8. Descriptive Statistics of Vocabulary Levels Test

Vocabulary Level	N	Min	Max	Mean	SD
1000 Level	53	18	30	27.68	3.030
2000 Level	53	5	30	23.89	6.393
3000 Level	53	5	30	20.08	8.211
4000 Level	53	1	30	20.34	8.665
5000 Level	53	2	30	21.53	7.934

As shown in Table 8, the highest vocabulary mastery is found at Vocabulary Level 1000 with the mean score of 27.68 and a standard deviation of 3.030. At vocabulary level 2000, the mean score decreased to 23.89 (SD = 6.393). The mean scores for Vocabulary Levels 3000, 4000, and 5000 were recorded between 20.08 and 21.53, with a standard deviation consistently above 7, indicating that students' scores at these levels were more widely spread than the previous levels.

Based on the mastery criteria proposed by Webb et al. (2017), students are classified into the Mastered and Not Mastered categories using different cut-off scores at each level (see Table 3). The distribution of students' vocabulary mastery classification is presented in Table 9.

Table 9. Distribution of Students' Vocabulary Mastery

Vocabulary Level	Category	Frequency (n)	Percentage (%)
1000 Level	Mastered	49	92.5
	Not Mastered	4	7.5
	Total	53	100
2000 Level	Mastered	33	62.3
	Not Mastered	20	37.7
	Total	53	100
3000 Level	Mastered	22	41.5
	Not Mastered	31	58.5
	Total	53	100
4000 Level	Mastered	31	58.5
	Not Mastered	22	41.5
	Total	53	100
5000 Level	Mastered	30	56.6
	Not Mastered	23	43.4
	Total	53	100

As shown in Table 9, 92.5% of students (n = 49) were classified as Mastered at the 1000-word level. This proportion decreased to 62.3% (n = 33) at the 2000-word level and further to 41.5% (n = 22) at the 3000-word level. At the 4000- and 5000-word levels, the proportion of students classified as Mastered rose again to 58.5% (n = 31) and 56.6% (n = 30), respectively. This apparent increase reflects the different mastery cut-off scores applied at each level (a score of 29-30 was required at the 1000-3000 levels, compared with 24-30 at the 4000-5000 levels) rather than an actual improvement in students' vocabulary knowledge.

3.2 Students Inferential Reading Comprehension

In order to answer the second research question, specifically, how is the performance of junior high school students in inferential reading comprehension. The Inferential Reading Comprehension Test was administered to 53 students, and the results were analyzed descriptively, as presented in Table 10.

Table 10. Descriptive Statistics of Inferential Reading Comprehension.

	N	Minimum	Maximum	Mean	Std. deviation
Inferential Reading Comprehension Test	53	15	95	57.83	22.415
Valid N (listwise)	53				

Based on the table above, the minimum score that the students got in the inferential reading test is 15 and the maximum score is 95, with a mean score of 57.83 (SD = 22.415).

Students' inferential reading ability levels are grouped into five achievement categories based on Arikunto (2018) (see Table 4). The frequency distribution of students' ability categories is presented in Table 11.

Table 11. Inferential Reading Ability Categories.

Category	Frequency	Percent	Valid Percent	Cumulative Percent
Very Bad	2	3.8	3.8	3.8
Bad	13	24.5	24.5	28.3
Sufficient	17	32.1	32.1	60.4
Good	9	17.0	17.0	77.4
Very Good	12	22.6	22.6	100.0
Total	53	100.0	100.0	

The data in Table 11 show that the most frequently achieved category by students is the Sufficient category, with a total of 17 students (32.1%). Furthermore, there are 13 students (24.5%) in the Bad category, followed by 12 students (22.6%) in the Very Good category. The remaining 9 students (17.0%) are in the Good category, and the smallest number, 2 students (3.8%) are in the Very Bad category.

3.3 The Relationship between Vocabulary Knowledge and Inferential Reading Comprehension

The final data analysis conducted in this study addresses the third research question; Is there a significant relationship between the level of vocabulary knowledge and inferential reading comprehension of junior high school students. To investigate the correlation between variables, this study employed the Spearman's rho correlation. The results of the correlation analysis are presented below.

Table 12. Correlation between Variables

		1000	2000	3000	4000	5000	Total VLT
Inferential Reading Comprehension	Spearman's rho	.513**	.541**	.424**	.323*	.349*	.488**
	Correlation						
	Sig. (2-tailed)	.000	.000	.002	.018	.010	.000
N		53	53	53	53	53	53
95% CI		[.271, .694]	[.306, .711]	[.135, .654]	[.059, .565]	[.077, .595]	[.248, .692]
**. Correlation is significant at the 0.01 level (2-tailed).							

As shown in Table 12, the correlation between the Total VLT Score and inferential reading comprehension was $r = .488$, 95% CI [.248, .692], $p = .000$. This shows a significant positive relationship between students' vocabulary knowledge and their inferential reading comprehension, classified as a moderate correlation. A more detailed analysis of each vocabulary level revealed varying degrees of correlation strength. At the 1000, 2000, and 3000 levels, the results showed a moderate positive correlation with correlation coefficients of $r = .513$, 95% CI [.271, .694], $p = .000$; $r = .541$, 95% CI [.306, .711], $p = .000$; and $r = .424$, 95% CI [.135, .654], $p = .002$, respectively. These three levels are categorized as a moderate degree of correlation. In contrast, the correlation results at the 4000 and 5000 levels showed a slightly different pattern. At the 4000 level, a positive correlation was found with a coefficient of $r = .323$, 95% CI [.059, .565], $p = .018$, and at the 5000 level, the coefficient was $r = .349$, 95% CI [.077, .595], $p = .010$. The correlation results of these two levels are categorized as a low degree of correlation. Notably, the confidence intervals for the 4000- and 5000-word levels are also the widest and extend closest to zero, indicating less precise estimates than those obtained for the 1000- and 2000-word levels. However, both levels still remain statistically significant, as their p -values are lower than the 0.05 threshold. In short, it can be concluded that all levels of vocabulary knowledge have a significant positive relationship with inferential reading comprehension, with the overall strength of the relationship predominantly falling into the moderate category.

Thus, the alternative hypothesis (H_1) stating that there is a significant relationship between vocabulary knowledge and students' inferential reading ability is accepted, while the null hypothesis (H_0) is rejected.

4. Discussion

4.1 Students' Vocabulary Knowledge

An interesting pattern emerges when the Mastered/Not Mastered classifications are examined across word frequency levels: mastery rates decline steadily from the 1000- to the 3000-word level but then appear to rise again at the 4000- and 5000-word levels. This apparent improvement, however, should not be read as evidence that students' actual vocabulary knowledge increased at these higher, less frequent word levels. Rather, it is largely an artifact of the different cut-off scores used across levels, following Webb et al.

(2017). A stricter cut-off of 29-30 out of 30 was applied at the 1000-3000 levels, while a more lenient cut-off of 24-30 out of 30 was applied at the 4000-5000 levels. A clearer picture emerges when the raw mean scores in Table 8 are considered instead of the Mastered/Not Mastered classification. Students' performance in fact continued to decline gradually across the higher levels, from a mean of 20.08 (SD = 8.211) at the 3000-word level to 20.34 (SD = 8.665) at the 4000-word level and 21.53 (SD = 7.934) at the 5000-word level, with means and standard deviations remaining broadly similar across the three levels. This suggests that students' vocabulary knowledge did not genuinely improve at the 4000- and 5000-word levels. Instead, a larger proportion of students simply crossed the lower threshold set for those levels. This distinction is important to avoid misinterpreting the higher Mastered percentages at the 4000- and 5000-word levels as evidence that students command rarer, lower-frequency vocabulary more successfully than they do the 3000-word level.

Evidence from the declining score reinforces the premise that vocabulary knowledge is a complex process because it involves a multifaced understanding of word forms and their usage in various different contexts (Nation, 2022). The low-test scores observed at levels 4000 and 5000 indicate that although students may have limited vocabulary breadth or the number of words they know, they are still significantly lacking in the aspect of vocabulary depth, which includes a deep understanding of semantic relationships and the appropriate use of words (McGregor et al., 2013). The condition where students find it difficult to achieve the mastered category at 3000 to 5000 levels reflects the principle that the lower the frequency of a word, the greater the demand on students to understand the depth of its meaning in order to achieve full mastery.

This phenomenon is supported with the findings of Alsowat (2022), which showed that students' vocabulary scores tend to decrease as the frequency level of words becomes more less common (from level 1000 to higher levels). Many students in the study also were still lacking in the 95% vocabulary mastery threshold required to understand texts independently. This finding is similar to the data from this study, where students' proficiency at the 3000 and above are still below the optimal category. Furthermore, *Kurikulum Merdeka* stipulates that junior high school students (Phase D) are targeted to reach CEFR level B1, a competency benchmark that independent vocabulary-size research (rather than the curriculum document itself) associates with roughly 2750 to over 3250-word families depending on the estimation used (Milton & Alexiou, 2009). Using the lower end of that range as an approximate reference point, the present results show that only 62.3% of students mastered 2000-word level (41.5% at the stricter 3000-word criterion). This suggests that the curriculum's communicative B1 target has not been fully achieved, and students' vocabulary knowledge is still developing toward the expected standard for junior high school level.

4.2 Students' Inferential Reading Comprehension

The mean inferential reading comprehension score of 57.83%, which fall into the Sufficient category, is comparable to the findings of Al Qunayeer (2021), who noted that

intermediate-level EFL learners generally demonstrate moderate performance in reading comprehension. However, since this specific focuses on inferential comprehension, it provides a clearer overview of why a large proportion of students (28.3%) still fall into the bad and very bad categories. This outcome is also consistent with the research by Alsowat (2022), which identified that many EFL students remain below mastery threshold required to comprehend texts independently. Interestingly, the average performance in the current study shows a slightly better result than the poor category in that previous research.

Furthermore, these results underscore the importance of readers' prior knowledge in constructing meaning from a text. As inferential comprehension requires students to integrate textual clues with their background knowledge to derive implied meanings, the moderate scores suggest that many students may face difficulties in activating their stored knowledge during the reading process (Brassell & Rasinski, 2008). This aligns with the principle that reading is an interactive process where understanding is achieved not only by recognizing words but also by establishing causal relationships and predicting outcomes based on what the reader already knows (Kispal, 2008; Noordman & Vonk, 2015). The sufficient achievement of the students indicates that while they can recognize some implicit details, they still need more linguistic and cognitive resources to consistently achieve the deeper understanding required by the *Kurikulum Merdeka*.

4.3 The Relationship between Vocabulary Knowledge and Inferential Reading Comprehension

The results show a positive, moderate relationship between vocabulary knowledge and inferential reading comprehension, indicating that students with higher vocabulary knowledge scores also tend to have higher inferential reading scores, while those with lower vocabulary knowledge scores tend to show lower inferential reading scores. Since this study employed a correlational design, this pattern indicates an association between the two variables rather than a demonstrated causal effect. It should not be read as evidence that vocabulary knowledge directly produces or determines inferential reading ability. Establishing the direction and mechanism of any causal influence would require further regression-based or experimental research.

Theoretically, this pattern of association is consistent with the Lexical Quality Hypothesis (Perfetti, 2007), which proposes that higher-quality lexical representations are associated with more efficient higher-level comprehension processes, including inference-making. When students hold a stronger vocabulary, they may be able to connect word meanings with textual context more efficiently, which is consistent with a greater capacity to read between the lines. This interpretation aligns with Al Qunayeer (2021) and Quines (2023), who similarly observed that vocabulary mastery at the basic-to-intermediate level tends to accompany stronger reading performance, plausibly because vocabulary provides access to the linguistic cues that support the recognition of abstract or implicit concepts.

A closer look at the level-specific correlation coefficients and their confidence intervals (Table 12) helps explain why vocabulary knowledge at the 1000- and 2000-word levels was more strongly associated with inferential reading comprehension than knowledge at the 3000- to 5000-word levels. This pattern can be interpreted through a lexical threshold perspective, which proposes that learners must first attain a sufficiently high level of mastery over the most frequent, high-utility vocabulary before that knowledge can meaningfully support higher-order comprehension processes such as inference-making (cf. Milton & Alexiou, 2009; Nation, 2022). Because high-frequency words occur far more often in running text and carry a disproportionately large share of the lexical coverage needed to understand a passage (Nation, 2022), strong mastery of the 1000- and 2000-word levels most directly reduces the burden of lexical decoding, potentially freeing cognitive resources for the bridging, elaborative, and evaluative inferential operations (Graesser et al., 1994; Perfetti, 2007; Singer et al., 1992). In this sample, students had already achieved consistently high and comparatively homogenous mastery of the 1000-word level (92.5% Mastered, $SD = 3.03$), which may have allowed this foundational vocabulary knowledge to function as a threshold that differentiates weaker from stronger inferential readers. By contrast, vocabulary at the 3000- to 5000-word levels is encountered far less frequently in everyday input and school-based instruction, so students' knowledge of it was not only lower on average but also considerably more uneven, as reflected in the larger standard deviations reported in Table 8 ($SD > 7$ at the 3000-5000 levels compared with $SD = 3.03$ at the 1000-word level). Because these lower-frequency words are less central to constructing the overall meaning of a passage, individual differences in students' mastery of them appear to matter comparatively less for inferential comprehension once the foundational, high-frequency threshold has been met. This interpretation aligns with the theoretical position that lexical quality at the level of frequently encountered words most directly underlies the inferential operations readers rely on to construct meaning beyond what is explicitly stated (Perfetti, 2007).

The results of this study are broadly similar to the report by Rabadi (2023), which found a moderate correlation among EFL students, in which reading comprehension performance was closely associated with vocabulary performance. However, there is a difference in the strength of the relationship compared to the study by Dagnaw (2023), which reported a much stronger correlation of 0.73. The difference in distribution scores is most likely due to the use of different test instruments, considering the current study specifically focuses on inferential reading, which requires more complex reasoning than general reading. Thus, these findings therefore reinforce the argument that vocabulary mastery is an important correlate of students' ability to interpret texts deeply, consistent with current literacy demands. Because the present design is correlational, however, this association should not be interpreted as evidence that vocabulary mastery directly causes stronger inferential comprehension.

5. Conclusion

According to the research objective, which aimed to investigate students' level of vocabulary knowledge, their performance in inferential reading comprehension, and the relationship between the two, the researcher administered tests to 53 8th graders at a private junior high school in Sidoarjo. This conclusion focuses on the theoretical contribution, practical implications, and directions for future research that follow from them.

Theoretically, the level-specific patterns of correlation found in this study extends the Lexical Quality Hypothesis (Perfetti, 2007) by suggesting that the strength of the vocabulary-inference relationship is not uniform across word-frequency levels; the association was strongest at the 1000- and 2000-word levels and weaker, though still significant, at the 3000-5000-word levels. This graded pattern supports a lexical-threshold interpretation, in which mastery of high-frequency vocabulary functions as a foundational resource that frees cognitive capacity for bridging, elaborative, and evaluative inference-making, while knowledge of lower-frequency words plays a comparatively smaller role once this threshold is met. This contribution also distinguishes the present study from previous correlational research, which has generally treated vocabulary as a single composite score and reading comprehension as an undifferentiated construct, most often among university-level or adults EFL learners. By disaggregating vocabulary knowledge into five frequency levels, isolating inferential comprehension specifically, and situating the investigation within the Indonesian junior high school context under the *Kurikulum Merdeka*, this study extends the existing literature into a population, construct, and level of analytic granularity that earlier correlational studies have not addressed.

From a practical perspective, the level-specific pattern of association suggests that English instruction at the junior high school level should prioritize consolidating students' mastery of high frequency words (1000-2000) vocabulary before extensive attention is given to lower-frequency words, since it is this foundational vocabulary that appears most closely tied to students' capacity to draw inferences. Teachers are encouraged to pair vocabulary instruction with explicit inference practice, for example through guided questioning and contextual-clue analysis embedded directly in reading activities, rather than treating vocabulary and reading-strategy instruction as separate strands.

Because this study employed a correlational design, these findings describe an association rather than a causal mechanism, and this limitation points toward directions for future research. Subsequent studies could employ regression-based or experimental design to test whether targeted vocabulary instruction produces measurable gains in inferential reading comprehension, incorporate direct measures of vocabulary depth alongside the breadth-based VLT used here, and extend the investigation to larger and more diverse samples, including students outside intensive English tracks, to examine whether the lexical-threshold pattern observed here holds more broadly.

5. Acknowledgments

We would like to acknowledge to the lectures for their constructive feedback and insightful suggestions which enabled the completion of this research. Moreover, appreciation is also conveyed to the teacher and all of the participants of this research, who were willing to invest their time and energy to participate in this study.

References

- Al Qunayeer, H. S. (2021). An Investigation of the Relationship between Reading Comprehension, Vocabulary Knowledge, and English Language Proficiency Level of Saudi EFL Learners. *Advances in Language and Literary Studies*, 12(2), 59. <https://doi.org/10.7575/aiac.all.s.v.12n.2.p.59>
- Alsowat, H. H. (2022). The Mediating Roles of Gender and Academic Field in the Relationship between Vocabulary Knowledge and Reading Comprehension of EFL Students. *Universal Journal of Educational Research*, 10(2), 146–159. <https://doi.org/10.13189/ujer.2022.100204>
- Arikunto, S. (2018). *Dasar-Dasar Evaluasi Pendidikan* (3rd ed.). Bumi Aksara.
- Ary, D., Jacobs, L. C., & Sorensen, C. (2009). *Introduction to Research in Education* (8th ed.). Wadsworth Cengage Learning.
- Binder, K. S., Cote, N. G., Lee, C., Bessette, E., & Vu, H. (2017). Beyond breadth: the contributions of vocabulary depth to reading comprehension among skilled readers. *Journal of Research in Reading*, 40(3), 333–343. <https://doi.org/10.1111/1467-9817.12069>
- Brassell, Danny., & Rasinski, T. V. . (2008). *Comprehension that works : taking students beyond ordinary understanding to deep comprehension*. Shell Education.
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars: Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest*, 19(2), 93. <https://doi.org/10.1177/1529100618786959>
- Clinton, V., Taylor, T., Bajpayee, S., Davison, M. L., Carlson, S. E., & Seipel, B. (2020). Inferential comprehension differences between narrative and expository texts: A systematic review and meta-analysis. *Reading and Writing* 33. <https://doi.org/10.1007/s11145-020>
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education* (6, Ed.; 6th ed.). Routledge.
- Cooter, R. B., Joseph, D. G., & Flynt, E. S. (1986). Eliminating the Literal Pursuit of Reading Comprehension. *Journal of Clinical Raeding: Research and Practice*, 2.
- Dagnaw, A. T. (2023). Revisiting the role of breadth and depth of vocabulary knowledge in reading comprehension. *Cogent Education*, 10(1). <https://doi.org/10.1080/2331186X.2023.2217345>

- ETS. (2014). *Official Guide to the TOEFL Junior® Standard Test* (1st ed.). Educational Testing Service.
- Field, Andy. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.
- Gallagher, M. A., Barber, A. T., Beck, J. S., & Buehl, M. M. (2019). Academic Vocabulary: Explicit and Incidental Instruction for Students of Diverse Language Backgrounds. *Reading and Writing Quarterly*, 35(2), 84–102.
<https://doi.org/10.1080/10573569.2018.1510796>
- Ghozali, Imam. (2018). *Aplikasi analisis multivariate dengan program SPSS* (9th ed.). Badan Penerbit Universitas Diponegoro.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing Inferences During Narrative Text Comprehension. In *Psychological Review* (Vol. 101, Number 3).
- Hahn, R. (2014). *Master TOEFL Junior Intermediate (CEFR Level B1): Reading Comprehension*. Nhân Trí Việt.
- Hall, C., Vaughn, S., Barnes, M. A., Stewart, A. A., Austin, C. R., & Roberts, G. (2020). The Effects of Inference Instruction on the Reading Comprehension of English Learners With Reading Comprehension Difficulties. *Remedial and Special Education*, 41(5), 259–270. <https://doi.org/10.1177/0741932518824983>
- Kementerian Pendidikan, K. R. dan T. R. I. (2022). *Capaian Pembelajaran Mata Pelajaran Bahasa Inggris Fase A-Fase F untuk SD/MI/Program Paket A, SMP/MTs/Program Paket B, dan SMA/MA/SMK/MAK/Program Paket C*.
- Kispal, Anne. (2008). *Effective teaching of inference skills for reading : literature review*. Department for Children, Schools and Families.
- Koda, K. (2005). *Insights into Second Language Reading*. Cambridge University Press.
<https://doi.org/10.1017/CBO9781139524841>
- Ma, Y., & Lin, W. (2015). A Study on the Relationship between English Reading Comprehension and English Vocabulary Knowledge. *Education Research International*, 2015, 1–14. <https://doi.org/10.1155/2015/209154>
- McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language and Communication Disorders*, 48(3), 307–319. <https://doi.org/10.1111/1460-6984.12008>
- Milton, J., & Alexiou, T. (2009). *Vocabulary Studies in First and Second Language Acquisition* (B. Richards, M. H. Daller, D. D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller, Eds.). Palgrave Macmillan UK. <https://doi.org/10.1057/9780230242258>

- Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, 22(1), 67–72. https://doi.org/10.4103/aca.ACA_157_18
- Nation, I. S. P. (2022). *Learning Vocabulary in Another Language* (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/9781009093873>
- Noordman, L. G. M., & Vonk, W. (2015). Inferences in Discourse, Psychology of. In *International Encyclopedia of the Social & Behavioral Sciences: Second Edition* (pp. 37–44). Elsevier Inc. <https://doi.org/10.1016/B978-0-08-097086-8.57012-3>
- Pallant, J. (2020). *SPSS Survival Manual*. Routledge. <https://doi.org/10.4324/9781003117452>
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- Quines, Z. M. (2023). Impact of Students' Vocabulary Level to their Reading and Writing Performance. *International Journal of English Language and Linguistics Research*, 11(2), 18–32. <https://doi.org/10.37745/ijellr.13/vol11n21832>
- Rabadi, R. I. (2023). Examining the Role of Breadth and Depth of Vocabulary Knowledge in Reading Comprehension of English Language Learners. *Jordan Journal of Modern Languages and Literatures*, 15(1), 327–345. <https://doi.org/10.47012/jjml.15.1.17>
- Rizki Asani, F. (2022). *Students' Difficulties Analysis in Reading Comprehension at The Second Grade of MTs Nurul Ikhlas*.
- Saeedi, G., Khanmohammadi, A., & Mahmmodi, H. (2016). Reading Comprehension and Its Importance in Learning English as a Second Language. *International Journal of EFL*, 1(1), 61. <https://doi.org/10.21462/ijefl.v1i1.1>
- Shi, Y. (2023). How Inference Making Mediates the Relation Between Vocabulary Knowledge and Reading Comprehension for Chinese EFL Middle School Students. *English Teaching(South Korea)*, 78(3), 85–104. <https://doi.org/10.15858/engtea.78.3.202309.85>
- Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. In *Memory & Cognition* (Vol. 20, Number 5).
- Sugiyono. (2021). *Metode penelitian kuantitatif, kualitatif, dan R&D* (3rd ed.). Alfabeta, CV.
- Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test. *ITL - International Journal of Applied Linguistics*, 168(1), 33–69.
<https://doi.org/10.1075/itl.168.1.02web>

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL - International Journal of Applied Linguistics*, 168(1), 34–70.

Unveiling the Macro-Micro Paradox: A Hermeneutic Phenomenology of Pre-Service English Teacher's Digital Praxis in a Micro-Class of a Rural Indonesian Islamic College

Ayu Meiratnasari¹⁾, Ema Wilianti Dewi²⁾

¹Tadris Bahasa Inggris, Fakultas Tarbiyah dan Ilmu Keguruan, Institut Agama Islam Hasyim Asy'ari Ogan Komering Ilir

²Tadris Bahasa Inggris, Fakultas Tarbiyah dan Ilmu Keguruan, UIN Syekh Nurjati Cirebon

Email Correspondence : ayumeiratnasari@gmail.com

ARTICLE INFO

Keywords:

Communicative Language Teaching
Digital Praxis
Generative AI
Hermeneutic Phenomenology
Language Teacher Identity
Macro-Micro Paradox in
Rural EFL Context

Article History:

Received : 21-07-2026
Revised : 08-08-2026
Accepted : 13-08-2026
Published: 20-08-2026

ABSTRACT

The integration of Generative Artificial Intelligence (GenAI) is transforming English as a Foreign Language (EFL) teacher education, yet empirical accounts examining how teacher candidates navigate digital tools within under-resourced contexts remain limited. This study investigated the lived experiences of pre-service English teachers at a rural, faith-based higher education institution in South Sumatra, Indonesia, as they utilized GenAI platforms during lesson preparation for microteaching. Framed within van Manen's hermeneutic phenomenological tradition, data were gathered through triangulated sources, comprising structured written reflections and GenAI-assisted lesson plan (RPP) artifacts from a naturally bounded micro-cohort. The thematic analysis indicates four interconnected essential structures within this micro-cohort: (1) an epistemological realignment from grammar-heavy instruction toward Communicative Language Teaching (CLT) principles, accompanied by un-proofread artifact glitches in AI-generated lesson designs; (2) the role of GenAI as a conditional preparation scaffold that revealed an internal digital asymmetry during live delivery, where participants bifurcated into digital distribution via mobile messaging and tactile, non-digital workarounds; (3) the Macro-Micro Paradox, wherein optimizing technical lesson blueprints (*techne*) accentuated live classroom frictions (*phronesis*) amid teacher target-language anxieties and learner passivity; and (4) a professional identity reconfiguration from rigid knowledge transmitters toward adaptive, empathetic facilitators. Rather than operating as an all-encompassing equalizer, these findings suggest that GenAI serves as a context-dependent scaffold whose pedagogical utility relies on personal digital readiness, material access, and critical engagement. Ultimately, the study indicates that cultivating practical wisdom

(*phronesis*) and digital empathy remains essential when translating digital lesson preparation into live classroom enactment in under-resourced teacher training settings.

1. Introduction

English as a Foreign Language (EFL) instruction is increasingly shaped by global connectivity and the growing availability of digital tools (Gruenhagen et al., 2024; Laoha et al., 2025). Alongside these developments, pedagogical approaches have increasingly incorporated Communicative Language Teaching (CLT) and student-centred inquiry, both of which place greater emphasis on interaction, collaboration, and learner participation (Adler & Shani, 2026; Tao et al., 2026). These developments also place new demands on pre-service English teachers, who are expected to develop pedagogical flexibility and to design learning experiences that encourage active engagement (Akyıldız, 2026). Micro-learning and other concise instructional formats have likewise attracted attention as teachers respond to changing patterns of digital media use (Kohnke & Moorhouse, 2026). At the same time, initial teacher education cannot be reduced to the acquisition of technical skills, because prospective teachers must also develop a professional identity and the practical judgment needed to respond to classroom situations (Zhang & Darwin, 2025). This issue may be particularly relevant in faith-based institutions, including Islamic colleges, where student teachers negotiate professional expectations alongside the particular institutional and educational contexts in which they learn to teach (Ashari et al., 2023).

Generative Artificial Intelligence (GenAI) has become increasingly visible in teacher education as a tool for supporting lesson preparation and other instructional tasks (Alam et al., 2026; Lo, 2025). The educational literature has identified several possible affordances of GenAI, including conversational role-play, personalized feedback, and support for language-learning activities (Khlaif et al., 2025; Kusuma et al., 2024; Tao et al., 2026). In teacher-candidate workflows, however, its use has often centred on practical tasks such as lesson planning, quiz generation, and the adaptation of teaching materials (Moorhouse et al., 2024; Omidvar & Meihami, 2025). AI-supported simulations may also provide relatively low-risk spaces in which novice teachers can rehearse pedagogical tasks (Gao et al., 2026; Zhuang & Zhang, 2025). Nevertheless, much of the existing literature has examined relatively well-resourced urban, Global North, or East Asian educational settings (Fathi, 2026; Gao et al., 2026; Kohnke et al., 2025; Moorhouse, 2024; Moorhouse et al., 2024; Wei et al., 2026; Wulandari & Purnamaningwulan, 2024). Less is known about how teacher candidates in digitally under-resourced settings incorporate GenAI into their preparation and subsequently translate AI-assisted plans into classroom practice (Brown, 2025; Davin et al., 2025; Li et al., 2026; Tan et al., 2025).

Translating academic preparation into classroom practice in rural contexts can be complicated by material constraints and uneven access to educational resources (Kristjánsson, 2024; Tsao, 2025). Pre-service teachers in rural settings may work with unstable internet connections, limited hardware, and inadequate classroom facilities, creating conditions in which digital resources cannot always be used as originally planned

(Boymau & Sari Hidayati, 2025; Hikmah & Astuti, 2025; Syuhra et al., 2025). Rural EFL learners may also have limited opportunities to encounter English beyond the classroom, while test-oriented and grammar-focused instruction can influence their participation and motivation (Boymau & Sari Hidayati, 2025; Carr, 2023; Hikmah & Astuti, 2025; Mandasari et al., 2025; Ramadhani & Fithriani, 2025; Tutyandari, 2022). These conditions may intersect with novice teachers' concerns about pronunciation, vocabulary, and the management of mixed-ability classrooms (Andriani & Subekti, 2022; Ashari et al., 2023; Kawengian & Subekti, 2023; Tutyandari, 2022), which are also related to teachers' capacity to adapt their instructional decisions (Dongying et al., 2026). Thus, possessing theoretical knowledge of digital tools does not necessarily mean that pre-service teachers can enact those tools smoothly in the unpredictable conditions of live classroom interaction (Klimova et al., 2023) (Klimova et al., 2023).

Despite increasing scholarship on digital integration in Indonesian English language teaching, several empirical, methodological, and conceptual gaps remain. Empirically, some Indonesian studies have focused on urban settings or have examined digital integration through relatively narrow indicators such as readability tools (Basori et al., 2024; Wulandari & Purnamaningwulan, 2024). Qualitative research involving rural pre-service teachers has also tended either to give limited attention to AI tools (Maningo et al., 2026) or to rely on quantitative designs that provide less insight into participants' situated experiences over time (Boonmoh & Kulavichian, 2025; Makrakis, 2026; Sperling et al., 2024). Conceptually, rural teacher preparation is sometimes treated primarily as a contextual condition rather than as a setting in which the relationship between technical preparation and classroom enactment can be examined more closely (Karina & Kastuhandani, 2024; Kusuma et al., 2024).

To engage with this gap, the present study uses the term "Macro-Micro Paradox" as a contextually grounded conceptualization rather than as a new theoretical paradigm. The concept draws on the classical distinction between technical skill (*techne*) and practical wisdom (*phronesis*), with the Macro dimension referring to AI-assisted lesson planning, institutional expectations, and available material conditions, and the Micro dimension referring to moment-to-moment classroom interaction, teacher linguistic anxiety, learner participation, and spontaneous pedagogical adjustment. The concept is used here to examine how technically well-structured preparation may coexist with difficulties that emerge during classroom enactment. Within this framing, GenAI may support aspects of technical preparation while leaving the situated interpersonal and pedagogical demands of teaching to be addressed by the teacher.

To address these issues, this study employs a hermeneutic phenomenological approach to explore the lived experiences of pre-service English teachers at Institut Agama Islam Hasyim Asy'ari Ogan Komering Ilir during microteaching. Specifically, the study addresses two research questions:

- a. How do pre-service English teachers in a rural Islamic college utilize GenAI and non-digital tools during the lesson preparation phase (Macro dimension)?
- b. How do these pre-service teachers navigate the Macro-Micro Paradox when translating AI-assisted lesson designs into live classroom interactions (Micro dimension), and how does this experience shape their pedagogical beliefs and professional identity?

2. Methods

This study employed a hermeneutic phenomenological research design (Van Manen, 2016) to interpret the lived experiences of pre-service English teachers as they incorporated generative artificial intelligence (GenAI) into microteaching. Hermeneutic phenomenology assumes that human experience is interpreted rather than simply observed and therefore recognizes the researcher's pre-understanding (*Vorverständnis*) as part of the interpretive process. This approach was appropriate for examining how novice teachers made sense of the relationship between academic preparation, AI-assisted planning, and classroom enactment in a resource-constrained setting. The analysis therefore focused not only on what participants did, but also on how they understood and interpreted their experiences.

The primary researcher was the course lecturer for the TEFL microteaching from which the participants were recruited. This position provided contextual familiarity with participants' learning trajectories but also created a potential power imbalance between researcher and participants. To reduce the possibility of socially desirable responses, data collection was anonymized and managed independently, and analysis was conducted only after final course grades had been submitted to the institutional registrar. These procedures were intended to separate participants' academic assessment from their participation in the study and to support the credibility of their accounts.

The study was conducted in the English Education (*Tadris Bahasa Inggris*) Study Program at Institut Agama Islam Hasyim Asy'ari Ogan Komering Ilir, a rural, faith-based higher education institution in South Sumatra, Indonesia. The setting has limited material and digital infrastructure. Participants did not generally have access to personal laptops and relied mainly on mobile devices, while the campus had limited access to classroom projectors. As a result, student teachers sometimes shared instructional documents and media through class WhatsApp groups and, when necessary, used physical and tactile materials instead of projected digital presentations.

Rather than using statistical sampling, the study recruited the entire bounded cohort of four pre-service English teachers enrolled in the TEFL Class during the academic year through criterion-based purposive sampling. The cohort was treated as a naturally bounded group rather than as a sample intended for statistical generalization. The small cohort size reflects the local context, in which foreign-language anxiety and other contextual factors may contribute to limited enrolment in English education. All four participants met the inclusion criteria and reported using GenAI platforms, including ChatGPT, to obtain instructional ideas, develop lesson plans (RPPs), and prepare teaching materials during the planning phase. Differences in classroom implementation were observed between participants who used tactile physical media and those who used digital materials through WhatsApp. These differences were considered in relation to participants' device access and classroom infrastructure rather than interpreted as differences in whether they used AI during preparation. The demographic and contextual profile of the participants is presented in Table 1.

Table 1. Demographic, Material, and Contextual Profile of Participants

Participant Code	Gender	Age	Academic Semester	Targeted Grade	Primary Pedagogical Approach	GenAI Planning Tool (Macro)	Live Instructional Tool Employed (Micro)
P1	Male	20	4th Semester	Junior High School	Project-Based Learning (PBL) & Dialogue Creation	Generative AI (ChatGPT) for PBL syntax & rubrics	Conventional Tactile Media / Board & Paper LKPD
P2	Male	20	4th Semester	Junior High School	Communicative Language Teaching (CLT) & Interactive Roleplay	Generative AI (ChatGPT) for scenario & dialogue design	Physical Interactive Media (Cardboard "Expression Dice")
P3	Female	20	4th Semester	Junior High School	Grammar Translation Method (GTM) & Deductive Grammar	Generative AI (ChatGPT) for GTM RPP & rule-based materials	Generative AI Content via Class WhatsApp Group & Whiteboard
P4	Female	20	4th Semester	Junior High School	CLT & Authentic Materials Use	Generative AI (ChatGPT) for PBL/CLT lesson design	Generative AI Tools & Digital Multimedia via WA Group (Video/Music)

To examine participants' experiences beyond self-reported reflections, data were collected from two complementary sources:

- Structured Written Reflections:** Participants completed end-of-microteaching reflective journals guided by open-ended prompts concerning their affective experiences, expectations, pedagogical decision-making, perceived changes in professional identity, and encounters with classroom difficulties.
- Pedagogical Artifacts (GenAI-Assisted Lesson Plans/RPPs):** Official lesson plans/teaching modules, and instructional materials produced by the four participants were collected as physical artifacts. These materials were examined to identify how AI-assisted planning was translated into technical lesson structures (*techne*) during the preparation phase.

Data were analyzed using Van Manen's (2016) four-stage interpretive phenomenological approach: holistic reading, selective reading, detailed line-by-line reading, and thematic clustering, within an ongoing hermeneutic circle. The process involved movement between individual

excerpts and the dataset as a whole, while the researcher's contextual understanding of rural English language teaching was treated as part of the interpretive process.

First, the researchers conducted a holistic reading of all written reflections and RPPs/teaching modules artifacts to develop an overall understanding of each participant's experience. Second, selective reading identified phrases associated with technical preparation (*techne*) and difficulties during classroom enactment (*phronesis*). Third, line-by-line reading was used to identify statements concerning teacher anxiety, adaptations to limited resources, classroom interaction, and developing professional identity. Finally, related statements were clustered into four themes. The interpretation was developed through repeated comparison between participants' AI-assisted lesson plans, the physical characteristics of their teaching artifacts, and their accounts of classroom implementation (Rony et al., 2024).

To support methodological trustworthiness, the study maintained an audit trail and used member checking. After final course grades had been submitted and published, all four participants took part in a member-checking session. Each participant received a summarized matrix of the thematic interpretations, anonymized quotations, and artifact analyses. Participants were invited to assess whether the interpretations represented their experiences and pedagogical realities accurately. All four participants confirmed the interpretations and did not request substantive changes. This procedure was used as one source of support for the credibility of the researchers' interpretations rather than as a claim of objective verification.

3. Result

The hermeneutic phenomenological analysis of participants' written reflections and GenAI-assisted lesson plans (RPPs), following van Manen's (2016) interpretive approach, resulted in four themes. Taken together, the findings describe a movement from changing conceptions of English teaching, through differences between AI-assisted preparation and classroom implementation, to difficulties in live interaction and emerging changes in professional identity. The findings also include variations within the bounded cohort, particularly in the use of digital and non-digital instructional materials. Table 2 summarizes the themes, their interpreted meanings, artifact evidence, and representative participant statements.

Table 2. Matrix of Emergent Themes, Essences, and Empirical Data

Core Themes	Emergent	Thematic (Lived Meaning)	Essence	Physical Evidence (RPPs)	Artifact (GenAI)	Representative Statements	Meaningful
Theme Epistemological Realignment and Techne-Related Artifact Issues	1:	Shifting from grammar memorization to CLT, alongside un-proofread artifact glitches.	static CLT, un-GenAI	• P2: Header/footer labeled "BAHASA INDONESIA SMP FASE D KELAS VII" on an English module. • P3: Unformatted LaTeX code ($\rightarrow$) and Class XI label for Class VII. • P4: Cover typo "PAST TANSE" and SMA Class XI reference for SMP VIII.		<i>"The most important thing I learned... is that learning English is not just about memorizing grammar formulas, but how to make students able to use it to communicate in reality..."</i> (Participant 3) <i>"Here I just realized that learning English is not just about memorizing grammar formulas... We can use authentic materials, for example, watching YouTube, songs..."</i> (Participant 4)	

Theme 2: GenAI Scaffolding and Non-Digital Adaptation: Internal Digital Asymmetry	An internal digital divide where all used AI for planning, but split into digital vs. tactile pathways in execution.	• P1 & P2: Paper LKPD, whiteboard, and physical cardboard "Expression Dice". • P3 & P4: Slide decks and videos distributed via class WhatsApp groups.	"Bikin lesson plan dan nyari ide aktivitas kelas jadi gampang dan cepat karena dibantu AI." (Participant 4) "Kekuatan: mampu menyusun lesson plan yang terstruktur dan adaptif dengan bantuan teknologi." (Participant 3) "Dengan cara dadu expressions... students mengguncang dadu dan mendapat expressions apa ya; itu yang mereka praktikkan..." (Participant 2) "...kurang pemahaman saya terhadap kosakata dan kalimat-kalimat bahasa Inggris lainnya... Bahasa yang kurang dipahami terkadang disertai pelafalan yang kurang tepat..." (Participant 1) "Yang kurang optimalnya... cara menyusun kata saya belum jelas, jadi saya bingung menjelaskannya... banyak yang takut sama grammar." (Participant 4) "Guru zaman sekarang bukan lagi sekadar pemberi materi, tetapi juga harus mampu menjadi fasilitator yang memanfaatkan teknologi AI..." (Participant 3) "...menjadi guru bukan sekadar mentransfer rumus/materi, melainkan bagaimana kita bisa membaca kebutuhan, karakter, serta emosi siswa..." (Participant 1)
Theme 3: The Macro-Micro Paradox: Teacher Anxiety and Learner Passivity	Friction between AI-optimized planning (<i>techne</i>) and live delivery (<i>phronesis</i>) due to mutual linguistic anxieties.	• Highly detailed PBL and GTM lesson plans containing advanced linguistic structures.	
Theme 4: Professional Identity Reconfiguration: Developing Adaptive Facilitation	Reconstructing identity from a rigid knowledge transmitter to an empathetic, tech-adaptive facilitator.	• Inclusion of Social Emotional Learning (SEL) rubrics and empathetic reflection sheets.	

3.1 Theme 1: Epistemological Realignment and *Techne* Artifact Glitches

The participants described a shift from viewing English primarily as a set of grammatical rules toward greater attention to communication and classroom use. Microteaching appeared to provide an opportunity for participants to reconsider how English could be taught through activities that encourage students to use the language in meaningful contexts. Participant 3 described this learning as a movement beyond memorizing grammar formulas toward enabling students to use English for communication:

"Hal terpenting yang saya pelajari dari mata kuliah TEFL adalah bahwa belajar bahasa Inggris bukan hanya menghafal rumus tata bahasa (grammar), tetapi juga tentang bagaimana membuat siswa mampu memakainya untuk berkomunikasi secara nyata..."

(The most important thing I learned from the TEFL course is that learning English is not just about memorizing grammar rules; it is also about helping students use English to communicate meaningfully in real-life situations.)

(Participant 3)

Participant 4 echoed this realignment, noting that authentic digital media (e.g., YouTube videos and songs) generated real-world communicative practice:

"Di sini saya baru ngeh kalau belajar bahasa Inggris itu bukan cuma soal menghafal rumus grammar atau teori di buku. Tetapi bagaimana caranya supaya siswa bisa langsung mempraktikkan bahasa tersebut secara nyata". (That was when I realized that learning English isn't just about memorizing grammar rules or textbook theories. It's about helping students actually use the language in real-life situations)

(Participant 4)

Analysis of the lesson-plan artifacts, however, indicated that the use of GenAI did not always result in materials that were ready for classroom use without further review. Several artifacts contained inconsistencies or formatting and language errors. These observations are important because they show that AI-assisted preparation still required human checking and contextual adaptation. The examples below illustrate the types of issues identified in the collected materials:

- a. Participant 2: A module for interpersonal dialogue retained document headers and footers labelled "BAHASA INDONESIA SMP FASE D KELAS VII", despite being prepared for English instruction.
- b. Participant 3: A Grammar Translation Method (GTM) module contained raw LaTeX coding symbols (`\rightarrow`) in the reading material and labelled the student LKPD as Class XI rather than Class VII.
- c. Participant 4: A PBL-SEL module contained a cover typo ("PAST TANSE") and referred to a Senior High School (SMA XI) textbook for a Junior High School (SMP VIII) class.

3.2 Theme 2: GenAI as an Operative Scaffold vs. Non-Digital Adaptation: Revealing Internal Digital Asymmetry

Course records indicated that all four participants used GenAI, particularly ChatGPT, during lesson preparation to generate ideas, structure activities, and develop assessment rubrics. Participants described the tool as useful for organizing and accelerating aspects of planning. Participant 4 stated:

"Bikin lesson plan dan nyari ide aktivitas kelas jadi gampang dan cepat karena dibantu AI." (With the help of AI, it became much easier and faster to create lesson plans and come up with ideas for classroom activities)

(Participant 4)

Participant 3 similarly connected her perceived planning confidence to AI-assisted customization, identifying her primary strength as being *"mampu menyusun lesson plan yang terstruktur dan adaptif dengan bantuan teknologi"*. (Technology helped me develop lesson plans that were both well-structured and adaptable).

Although all participants used AI during preparation, their classroom implementation differed. This variation suggests an internal asymmetry between AI-assisted planning and the resources available for classroom delivery. Differences in device access and the limited availability of classroom projection facilities were associated with two broad implementation pathways:

- a. Digital Distribution Pathway (P3 & P4): Participants 3 and 4 shared AI-assisted slide decks, reading materials, and video links with students through class WhatsApp groups.
- b. Tactile Non-Digital Pathway (P1 & P2): Participants 1 and 2 relied more heavily on physical materials. Participant 1 used printed LKPDs and board writing, while Participant 2 developed a cardboard "Expression Dice" to support oral interaction.

"Dengan cara dadu expressions, students mengguncang dadu dan mendapat expressions apa ya; itu yang mereka praktikkan dengan pasangannya." (Using the Expression Dice activity, students rolled the die and then practiced the expression indicated by the result with a partner)
(Participant 2)

The difference between these pathways suggests that AI-assisted planning does not necessarily lead to the same form of digital classroom implementation. Within this cohort, the translation of AI-assisted preparation into practice was shaped by available devices, classroom infrastructure, and the instructional choices made by individual participants.

3.3 Theme 3: The Macro-Micro Paradox: Intersecting Teacher Linguistic Anxiety and Learner Passivity

The transition from AI-assisted preparation (Macro dimension) to live classroom delivery (Micro dimension) brought challenges that were less visible during lesson planning. Participants described difficulties related to English vocabulary, pronunciation, public speaking, classroom management, and student participation. These experiences suggest that a technically structured lesson plan does not remove the need for situated judgment during classroom interaction.

"...kurang pemahaman saya terhadap kosakata dan kalimat-kalimat bahasa Inggris lainnya... Bahasa yang kurang dipahami terkadang disertai pelafalan yang kurang tepat..." (...I still had limited vocabulary and struggled to understand some English expressions. When I was unfamiliar with certain words or phrases, I sometimes pronounced them incorrectly as well...)
(Participant 1)

Participant 2 similarly identified vocabulary limitations and limited experience with public speaking as challenges during implementation: *"Kekurangan kosa kata dan belum terbiasanya dalam public speaking... Aspek yang paling menantang... adalah cara*

menyampaikannya...".(One of the challenges was my limited vocabulary and lack of experience with public speaking. The most challenging part was figuring out how to deliver the material effectively...)

Participants also described limited student participation and concerns about grammar. Participant 4 explained that uncertainty in her own delivery coincided with students' limited engagement:

"Yang kurang optimalnya, saat saya praktik mengajar cara menyusun kata saya belum jelas, jadi saya bingung menjelaskannya, dan teori saya kurang seru, jadi muridnya pada diam... Tantangan terbesarnya adalah kurangnya motivasi dan rasa percaya diri murid... Ditambah lagi, banyak yang takut sama grammar." (What didn't go as well as I had hoped was the way I expressed my ideas during the teaching practice. I wasn't always clear when explaining the material, and my explanation of the theory wasn't engaging enough, so the students mostly stayed quiet. The biggest challenge was that many of the students lacked motivation and confidence. On top of that, a lot of them were afraid of making grammar mistakes)
(Participant 4)

Participant 3 reported a related difficulty in classroom management, noting that although AI had supported lesson design, she still found it difficult to encourage active participation among students who were relatively passive. These accounts indicate a distinction between preparing an instructional sequence and responding flexibly to interaction as it unfolds.

3.4 Theme 4: Professional Identity Reconfiguration: Shifting from Knowledge Transmitters to Adaptive Facilitators

The challenges encountered during teaching also appeared to influence how participants understood their developing professional roles. Rather than describing teaching solely as the transmission of subject knowledge, participants increasingly emphasized adaptation, facilitation, and attention to learners' needs. These accounts suggest an emerging professional identity in which technological competence was considered alongside interpersonal and situational responsiveness.

"Guru zaman sekarang bukan lagi sekadar pemberi materi, tetapi juga harus mampu menjadi fasilitator yang memanfaatkan teknologi AI untuk menciptakan pembelajaran yang lebih interaktif, adaptif, dan berpusat pada siswa." (Today's teachers are no longer merely providers of knowledge; they are also expected to serve as facilitators who harness AI technology to create learning experiences that are more interactive, adaptive, and student-centered.)
(Participant 3)

Participant 4 redefined effective teaching as an empathetic, communicative experience prioritizing learner enjoyment over rigid textbook coverage: *"Dari mata kuliah ini saya sadar bahwa menjadi guru Bahasa Inggris zaman sekarang tidak bisa cuma bermodal bisa Bahasa Inggris dan pakai metode ceramah, tetapi... harus adaptif, bisa memanfaatkan AI,*

dan kreatif... esensi ngajar itu tidak semudah memindahkan buku ke dalam pikiran siswa. Tetapi bagaimana agar mereka bisa enjoy dan paham". (This course made me realize that being an English teacher today is about much more than simply knowing English and relying on traditional lecture-based methods. Teachers need to be adaptable, creative, and able to make meaningful use of AI. Teaching is not simply about transferring information from a textbook into students' minds; it is about creating learning experiences that help students enjoy the process, engage with the material, and truly understand what they are learning)

(Participant 4)

For Participant 1, reflective teaching cultivated an identity grounded in situational awareness and emotional connection: *"Pelajaran berharga yang saya dapatkan... adalah bahwa menjadi guru bukan sekadar mentransfer rumus/materi, melainkan bagaimana kita bisa membaca kebutuhan, karakter, serta emosi siswa di kelas dan menempatkan posisi saya sesuai dengan keadaan yang ada."* (One of the most valuable lessons I have learned is that being a teacher is not simply about transferring knowledge or teaching formulas. It is about understanding students' needs, personalities, and emotions in the classroom, and knowing how to adapt my approach to the situation and meet them where they are)

4 Discussion

The hermeneutic phenomenological interpretation offers a contextual account of how pre-service English teachers experienced GenAI-assisted preparation in a resource-constrained rural Islamic higher education setting. The findings are particularly relevant to three interconnected issues: the conditional nature of AI-supported preparation, the relationship between technical lesson design and classroom enactment, and the development of professional identity. Because the study involved a bounded cohort of four participants, these interpretations are situated rather than intended to represent rural teacher education more generally.

4.1 GenAI as a Conditional Scaffold vs Universal Equalizer: Internal Digital Asymmetry

While educational technology literature has increasingly identified Generative AI as a useful tool for supporting administrative tasks and lesson preparation (Alam et al., 2026; Khlaif et al., 2025; Omidvar & Meihami, 2025), the present findings suggest that its usefulness is shaped by the material and contextual conditions in which it is used. Rather than functioning as a universal equalizer that can independently overcome geographical or infrastructural disparities, GenAI appeared in this study as a conditional scaffold whose classroom application depended on participants' access to devices, available institutional infrastructure, and their capacity to adapt AI-generated materials to local conditions. This interpretation is consistent with the broader digital divide literature, which emphasizes that access to digital technologies alone does not necessarily result in equivalent opportunities for educational use (Brown, 2025; Davin et al., 2025; Syuhra et al., 2025).

The variation within the present cohort further illustrates this contextual dimension. Although all four participants reported using ChatGPT during the lesson preparation phase (Macro dimension), their classroom implementation (Micro dimension) followed different pathways. Participants 3 and 4 used digital materials, including slides and videos distributed through WhatsApp, whereas Participants 1 and 2 relied more heavily on non-digital resources such as paper-based LKPDs, whiteboard activities, and the cardboard "Expression Dice." This variation suggests that AI-assisted lesson preparation does not necessarily translate into comparable forms of digital classroom practice. Instead, the transition from preparation to implementation was mediated by the resources available to individual participants and by the material conditions of the classroom. In this sense, the findings point to an internal dimension of digital inequality within the small cohort: differences in access to devices and classroom facilities shaped how AI-assisted preparation could be enacted in practice. Accordingly, the instructional value of GenAI in this context appears to depend not only on the capabilities of the technology but also on teachers' ability to evaluate, adapt, and translate its outputs within the material conditions of their teaching environment.

4.2 Epistemological Realignment and Risk of Uncritical Techne

The participants' accounts suggest a shift from an exclusively grammar-focused view of English teaching toward greater attention to Communicative Language Teaching (CLT) principles. This shift was reflected in their recognition of the importance of authentic communication and meaningful language use, as well as in their use of multimedia resources such as YouTube videos and songs. Such experiences appeared to broaden participants' understanding of how English could be taught through more student-centred and communicative activities. This reflection is consistent with previous research highlighting the potential of authentic digital resources and communicative approaches to support more interactive forms of language learning (Adler & Shani, 2026; Gruenhagen et al., 2024; Kohnke & Moorhouse, 2026; Laoha et al., 2025; Tao et al., 2026; Akyıldız, 2026). In the present context, these experiences provide a situated perspective on how pre-service teachers in a rural, faith-based institution interpreted and incorporated such approaches into their developing teaching practices (Basori et al., 2024; Kusuma et al., 2024).

At the same time, analysis of the physical lesson-plan artifacts revealed several limitations in the use of GenAI-generated materials. These included Indonesian-language headers in an English module (Participant 2), raw LaTeX coding symbols (`\rightarrow`) and grade-level inconsistencies (Participant 3), and typographical errors alongside a mismatch between the cited textbook level and the intended junior high school class (Participant 4). These examples suggest that AI-assisted generation did not eliminate the need for careful proofreading and contextual evaluation. In this sense, GenAI appeared to support the technical organization of lesson materials (techne), while the evaluation of accuracy, relevance, and pedagogical appropriateness still required teacher judgment.

The artifact produced by Participant 3 provides a further illustration of this issue. Although her reflection expressed an orientation toward CLT, the corresponding AI-generated module was structured around the Grammar Translation Method (GTM). Rather than interpreting this discrepancy as evidence that GenAI necessarily reproduces traditional pedagogy, it may be understood as an indication that AI-generated instructional materials can reflect different pedagogical orientations from those intended by the teacher. This discrepancy highlights the importance of critically reviewing and adapting AI-generated content. Thus, the findings suggest that GenAI may facilitate the technical aspects of lesson preparation, but its pedagogical implications depend on how teacher candidates evaluate, modify, and contextualize the materials it produces.

4.3 The Macro-Micro Paradox: Disentangling Techne from Phronesis

The Macro-Micro Paradox is used in this study to describe the tension between technical lesson preparation and situated classroom enactment. Drawing on the distinction between *techne*, understood as technical or rule-based knowledge, and *phronesis*, understood as practical judgment in particular situations (Carr, 2023; Kristjánsson, 2024; Makrakis, 2026; Sperling et al., 2024), the concept helps interpret why a well-structured lesson plan may still require substantial adaptation during teaching. The present findings suggest that AI-assisted preparation can strengthen the technical organization of a lesson while leaving unresolved the interpersonal and contextual demands that arise during live interaction.

During classroom implementation, participants' AI-assisted lesson structures were accompanied by concerns about vocabulary, pronunciation, public speaking, and classroom interaction (Andriani & Subekti, 2022; Ashari et al., 2023; Kawengian & Subekti, 2023; Tutyandari, 2022). Participants also described limited learner participation and concerns about grammar (Boymau & Sari Hidayati, 2025; Hikmah & Astuti, 2025; Mandasari et al., 2025; Ramadhani & Fithriani, 2025). These experiences illustrate a distinction between what can be prepared in advance and what must be negotiated during teaching. Rather than indicating that AI preparation is ineffective, the findings suggest that its usefulness is partial: it can support lesson design, while situated pedagogical judgment remains important when teachers respond to learners and classroom conditions in real time. This interpretation complements research on the relationship between digital adoption and situated professional knowledge (Li et al., 2026) and adds a qualitative perspective to studies that have examined rural AI use through other methodological approaches (Boonmoh & Kulavichian, 2025; Maningo et al., 2026).

4.4 Reconfiguring Language Teacher Identity through Digital Empathy

The participants' reflections also provide insight into the development of their professional self-concepts in a rural, faith-based higher education context. Previous studies have examined changes in Language Teacher Identity (LTI) across different institutional and technological settings, including contexts with greater access to digital resources (Fathi, 2026; Guan et al., 2025; Wei et al., 2026). In the present study, participants described their emerging professional identities in relation to the material

conditions of their teaching practice and their experiences of using technology alongside more accessible instructional resources (Ashari et al., 2023; Zhang & Darvin, 2025). These accounts suggest that professional identity development was closely connected to how participants responded to the practical and interpersonal demands they encountered during classroom teaching.

The classroom difficulties described by participants also appeared to influence how they understood the role of the English teacher. Rather than viewing teaching primarily as the transmission of knowledge, participants increasingly emphasized facilitation, adaptability, and attention to learners' needs. Participant 1, for example, emphasized the importance of recognizing students' needs, characteristics, and emotions, while Participant 4 highlighted the importance of helping students understand and enjoy the learning process. These reflections indicate an emerging understanding of teaching as a relational and situational practice rather than solely as the delivery of subject matter. The participants' experiences therefore suggest that their developing professional identities involved not only subject and technological knowledge but also the capacity to respond to learners and classroom conditions.

This interpretation also offers a more nuanced perspective on the relationship between digital competence and teacher identity. The findings do not suggest that technological proficiency alone determines professional identity development. Instead, within this small cohort, participants associated effective teaching with the ability to combine available technological resources with interpersonal awareness and pedagogical adaptation. As Participant 1 reflected, being a teacher involves understanding students' needs, characteristics, and emotions and adjusting one's position according to the classroom situation. Similarly, Participant 4 emphasized that teaching cannot be reduced to transferring information but involves helping learners engage with and understand the material. These accounts suggest that, in this particular resource-constrained context, the development of a digitally informed teacher identity was intertwined with adaptability, empathy, and situated pedagogical judgment (Alam et al., 2026; Brown, 2025).

5 Conclusion

This hermeneutic phenomenological study examined how four pre-service English teachers in a rural, faith-based higher education institution experienced GenAI-assisted lesson preparation and classroom teaching. The findings suggest a movement in participants' reported understandings from predominantly grammar-focused views of English teaching toward greater attention to communication and authentic language use. GenAI was also described as a useful resource for organizing lesson plans, generating activities, and preparing instructional materials. However, its classroom use was shaped by material and infrastructural conditions, with participants adopting different combinations of digital and non-digital resources. The findings therefore indicate that AI-assisted preparation should be understood as a contextual support rather than as a guarantee of effective classroom implementation.

The study also uses the concept of the "Macro-Micro Paradox" to interpret the relationship between technical lesson preparation (techne) and situated classroom judgment (phronesis). In this bounded context, AI-assisted planning could provide structure and efficiency, but classroom implementation still involved language anxiety, learner participation, classroom management, and adaptation to available resources. Several lesson-plan artifacts also contained errors or inconsistencies, indicating the importance of critical review and contextualization. Taken together, the findings suggest that the development of professional digital competence involves more than access to AI tools. For these participants, it also involved learning to evaluate AI-generated materials, adapt instructional plans to local conditions, and respond to learners during live interaction. These findings are context-specific and should be interpreted as an account of the experiences of this small rural cohort rather than as a basis for broad generalization.

References

- Adler, I., & Shani, Y. (2026). Hybrid Human-GenAI cognitive apprenticeship: Encouraging pre-service teachers to implement instructional practices to support students' self-regulated learning. *Teaching and Teacher Education*, 169, 105268. <https://doi.org/10.1016/j.tate.2025.105268>
- Alam, M. M., Farhaz, S., Haq, S. M. A., & Ferdous, M. (2026). Generative Artificial Intelligence integration in higher education: A constructivist learning theory approach. *Computers and Education Open*, 10, 100378. <https://doi.org/10.1016/j.caeo.2026.100378>
- Andriani, R., & Subekti, A. S. (2022). *Indonesian Pre-Service English Teachers' Teaching Anxiety in Teaching English for Occupational Purposes Classes*.
- Ashari, A., Hamiddin, H., Ainiyah, I., Diana Khalidah, N., Ubaidillah, M. F., & Mac Donald, K. (2023). Pre-Service English Teachers' Anxiety in Speaking English: A Case from an Indonesian Islamic College. *Jurnal Ilmu Pendidikan*, 29(1), 25. <https://doi.org/10.17977/um048v29i1p25-31>
- Basori, B., Wihardini, D., Rofiki, I., & Maharany, E. R. (2024). Investigating pre-service teachers' perceptions on readability tools: A case study in Indonesia. *International Journal of Evaluation and Research in Education (IJERE)*, 13(1), 589. <https://doi.org/10.11591/ijere.v13i1.26333>
- Boonmoh, A., & Kulavichian, I. (2025). Self-directed professional development of Thai pre-service English teachers in rural Thailand: Adapting to limited resources and technology integration. *Australian Journal of Applied Linguistics*, 8(4), 103076. <https://doi.org/10.29140/ajal.v8n4.103076>
- Boymau, L. T. & Sari Hidayati. (2025). Unpacking Rural English Teachers' Perspectives on Digital Learning Tools. *JPI (Jurnal Pendidikan Indonesia)*, 14(1), 101–110. <https://doi.org/10.23887/jpiundiksha.v14i1.84506>
- Brown, J. R. (2025). *Marginalized Pre-Service Teachers' Digital Competence: An Exploration of Epistemic Beliefs and Lived Experiences* [Old Dominion University]. <https://doi.org/10.25777/XT0D-7421>
- Carr, D. (2023). The Practical Wisdom of *Phronesis* in the Education of Purported Virtuous Character. *Educational Theory*, 73(2), 137–152. <https://doi.org/10.1111/edth.12570>
- Davin, K., Kissau, S., Haudeck, H., Ade-Thurrow, B., Gómez-Pereira, D., Oon, P. T., & Dornburg, A. (2025). *Global and Educational Disparities in AI Integration: A Study of*

- L2 Teacher Training and Usage Patterns*. SocArXiv. https://doi.org/10.31235/osf.io/c5u8r_v1
- Dongying, S., Gai, M., & Simayilijiang, A. (2026). Cognitive load and teachers' innovative behavior in AI-enhanced English language instruction: A mediation analysis of technological adaptability. *PLOS One*, 21(3), e0343002. <https://doi.org/10.1371/journal.pone.0343002>
- Fathi, J. (2026). Becoming with the algorithm: A posthumanist narrative of EFL teacher identity in the age of AI. *System*, 140, 104069. <https://doi.org/10.1016/j.system.2026.104069>
- Gao, A., Yu, A., Xu, G., Trainin, G., & Luo, X. (2026). Developing an AI-assisted learning flow framework for pre-service teachers of Chinese as a second language: Investigating the influence of teacher educators' support on AI-flow and self-efficacy in teaching. *System*, 139, 104025. <https://doi.org/10.1016/j.system.2026.104025>
- Gruenhagen, J. H., Sinclair, P. M., Carroll, J.-A., Baker, P. R. A., Wilson, A., & Demant, D. (2024). The rapid rise of generative AI and its implications for academic integrity: Students' perceptions and use of chatbots for assistance with assessments. *Computers and Education: Artificial Intelligence*, 7, 100273. <https://doi.org/10.1016/j.caeai.2024.100273>
- Guan, L., Zhang, Y., & Gu, M. M. (2025). Pre-service teachers preparedness for AI-integrated education: An investigation from perceptions, capabilities, and teachers' identity changes. *Computers and Education: Artificial Intelligence*, 8, 100341. <https://doi.org/10.1016/j.caeai.2024.100341>
- Hikmah, W. N., & Astuti, I. W. (2025). EFL pre-service teachers' challenges in teaching English to young learners during Kampus Mengajar program in rural areas. *English Learning Innovation*, 6(1), 37–48. <https://doi.org/10.22219/englie.v6i1.38340>
- Karina, B. D., & Kastuhandani, F. C. (2024). Pre-Service English Teachers' Lived Experience in Using AI in Teaching Preparation. *Edunesia : Jurnal Ilmiah Pendidikan*, 5(1), 550–568. <https://doi.org/10.51276/edu.v5i1.767>
- Kawengian, G. E., & Subekti, A. S. (2023). Thesis writing anxiety among Indonesian pre-service teachers of English: A survey study. *Englisia: Journal of Language, Education, and Humanities*, 11(1). <https://doi.org/10.22373/ej.v11i1.18940>
- Khlaif, Z. N., Hamamra, B., Bensalem, E., Mitwally, M. A. A., & Sanmugam, M. (2025). Factors Influencing Educators' Technostress While Using Generative AI: A Qualitative Study. *Technology, Knowledge and Learning*. <https://doi.org/10.1007/s10758-025-09885-z>
- Klimova, B., Pikhart, M., Polakova, P., Cerna, M., Yayilgan, S. Y., & Shaikh, S. (2023). A Systematic Review on the Use of Emerging Technologies in Teaching English as an Applied Language at the University Level. *Systems*, 11(1), 42. <https://doi.org/10.3390/systems11010042>
- Kohnke, L., & Moorhouse, B. L. (2026). Principles of GenAI-assisted microlearning: Perspectives of pre-service language teachers. *Elt Journal*, 80(1), 148–159. <https://doi.org/10.1093/elt/ccaf066>
- Kohnke, L., Zou, D., Ou, A. W., & Gu, M. M. (2025). Preparing future educators for AI-enhanced classrooms: Insights into AI literacy and integration. *Computers and Education: Artificial Intelligence*, 8, 100398. <https://doi.org/10.1016/j.caeai.2025.100398>
- Kristjánsson, K. (2024). Aristotelian Practical Wisdom (Phronesis) as the Key to Professional Ethics in Teaching. *Topoi*, 43(3), 1031–1042. <https://doi.org/10.1007/s11245-023-09974-7>

- Kusuma, I. P. I., Roni, M., Dewi, K. S., & Mahendrayana, G. (2024). Revealing the potential of ChatGPT for English language teaching: EFL preservice teachers' teaching practicum experience. *Studies in English Language and Education*, 11(2), 650–670. <https://doi.org/10.24815/siele.v11i2.34748>
- Laoha, R., Chomthong, W., & Pongpanich, W. (2025). Artificial Intelligence and English as a Foreign Language (EFL) Teachers' Competencies: A Systematic Review. *Higher Education Studies*, 15(3), 262. <https://doi.org/10.5539/hes.v15n3p262>
- Li, B., Exter, M., Feng, W., Tang, G., & Xu, K. (2026). A systematic review of language educators' practices and development with GenAI. *International Journal of Artificial Intelligence in Education*, 36(1–2), 100010. <https://doi.org/10.1016/j.ijaiied.2026.100010>
- Lo, A. W. T. (2025). The educational affordances and challenges of generative AI in Global Englishes-oriented materials development and implementation: A critical ecological perspective. *System*, 130, 103610. <https://doi.org/10.1016/j.system.2025.103610>
- Makrakis, V. (2026). Aristotle and AI in Education: Virtue, Wisdom, Human Flourishing and the Common Good. *Encyclopedia*, 6(6), 116. <https://doi.org/10.3390/encyclopedia6060116>
- Mandasari, B., Basthomi, Y., Hastomo, T., Afrianto, Hamzah, I., & Aminatun, D. (2025). The Snapshots of Indonesian Pre-Service English Teachers' Perspectives on Integrating Technology-Based Tools to Rural Schools. *Voices of English Language Education Society*, 9(1), 42–57. <https://doi.org/10.29408/veles.v9i1.27965>
- Maningo, M. P. S., Ochotorena, C., Catalbas, V. B., Empeynado, M., Recamara, J., & Anito, J. (2026). Agency at the margins: How pre-service teachers navigate poverty, care, and community leadership. *Social Sciences & Humanities Open*, 14, 103175. <https://doi.org/10.1016/j.ssaho.2026.103175>
- Moorhouse, B. L. (2024). Beginning and first-year language teachers' readiness for the generative AI age. *Computers and Education: Artificial Intelligence*, 6, 100201. <https://doi.org/10.1016/j.caeai.2024.100201>
- Moorhouse, B. L., Wan, Y., Wu, C., Kohnke, L., Ho, T. Y., & Kwong, T. (2024). Developing language teachers' professional generative AI competence: An intervention study in an initial language teacher education course. *System*, 125, 103399. <https://doi.org/10.1016/j.system.2024.103399>
- Omidvar, S., & Meihami, H. (2025). Exploring the “what” and “how” of opportunities and challenges of AI in EFL teacher education. *Computers and Education: Artificial Intelligence*, 9, 100443. <https://doi.org/10.1016/j.caeai.2025.100443>
- Ramadhani, P. F., & Fithriani, R. (2025). Breaking Barriers: Indonesian EFL Pre-Service Teachers' Challenges and Solutions in TEFL. *Jurnal Ilmiah Global Education*, 6(2), 493–510. <https://doi.org/10.55681/jige.v6i2.3708>
- Rony, M. K. K., Numan, S. Md., Akter, K., Tushar, H., Debnath, M., Johra, F. T., Akter, F., Mondal, S., Das, M., Uddin, M. J., Begum, J., & Parvin, Mst. R. (2024). Nurses' perspectives on privacy and ethical concerns regarding artificial intelligence adoption in healthcare. *Heliyon*, 10(17), e36702. <https://doi.org/10.1016/j.heliyon.2024.e36702>
- Sperling, K., Stenberg, C.-J., McGrath, C., Åkerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: A scoping review. *Computers and Education Open*, 6, 100169. <https://doi.org/10.1016/j.caeo.2024.100169>

- Syuhra, M. N., Chandra, N. E., & Rosalina, E. (2025). Artificial Intelligence in English Language Teaching: A Systematic Literature Review of Tools, Impact, and Challenges. *Voices of English Language Education Society*, 9(1), 193–205. <https://doi.org/10.29408/veles.v9i1.29745>
- Tan, X., Cheng, G., & Ling, M. H. (2025). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education: Artificial Intelligence*, 8, 100355. <https://doi.org/10.1016/j.caeai.2024.100355>
- Tao, L., Song, Y., & Fu, J. (2026). Exploring students' self-regulated learning behavioural patterns and perceptions in an English speaking task within a generative AI-supported immersive VR environment. *Computers & Education*, 242, 105515. <https://doi.org/10.1016/j.compedu.2025.105515>
- Tsao, J. (2025). Trajectories of AI policy in higher education: Interpretations, discourses, and enactments of students and teachers. *Computers and Education: Artificial Intelligence*, 9, 100496. <https://doi.org/10.1016/j.caeai.2025.100496>
- Tümen Akyıldız, S. (2026). Prompt engineering and creative pedagogical thinking: Insights from a qualitative study with teacher candidates. *Thinking Skills and Creativity*, 60, 102105. <https://doi.org/10.1016/j.tsc.2025.102105>
- Tutyandari, C. (2022). English Language Pre-Service Teachers' Sense Of Preparedness For Teaching: An Indonesian Case. *Teflin Journal - A Publication on the Teaching and Learning of English*, 33(2), 367. <https://doi.org/10.15639/teflinjournal.v33i2/367-385>
- Van Manen, M. (2016). *Researching lived experience: Human science for an action sensitive pedagogy* (Second edition). Routledge. <https://doi.org/10.4324/9781315421056>
- Wei, X., Yang, J., & Yuan, R. (2026). Evolving EFL teachers' professional identities in AI-mediated teaching: A longitudinal narrative inquiry. *Teaching and Teacher Education*, 181, 105630. <https://doi.org/10.1016/j.tate.2026.105630>
- Wulandari, M., & Purnamaningwulan, R. A. (2024). Exploring Indonesian Efl Pre-Service Teachers' Experiences In Ai-Assisted Teaching Practicum: Benefits And Drawbacks. *Llt Journal: A Journal on Language and Language Teaching*, 27(2), 878–894. <https://doi.org/10.24071/llt.v27i2.8690>
- Zhang, Y., & Darvin, R. (2025). Negotiating gender ideologies and investing in teacher identities: The motivation and investment of EFL pre-service teachers. *System*, 133, 103669. <https://doi.org/10.1016/j.system.2025.103669>
- Zhuang, Y., & Zhang, S. (2025). Integrating ChatGPT in Mathematics Teacher Education: AI-Based Simulation Role-Playing To Support Practice-based Teaching. *International Journal of Artificial Intelligence in Education*, 35(6), 3873–3895. <https://doi.org/10.1007/s40593-025-00519-0>

Bilingual Curriculum in EFL Speaking: Teachers' Strategies at Islamic Boarding School

Silmi Kaffah¹⁾, Wahid Hasim²⁾, Anne Ratna Suminar³⁾

¹²³English Education Program, Faculty of Social Sciences, Languages, and Literature Education, Institut Pendidikan Indonesia Garut, Indonesia

Email Correspondence: Silmikf2003@gmail.com

ARTICLE INFO

*Keywords: Keywords:
[Bilingual Curriculum,
EFL Speaking Class,
Teachers' Strategies,
Islamic Boarding
School]*

Article History:

Received : 21-07-2026
Revised : 13-08-2026
Accepted : 15-08-2026
Published: 20-08-2026

ABSTRACT

English language proficiency was one of the most important skills in education, prompting many schools to implement bilingual curricula to improve students' English skills, particularly their speaking skills. Although the implementation of bilingual curricula had become increasingly common, understanding of the strategies teachers used to implement them in English as a Foreign Language (EFL) speaking classes especially in Islamic boarding schools remained limited. This study employed a qualitative approach using a case study design involving three English teachers at an Islamic boarding school in Garut Regency, West Java. Data were collected through semi-structured interviews and non-participant observation and were analyzed using the interactive analysis model by Miles et al. (2014), which consisted of data condensation, data presentation, and drawing and verifying conclusions. The results indicated that the teachers implemented three main strategies in delivering the bilingual curriculum: instructional strategies, psychological support strategies, and classroom management strategies. These strategies were manifested through eight specific features, including combining direct and indirect teaching approaches, pairing students based on proficiency levels, providing spontaneous praise, creating a psychologically safe classroom environment, and consistently using English as a model for instruction. These findings provided a more comprehensive understanding of the strategies teachers used to implement the bilingual curriculum in EFL speaking instruction in Islamic boarding schools.

1. Introduction

English proficiency has become one of the essential requirements in 21st-century global education. In the Indonesian context, English is learned as a foreign language (English as a Foreign Language/EFL), encompassing four core skills: listening, speaking, reading, and writing (Husna et al., 2023; Quimosing, 2022). Among these skills, speaking occupies a particularly important position because it directly reflects students' ability to use the language for authentic communication (Brown, 2015). Speaking proficiency is determined not only by mastery of vocabulary, grammar, and pronunciation, but is also shaped by psychological factors such as confidence, motivation, anxiety, and the courage to express ideas (Osama et al., 2026).

This growing demand for English proficiency has also prompted Islamic educational institutions, including Islamic boarding schools and integrated Islamic schools, to combine religious education with bilingual curricula (Amalia & Marzuki, 2023). In practice, English is used not only as a subject taught in class but also as a medium of daily interaction, through policies such as the English Area, routine conversational practice, the use of English in academic activities, and intensive teacher–student interaction within the boarding environment (Hasyraf et al., 2026; Silviyanti et al., 2024). This suggests that the bilingual curriculum in boarding functions not merely as a classroom instructional tool, but also as part of a broader effort to build a language-rich ecosystem.

Nevertheless, the implementation of bilingual curricula in Boarding school settings faces distinct challenges. Boarding school are characterized by an educational model that integrates general education, religious instruction, and boarding-school life, meaning that English instruction must compete with the dominance of Arabic and religious studies. Many students also still have limited foundational English proficiency and are prone to anxiety when required to speak, particularly due to fear of making mistakes or being negatively judged by teachers or peers (Habibi et al., 2018; Wang et al., 2022). Such conditions may lead students to become passive, avoid English communication, or revert to their first language during speaking activities (Surotun et al., 2022).

Under these circumstances, the success of a bilingual curriculum depends not only on institutional policy but also on teachers' ability to translate curricular demands into everyday instructional practice. Teachers play a central role in selecting methods, designing speaking activities, managing student interaction, providing feedback, and creating a safe and supportive learning atmosphere particularly when confronted with students' varying proficiency levels, speaking anxiety, limited vocabulary, and tendency to rely on their first language (Marzulina, 2022). Teachers' strategies are therefore an important area of inquiry, as they serve as the link between the formal goals of the bilingual curriculum and the actual conditions students face in boarding.

Several prior studies have addressed issues related to this theme, as summarized in Table 1.

Table 1. Summary of Relevant Previous Studies

Research (year)	Focus	Method	Context	Research Gap
--------------------	-------	--------	---------	--------------

Silviyanti et al. (2024)	Daily practice of English and Arabic among students and teachers in an Islamic boarding school	Qualitative	Islamic bilingual middle boarding school, Aceh	Examined bilingual language practices broadly; did not focus on the specific pedagogical strategies teachers use in EFL speaking instruction
Fitria (2023)	Strategies and challenges of integrating English Language Teaching into Islamic Boarding Schools	Qualitative	Bilingual Schools	It has not yet specifically outlined the concrete strategies teachers use to address these implementation challenges.
Wang et al. (2022)	Psychological dimensions (anxiety, motivation) in speaking	Quantitative	General context	EFL Has not yet explored how teachers respond to and manage students' psychological factors in speaking instruction.
Surotun et al. (2022)	Teachers' challenges in implementing English instruction	Qualitative	EFL School	Has not specifically examined teachers' strategies for overcoming these challenges in the context of Islamic boarding schools.
Hasyraf et al. (2026)	The "English Area" policy and teacher-student interactions	Qualitative	Islamic Boarding Schools	Has not elaborated in detail on teachers' instructional strategies in EFL speaking instruction.

Based on a review of previous studies presented in Table 1, a research gap can be identified that justifies the need for the present study. Previous studies have examined the use of English in Islamic boarding schools, the challenges of implementing bilingual curricula, psychological factors affecting speaking skills, and the obstacles teachers encounter in using English. However, limited attention has been given to the concrete strategies teachers use to implement a bilingual curriculum in EFL speaking instruction in Islamic boarding schools, particularly in addressing students' varying levels of initial English proficiency, speaking anxiety, and tendency to rely on their first language. This gap provides a basis for the present study to examine in depth how teachers implement a bilingual curriculum in EFL speaking instruction and how they adapt their strategies to students' actual learning conditions. The novelty of this study lies in its explicit focus on three interconnected dimensions of teacher strategies; instructional strategies, psychological support strategies, and classroom management strategies in implementing a bilingual curriculum in EFL speaking classes within a boarding setting. The study also examines how these strategies are adapted to support students' participation and English use despite differences in proficiency, speaking anxiety, and reliance on their first language.

2. Methods

This study employed a qualitative approach using a case study design to examine the phenomenon in depth within its natural context Yin (2018). This design was chosen to gain an understanding of the participants' experiences in the context of their daily lives

Creswell (2014), particularly regarding teachers' strategies in implementing a bilingual curriculum for English as a Foreign Language (EFL) speaking instruction in a boarding setting.

The research was conducted at a boarding school in Garut Regency, West Java, which was selected because it consistently implemented a bilingual curriculum. English was not only taught as a subject but was also used as a medium of daily communication between teachers and students. The participants consisted of three English teachers (T1, T2, and T3) who were selected through purposive sampling based on their active involvement in EFL speaking instruction and the implementation of the bilingual system at the boarding school. All three participants had a background in English language education and relevant teaching experience.

Data were collected through semi-structured interviews and non-participant observation. Interviews were conducted individually in a single session for each participant, lasting 20–30 minutes, using 14 questions covering teaching strategies, motivation, and classroom management in a bilingual context. All interviews were recorded with the participants' consent and transcribed verbatim. Observations were conducted once in each classroom, lasting approximately 60 minutes, and were continued in the dormitory setting to observe the use of English outside the classroom. The observed aspects included teacher instruction, the use of code-switching, teacher responses to student anxiety, and English-language interactions among students. Observation results were recorded in field notes as supporting data for the interviews.

The researcher served as the primary instrument in the data collection and analysis process. During observations, the researcher acted as a non-participant observer without intervening in either classroom learning activities or dormitory activities. This approach was adopted to ensure that interactions occurred naturally and to minimize the researcher's influence on participant behavior. Self-reflection was also practiced throughout the research process to reduce potential interpretive bias.

Data were analyzed using the interactive model proposed by Miles et al. (2014), which included data condensation, data display, and conclusion drawing and verification. The validity of the findings was strengthened through source triangulation by comparing interview results across participants with observational findings. In addition, member checking was conducted by having participants review and confirm the interview transcripts to ensure that the data aligned with their experiences and intended meanings.

3. Result

This study was conducted in three EFL Speaking classes taught by T1, T2, and T3. The results of the observations and interviews showed that the three teachers implemented three main strategies in delivering the bilingual curriculum: learning strategies, psychological support strategies, and classroom management strategies. These strategies were reflected in eight strategic features, each of which had distinct characteristics and implementation techniques across the classes. The data also showed that the students did not yet consistently use English independently during speaking activities. They still

needed encouragement from their teachers and exhibited shyness and anxiety about making mistakes when speaking. A summary of these findings is presented in Table 2 below.

Table. 2 Summary of Teachers' Strategies in Implementing a Bilingual Curriculum in the EFL Speaking Class

NO	Teachers' Strategies	Strategy Features	Characteristic	Representative Quotes
1	Learning Strategies	Direct Approach	1) Materials delivered sequentially 2) Clear objectives stated at the outset	"I usually begin the lesson with a short greeting, then I state the learning objectives so students understand the competencies they need to achieve." (T1)
		Indirect/Communicative Approach	1) Warm-up and elicitation activities 2) Ability-based peer pairing 3) Communicative activities (dialogue, role play, discussion)	"I usually start with something close to students' daily lives first. I also often use short videos or audio as a trigger before moving into the core material." (T1) "I prefer to draw students in first through apperception, In my experience, this approach is far more effective for boarding students." (T2) "I usually pair students who are already proficient with those who are still less proficient, so the less proficient ones can learn while building confidence." (T3)
2	Psychological Support Strategies	Fostering Personal Meaning	1) Relating the material to real-life situations. 2) Emphasizing the communicative value of English.	"I instill in students an understanding of the importance of speaking skills gradually and continuously (pause) all four skills are equally important to master." (T1) "I often give examples of situations close to their experience, such as watching English-language films or videos." (T2) "I always relate the material to their daily lives, so students realize English is easy and can be used directly, not merely theory in the classroom." (T3)
		Affective Reinforcement	1) Providing spontaneous praise 2) Providing praise without prior promise 3) Providing praise for tangible achievements	"Appreciation is not promised at the beginning as a reward. It is given spontaneously when a student successfully demonstrates good ability." (T1) "I sometimes give extra points, reward stickers, or other simple forms of appreciation." (T2) "I give verbal appreciation, for

				example by saying 'Masya Allah,' 'Good job', on some occasions, I also give small snacks." (T3)
	Psychologically Safe Atmosphere	1) Mistakes normalized as part of learning 2) Personal approach used with low-confidence students		"I try to build students' confidence gradually so they dare to try." (T1) "I always explain that making mistakes is natural in language learning (pause) so students feel comfortable trying to speak without fear of being mocked." (T2) "If everyone here were already smart, I'm sure you wouldn't be here, it is precisely through mistakes that we know what needs to be improved." (T3)
3	Classroom Management Strategies	Procedure & Routines 1) Consistent routines 2) Behavioral rules	opening	"I usually begin the lesson with a short greeting such as 'good morning,' 'how are you today.'" (T1) "I habituate students to use basic expressions such as Yes, No, thank you, Excuse me, I think." (T2) "The rule is simple but I apply it consistently: when a peer is presenting, no one may make noise, laugh, or fail to pay attention." (T3)
	Group Management & Turn-Taking	1) Facilitating inclusive participation 2) The use of varied grouping methods was observed to support students' participation and provide opportunities for more consistent English use during classroom activities.		"I usually first explain the assessment rubric that will be used, it does not only focus on the final outcome but also includes participation." (T1) "I usually arrange groups in rotation and give every student an equal opportunity to participate." (T2) "Quiet or less confident students are paired with more proficient ones, so they can learn directly from their peers." (T3)
	Time Management & Teacher-Student Relationship	1) Flexible adjustment of activities 2) Consistent post-lesson reflection		"I make adjustments by changing an activity originally intended for all students into one carried out by just a few student representatives." (T1) "When students appear tired, unfocused, or unenthusiastic. I replace the planned activity with a more engaging one." (T2) "When students appear tired or unfocused, I replace the main activity with a lighter, more engaging one such as language games or songs." (T3)

4. Discussion

The findings of this study show that the implementation of the bilingual curriculum in EFL speaking instruction at the boarding does not rely on a single strategy, but rather involves a set of interrelated strategies encompassing learning strategies, psychological support, and classroom management. These findings confirm that students' willingness to speak is related not only to linguistic ability but is also strongly influenced by the quality of interaction and the emotional comfort teachers build. This deepens the perspectives of Asty et al. (2025) and Lintang Sari & Kurniawati (2024), positioning teacher support as a crucial foundation for students in boarding-school environments to begin expressing themselves.

Within the psychological support strategy, the teacher strategies observed focused on fostering personal meaning and reinforcing affective engagement. Through a policy of normalizing mistakes and appreciating students' courage to speak, teachers sought to lower the psychological barriers students often experience when using a foreign language. This aligns with Krashen (1982) Affective Filter theory, in which the teacher acts as a provider of a sense of safety so that student anxiety does not hinder language acquisition. These findings reinforce Khan (2023) argument that a non-judgmental atmosphere gives students room to express themselves. In particular, teachers' verbal appreciation is seen as an instrument for building students' confidence amid their concerns about social judgment.

In the learning strategy, teachers linked the material to students' daily dormitory routines and religious activities. This strategy helped students connect abstract language concepts to real situations they encounter in their daily lives at the boarding school. As facilitators, teachers provided scaffolding through step-by-step explanations and the use of communicative media to accommodate students' varying starting abilities. This approach aligns with Ly (2024), who emphasizes the importance of teacher flexibility in facilitating students' learning experiences according to their needs. Fitria (2023) similarly highlights the importance of flexible learning strategies in addressing differences in students' readiness and language ability. In the Boarding school densely scheduled environment, teachers applied practical, contextual strategies by linking English use to dormitory routines and students' religious activities. This approach is relevant to students' psychosocial conditions, as it allows them to practice English in situations close to their daily lives.

Furthermore, classroom management was carried out through consistent teacher role modeling both inside and outside the classroom. Based on Bandura (1977) Social Learning Theory, the teacher's presence as an active language user serves as a key stimulus for students to imitate communication. Rayes & Albelaihi (2023) also emphasize that this modeling is essential in boarding-school environments that prioritize habituation. In addition, the strategy of distributing speaking turns evenly guaranteed participation for passive students, ensuring that communicative space is not dominated by particular individuals but becomes a collective right within the classroom.

However, this study also noted a degree of limited language autonomy, in which student participation was generally responsive to teacher instruction rather than self-initiated. This phenomenon relates not only to language mastery but reflects the Boarding school strong sociocultural dimension. As examined by Uktolseja (2026), the deep obedience, respect, and tawadhu (humility) students hold toward teachers tend to make them wait for direction rather than take their own initiative. This reality offers a new perspective that a highly hierarchical-religious school culture is an important context that can shape the dynamics of student engagement.

Overall, EFL speaking instruction at the Boarding school is related not only to linguistic mastery but is also shaped by students' psychological conditions, teaching strategies, classroom management, and the dormitory communication environment. The novelty of this study lies in depicting three integrated aspects: contextual learning, psychological support, and classroom management. Accordingly, the implementation of the bilingual curriculum needs to be adapted to the characteristics of boarding-school life and students' need to develop into more autonomous language users.

5. Conclusion

This study concludes that teachers implemented the bilingual curriculum in EFL Speaking instruction through three interrelated strategies: instructional strategies, psychological support strategies, and classroom management through language modeling. The instructional strategies involved scaffolding and ability-based grouping, while the psychological support strategies included allowing students to make mistakes and providing encouragement to reduce anxiety and affective barriers. Classroom management was reflected in teachers' consistent use of English in daily classroom communication, which provided students with opportunities to encounter and use English in meaningful contexts. The findings further showed that the integration of these strategies helped create a safer and more supportive learning environment, enabling students to respond orally with greater confidence when prompted or encouraged by their teachers. However, students still showed limited initiative in communicating independently without external prompts. This finding suggests that although teacher support facilitated students' willingness to speak, greater attention is needed to foster learner autonomy in initiating English communication.

The main contribution of this study lies in proposing an integrative understanding of teacher strategies that connects the implementation of a bilingual curriculum with the psychosocial conditions of EFL speaking instruction in an Islamic boarding school context, where English coexists with Arabic and students' native languages. Practically, the findings suggest that teachers in similar contexts should prioritize students' psychological comfort and opportunities for meaningful language use alongside linguistic accuracy. This study was limited by its small number of participants and its focus on a single Islamic boarding school, which limits the transferability of the findings to other Boarding Schools. In addition, the study relied primarily on teachers' perspectives and therefore did not comprehensively capture students' experiences. Future research should involve more participants and multiple institutions and incorporate students'

perspectives, either through a phenomenological approach or a mixed-methods design, to provide a broader understanding of teacher strategies and their effectiveness in supporting sustained English Speaking development.

6. Acknowledgments

The researcher would like to express sincere gratitude to the Institute of Education Indonesia (IPI) Garut, particularly the English Education Study Program, for its support throughout this research. The researcher also extends sincere appreciation to the English teachers at Al-Misykah Islamic Boarding School for their participation in this study, as well as to the supervisors, family, and friends for their guidance, support, and encouragement.

References

- Amalia, I., & Marzuki, A. G. (2023). *The Implementation of Bilingual Class Program in Islamic Elementary School*. 30(1), 74–84. <https://doi.org/10.15548/jt.v30i1.738>
- Asty, N. A., Sawaluddin, Sabilla, D., & Nasution, S. (2025). Analisis Penerapan Metode Berbasis Tugas dalam Meningkatkan Kompetensi Berbahasa Asing di Sekolah Menengah: Perspektif Linguistik Terapan. *Pediaqu : Jurnal Pendidikan Sosial Dan Humaniora*, 4(2).
- Bandura, A. (1977). *Social Learning Theory*.
- Brown, H. Douglas. (2015). *Teaching by Principles : an Interactive Approach to Language Pedagogy*. Pearson Longman.
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (fourth edition).
- Fitria, T. N. (2023). Integrating English Language Teaching (ELT) Into Islamic Boarding Schools: A Review of Strategy and Challenges. *Journal of English Language and Pedagogy (JELPA)*, 1(2), 64–78. <https://doi.org/10.51826/jelpa.v1i2.772>
- Habibi, A., Mukminin, A., Najwan, J., Haswindy, S., & Marzulina, L. (2018). *Investigating EFL Classroom Management in Pesantren: A Case Study*. 23(9), 9–17. <https://nsuworks.nova.edu/tqr/vol23/iss9/6>
- Hasyraf, M. F., Aryasatya, M. K., Nurfaiz Lubis, M. D., Simarmata, R. I., Arrowi Pane, M. F., & Azmi, M. N. (2026). The Role of Daily English Conversation in Developing Students' English Communication Skills in a Boarding School Environment. *Jurnal Pengabdian Masyarakat Dan Riset Pendidikan*, 4(3), 19144–19149. <https://doi.org/10.31004/jerkin.v4i3.5264>
- Husna, A. H., |43, H., & Hindriyastuti, S. (2023). Need Analysis on therapeutic Communication of Palliative Nursing in English for Specific Purposes (ESP). *Journal of English Teaching and Research*, 8(1).
- Khan, S. S. (2023). Reflections of Krashen's Affective Filter Hypothesis in Language Classroom: A Study amongst Undergraduates. *British Journal of Arts and Humanities*, 142–149. <https://doi.org/10.34104/bjah.02301420149>
- Krashen, S. D. . (1982). *Principles and practice in second language acquisition*. Pergamon.

- Lintangsari, A., & Kurniawati, A. (2024). Reducing Students' Anxiety in Speaking English: Teacher's Strategies. *Journal of Nusantara Education*, 4(1), 70–77. <http://journal.unu-jogja.ac.id/fip/index.php/JONED>
- Ly, C. K. (2024). Teachers' Roles on English Language Teaching for Promoting Learner-Centered Language Learning: A Theoretical Review. *International Journal of TESOL & Education*, 4(2), 78–98. <https://doi.org/10.54855/ijte.24425>
- Marzulina, L. (2022). Challenges in Teaching English for EFL Learners at Pesantren: Teachers' Voices. *Academi Publication*.
- Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis* (3rd ed.).
- Osama, Aslan, & Effiyadi. (2026). Implementasi Guru Dalam Mengatasi Kesulitan Keterampilan Berbicara Siswa Pada Pembelajaran Bahasa Indonesia Kelas Ix D Di Mts Muhammad Basiuni Imran Sambas. *ADIBA: JOURNAL OF EDUCATION*, 5(3), 97–109.
- Quimosing, A. B. (2022). Learning English as A Foreign Language (EFL): A Narratology. *SALTeL Journal (Southeast Asia Language Teaching and Learning)*, 5(2), 13–21. <https://doi.org/10.35307/saltel.v5i2.85>
- Rayes, E. T., & Albelaihi, N. K. (2023). Investigating Bandura's Processes of Observational Learning Implementations from EFL Faculty Perspective at Umm Al-Qura University English Language Centre. *Arab World English Journal*, 14(3), 28–49. <https://doi.org/10.24093/awej/vol14no3.3>
- Silviyanti, T. M., Yusuf, Y. Q., Samad, I. A., & Mukminin, A. (2024). The implementation of English and Arabic practices at an Islamic bilingual boarding school. *Cakrawala Pendidikan: Jurnal Ilmu Pendidikan*, 43(3), 630–644. <https://doi.org/10.21831/cp.v43i3.72204>
- Surotun, I. A., Rozi, F., & . S. (2022). The Implementation of a Bilingual Program to Improve the English Skills at Al Azhar Islamic Elementary School 25 Semarang. *International Journal of Research and Review*, 9(12), 506–518. <https://doi.org/10.52403/ijrr.20221256>
- Uktolseja, L. J. (2026). EFL Students' Willingness to Communicate: A Survey of Verbal Participation in Speaking Classes. *Jurnal Pendidikan Bahasa*, 13(1).
- Wang, M., Wang, H., & Shi, Y. (2022). The role of English as a foreign language learners' grit and foreign language anxiety in their willingness to communicate: Theoretical perspectives. In *Frontiers in Psychology* (Vol. 13). Frontiers Media S.A. <https://doi.org/10.3389/fpsyg.2022.1002562>
- Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods. Sixth Edition* (L. Fargotstein, C. Neve, K. DeRosa, & G. Dickens, Eds.; Sixth Edition). SAGE Publication, Inc.

The Effectiveness of Kumon Learning Method Toward Student's Reading Comprehension Skill

Alianda Luthfan Dzaki¹⁾, DiskaFatima Virgiyanti²⁾, Aries Fachriza³⁾, Annisa Kurnia Asri⁴⁾

^{1, 2, 3}Tadris Bahasa Inggris (TBI), Fakultas Tarbiyah, Universitas DarussalamGontor

⁴ Pendidikan Bahasa Inggris, (PBI), Fakultas Ilmus Sosial dan Humaniora, Universitas Ibrahimy (UNIB) Situbondo

Email Correspondence : diska.fatima@unida.gontor.ac.id

ARTICLE INFO	ABSTRACT
Keywords: <i>[Kumon Learning Method Reading Comprehension English as a Foreign Language (EFL) Reading Skills English Education]</i>	English proficiency has become increasingly important in supporting students' academic and future professional success, encouraging parents to seek additional English learning opportunities outside formal education. One of the widely adopted institutions is Kumon, which applies an individualized and self-paced learning approach. However, studies discussing the effectiveness of the method in Kumon course in improving reading comprehension skills in English as a Foreign Language (EFL) classes remain limited, particularly in the Indonesian students. Therefore, this study aims to examine the effectiveness of the Kumon Learning Method toward students' reading comprehension skills at Kumon course that is located in Ahmad Yani street, Purwokerto, East Java, Indonesia. This research employed a descriptive qualitative design involving six students aged 11–12 years old enrolled in the EFL Class. Data were collected through interviews, observations, questionnaires, and documentation. The findings reveal that the Kumon Learning Method effectively enhances students' reading comprehension through individualized learning, leveled worksheets, repetitive practice, and Oral Reading Check (ORC) activities. Students demonstrated improvements in vocabulary acquisition, reading fluency, comprehension ability, confidence, and learning independence. In conclusion, the Kumon Learning Method provides an effective supplementary approach for improving reading comprehension skills in EFL learning and may serve as an alternative model for supporting English language development among young learners.

Article History:

Received : 22-06-2026

Revised : 13-08 2026

Accepted : 15-08-2026

Published : 20-08-2026

1. Introduction

It is an essential necessity for the future of children, both in their education and in the workforce. Particularly as time progresses, the advancement of more complex technology necessitates using English as the medium of instruction, which is extremely important (Husna & Murtini, 2019). Basic English proficiency is essential for students who plan to pursue middle and higher school. Therefore, they won't be shocked when junior high and high school pupils are compelled to take English classes.(Jumaira et al., n.d.).

With English no longer being a mandatory subject in primary schools currently, this decision is viewed as increasingly significant. Parents worry that their children will lose the chance to establish a solid groundwork in English if they don't get extra tutoring (Camerroz et. al., 2024). This groundwork is essential for gaining entry to advanced education and future employment opportunities. Parents think that having their children participate in tutoring or language classes will provide them with extra, targeted support, particularly in developing English language abilities, which can frequently pose difficulties in formal education. Parents are eager to rely on tutoring organizations that assist their children in learning more effectively than in school (Halimah, 2019). They see this as a calculated step to provide their kids with a future competitive advantage.

Kumon is one of the English language tutoring facilities that has expanded globally. Toru Kumon established Kumon, also known as Kumon Educational Japan Co. Ltd., a non-formal tutoring organization in Osaka, Japan, in 1958. It grew thru a franchise structure. (Ardiansyah et al., 2022). In essence, math is taught in Kumon more than English. The founder was looking for a way to help his primary school-aged son, Takashi, with his learning issues. As a father and math teacher, he wanted to create a model or learning system that would make it easier for his kid to grasp and learn. (Indah, 2023). This led to the creation of the Kumon method, which is currently being studied by over 4 million 200,000 students in 51 countries. With its slogan "Bring Success," Kumon eventually made English the Institute's second focus because it is thot to be the key to success that must be learned and mastered.

Kumon stands out from other educational establishments thanks to its distinctive teaching approach. As a result, the researcher chose Kumon to be the subject of this investigation. Kumon employs an individual or autonomic approach, which allows each student to get instruction based on their own aptitude and speed rather than age or grade level. This enables pupils to fully comprehend subjects without placing undue pressure on them. Thru regular self-practice, Kumon's English program also seeks to assist pupils in gradually improving their vocabulary, grammar, and reading abilities. Its emphasis on teaching solid principles and enabling students to learn independently sets it apart from other educational institutions that still employ old methods (Indah, 2023). Children studying English benefit greatly from this approach since it fosters better comprehension, self-learning, focus, study habits, self-confidence, and motivation. (Orcos et al., 2019).

Kumon in Indonesia emphasizes three educational programs: Mathematics, English, and Indonesian. In its English language learning program, Kumon offers the EFL (English

as a Foreign Language) class aimed at children for whom English is a secondary language, learning it as a foreign language. Students will enhance their comprehension of English fundamentals, particularly in reading skills, through the resources provided in this program. In nations like Indonesia, Japan, Korea, and others where English isn't the main language, this program enjoys significant popularity. Kumon Learning Institute began developing English lessons referred to as EFL class aimed at students from countries where English is not the native language, such as Indonesia, after noticing the increasing prevalence of English as a foreign language and the growing focus on children learning it (Farani et al., 2024). Kumon's EFL program is specifically designed for kids whose first language is not English.

With the pressure on students to master English and the fierce competition for admission to elite universities, many view Kumon Courses as a way to achieve their goals. This aligns with the belief that "the language environment of a child will influence the quality and amount of the language they use." The initial Kumon Method class in Indonesia launched in 1991, and since then, the method has expanded to 50 nations and territories, covering numerous regions throughout Indonesia. By December 2016, Kumon had educated 4.26 million children (Yuliasih et al., 2018).

The presence of English is crucial in many areas, particularly in globalization and communication as an international language. The idea of language acquisition encompasses four fundamental elements that constitute the basis of the learning process. The four concepts are reading, writing, listening, and speaking. As stated by Nunan, reading is a smooth activity in which readers combine textual information with their prior knowledge to construct meaning (Nunan, D. & Carter, 2001). It describes the reading comprehension process, where readers engage with the text by combining the information given with their previous knowledge and experiences to create meaning. Proficient readers carry out reading comprehension in a smooth, effective, and instinctive way.

Reading comprehension is widely recognized as one of the most essential skills in English language learning because it enables learners to construct meaning from written texts rather than merely decode words. Reading is an interactive cognitive process in which readers combine textual information with their existing knowledge to interpret and understand a text (Anderson & Pearson, 1984). This perspective is explained through Schema Theory, which proposes that comprehension occurs when readers activate their prior knowledge or schemas and integrate new information into existing cognitive structures. Therefore, successful reading comprehension depends not only on vocabulary knowledge but also on learners' ability to connect previous experiences with new textual information.

Interactive Reading Theory, which posits that reading comprehension is the outcome of the interplay between top-down processes like background information, prediction, and inference and bottom-up processes like word recognition and decoding, further supports this viewpoint (Rumelhart, 1977; Stanovich, 1980). Rather than relying solely on linguistic knowledge, effective readers continuously integrate these processes to construct meaning. Consequently, reading instruction should provide opportunities for

learners to practice decoding skills while simultaneously encouraging active meaning-making through continuous engagement with texts.

The Kumon Method's features strongly correspond with theories of self-regulated learning (SRL). Self-regulated learning, according to Zimmerman (2002), is a process in which students actively control their own learning thru goal-setting, progress tracking, strategy selection, and learning outcome evaluation. Kumon's individualized worksheets, mastery-based progression, and self-paced learning environment encourage students to become independent learners who gradually develop responsibility for their own academic progress. Through repeated practice and continuous feedback, learners are expected to improve not only their reading ability but also their learning autonomy and confidence.

In addition, the Kumon Method reflects the principles of Constructivist Learning Theory, which emphasizes that learners actively construct knowledge through meaningful learning experiences rather than passively receiving information from teachers (Vygotsky, 1978). Constructivist learning encourages students to solve problems independently while receiving appropriate guidance when necessary. In Kumon classrooms, instructors function primarily as facilitators who provide assistance only when students encounter difficulties, allowing learners to discover patterns, construct understanding, and develop independent learning habits. Such an approach supports deeper cognitive processing that is essential for reading comprehension.

Another important aspect supporting the Kumon Method is individualized learning. Individualized instruction allows students to learn according to their own abilities, learning speed, and readiness rather than following a uniform classroom pace. Research has shown that personalized instruction improves learning motivation, academic achievement, and learner engagement because instructional materials are matched to individual learning needs (Pane et al., 2017). Likewise, self-paced learning enables students to spend additional time mastering difficult concepts before progressing to more advanced materials, reducing learning anxiety and promoting long-term retention (Means et al., 2014). These characteristics closely resemble the instructional principles implemented in Kumon English classes.

Within the context of English as a Foreign Language (EFL), these theoretical perspectives provide a strong rationale for examining the effectiveness of the Kumon Method in improving reading comprehension. Reading comprehension requires learners to activate prior knowledge, apply effective reading strategies, monitor their understanding, and continuously build new knowledge from texts. Since Kumon integrates individualized instruction, repeated reading practice, mastery learning, and self-directed learning, the method is theoretically expected to facilitate the development of reading comprehension skills among EFL learners.

Additionally, according to Anderson et al., reading is a method for interpreting a text's significance by utilizing one's prior reading knowledge to build meaning (Neil, 2004). Tarigan, as cited by Panjaitan, defines reading as a process where a reader receives the message communicated by the writer via written language (Panjaitan et al., 2022). This process requires engaging the brain to think and deduce, merging word recognition with

existing knowledge to comprehend the meaning of the written material. Simultaneously, Goodman in Banditvilai states that reading is an active process where readers employ effective techniques to derive meaning from written material (Banditvilai, 2020). Therefore, reading is a crucial ability that everyone should have, as it allows us to communicate information effectively and precisely.

Not to mention the demands of the workplace, particularly in well-known and advanced firms where proficiency in the English language is a must for employment. This is crucial because, in addition to being an academic topic, English is a necessary life skill in the age of globalization and the workforce. Early English proficiency will give pupils a solid basis for comprehending increasingly difficult ideas at later educational levels (Kustini et al., 2012). Additionally, the teachers stated that teaching English at the elementary school level fosters critical thinking, creativity, and problem-solving abilities. They contend that teaching strategies that incorporate interactive communication and vocabulary expansion can boost students' self-assurance when speaking English. Since English is the dominant language used in many educational fields and the only language used to preserve information in books and journals in both printed and electronic versions, it has been widely used by students, teachers, and researchers worldwide (Rao, 2019). Additionally, learning English at a young age gives pupils wider access to global educational resources, like articles, films, and online platforms, which expands their learning options.

Students still find learning English intimidating, which makes them unwilling to do so. Nevertheless, a number of students assert that learning English is difficult, and they provide convoluted justifications for this (Indah, 2023). The fundamental English knowledge they learn in school is at odds with the demands of the increasingly advanced technology world and the increasingly dominant period of globalization. They also become less enthusiastic and interested in learning English as a result of this phenomena. Furthermore, not every parent is capable of teaching their children English at home. In order to maximize their children's potential and English proficiency, parents adopt a variety of viewpoints and approaches to these issues. Enrolling their kids in a Kumon tutoring center is one of them.

The topic of Kumon has been studied by a number of researchers; we will give a brief summary of their findings. According to the first study, Kumon's English program seeks to make pupils more self-reliant and accountable. It does not rely on direct instruction, which sets it apart from conventional classroom teaching methods. Instead, well-organized resources, worksheets, and assistance are given to pupils so they can learn at their own pace (Indah, 2023). The second study investigated the methods of teaching English at Kumon. Specifically, it examined the curriculum, learning strategies, learning materials, learning media and assessment procedures used at the English course. By concentrating on more than just Kumon's learning approach, the researcher delved deeper into the process of teaching English at Kumon (Trihapsari, 2021). Using qualitative approaches, their study examines English language teaching methods and self-directed learning. In contrast, the author's research focuses on how the Kumon curriculum affects students' reading comprehension abilities in foreign language classes.

As can be seen from the explanation above, the author selected Kumon as the tutoring facility to be the subject of the study because it was founded in 1958 and has a large global footprint. Parents may believe that their children are looking for additional assistance outside of the classroom since English is one of the subjects that students find most difficult and because mastery in the language is becoming increasingly important in both education and the workforce. Many parents enroll their children in Kumon courses in an attempt to address the issue because they have high expectations for the quality of education in schools, especially when it comes to EFL (Trihapsari, 2021). There are Kumon courses in practically every area of Indonesia, which is evidence that Kumon is a very popular tutoring service and that parents are interested in enrolling their kids there. Additionally, Kumon offers a reliable, organized, and tried-and-true learning approach called the "Kumon Method."

Regretfully, the researcher only discovered a small number of studies that addressed reading comprehension abilities in EFL classes among the numerous earlier Kumon studies. The majority of the earlier research on Kumon had to do with math courses. According to researchers' Google Scholar searches, only one of the 70 articles published with the keyword "EFL reading comprehension" directly addressed reading comprehension skills in Kumon EFL classes, and only twelve of the 97 articles with the keyword "Kumon learning method in EFL" addressed the Kumon learning method in EFL classes. Out of over 100 articles, over 50 of them address the "Kumon program in mathematics." This indicates that there is still a dearth of study on the English language subject of reading comprehension skills in EFL classrooms, particularly in Indonesia. To support this claim, the researcher will provide an outline in the form of a table.

Table 1. The Comparison of Kumon's researches

Keyword	Total Results Found	Relevant Results	Explanation
Kumon EFL reading comprehension	70	1	This topic is rarely discussed, with only 1 relevant article out of 70 results.
Kumon learning method in EFL	97	12	Research focuses on Kumon learning methods in EFL, with broader coverage.
Kumon program at mathematics	100 50		This topic is the most researched, showing the dominance of Kumon in mathematics.

They frequently emphasize the growth of mathematical abilities based on the narratives and tables detailing earlier research on the subject of the Kumon program. In the meanwhile, there are still very few research publications that address Kumon in EFL classes, particularly when it comes to the topic of reading comprehension. This implies that research on the reading comprehension abilities of students enrolled in EFL classes at Kumon tutoring centers is necessary, especially in Indonesia.

2. Methods

2.1 Research Design

This study employed a descriptive qualitative design to explore how the Kumon Learning Method supports students' reading comprehension in an English as a Foreign Language (EFL) context. A descriptive qualitative approach was chosen because it enables researchers to obtain an in-depth understanding of participants' experiences and perceptions in natural settings while preserving the authenticity of educational practices (Creswell & Poth, 2018; Sandelowski, 2000).

2.2 Research Site and Participants

The study was conducted at Kumon Ahmad Yani Purwokerto, Banyumas Regency, Central Java, Indonesia. The site was selected purposively because it implements the standardized Kumon English program and provides a suitable context for examining individualized learning. Participants were selected using purposive sampling (Patton, 2015). They consisted of six students enrolled in the Kumon English program (five Grade 6 students and one Grade 5 student), one English tutor, and the owner of the learning center. These participants were chosen because they possessed direct experience with the implementation of the Kumon Learning Method and could provide rich information relevant to the research objectives (Creswell & Poth, 2018).

2.3 Data Collection

Data were collected through semi-structured interviews, non-participant classroom observations, and document analysis to obtain comprehensive evidence and enhance methodological triangulation (Lincoln & Guba, 1985). Semi-structured interviews were conducted with the students, tutor, and owner to explore their experiences, perceptions, and views regarding the implementation of the Kumon Learning Method and its contribution to reading comprehension. The flexible interview format enabled the researcher to probe participants' responses while maintaining consistency across interviews (Kallio et al., 2016).

Non-participant classroom observations were carried out during regular learning sessions to examine students' learning behaviors, tutor guidance, classroom interactions, and the implementation of individualized worksheet-based learning. Observation findings were documented through field notes. Document analysis complemented the interview and observation data. The analyzed documents included students' worksheets, learning progress records, instructional materials, and classroom photographs, which

provided additional evidence of students' reading development and the implementation of the Kumon program.

2.4 Research Instruments

In qualitative research, the researcher served as the primary research instrument responsible for collecting, interpreting, and analyzing the data (Creswell & Poth, 2018). Supporting instruments included semi-structured interview guides, observation sheets, field notes, audio recordings, and documentation checklists. These instruments facilitated systematic data collection and strengthened the credibility of the findings through multiple sources of evidence (Asipi et al., 2022).

2.5 Data Analysis

The data were analyzed using the interactive model proposed by Miles, Huberman, and Saldaña (2014), consisting of data collection, data condensation, data display, and conclusion drawing and verification. Data from interviews, observations, field notes, and documents were transcribed, organized, and repeatedly reviewed before being coded inductively. Similar codes were grouped into categories and synthesized into four themes: (1) students' initial reading comprehension challenges, (2) individualized worksheet-based learning, (3) the collaborative role of tutors in supporting independent learning, and (4) increased motivation and reading confidence (Miles et al., 2014; Saldaña, 2021).

The identified themes were then organized into thematic displays to facilitate comparison across participants and data sources. The findings were continuously verified by revisiting the original data and comparing evidence obtained from interviews, observations, and documentation to ensure consistency and credibility (Lincoln & Guba, 1985; Creswell & Poth, 2018).

2.6 Trustworthiness

To enhance the trustworthiness of the findings, methodological triangulation was applied by comparing evidence obtained from interviews, classroom observations, and document analysis. Data were analyzed iteratively through repeated review of transcripts, field notes, and documentary evidence to ensure that the interpretations accurately reflected participants' perspectives. This process strengthened the credibility and dependability of the findings (Lincoln & Guba, 1985; Creswell & Poth, 2018).

3. Result

This section presents the findings obtained from classroom observations, semi-structured interviews with tutors, parents, and students, as well as document analysis conducted at Kumon Ahmad Yani, Purwokerto. The data analysis generated four major themes regarding the contribution of the Kumon Learning Method to students' reading comprehension development, they are (1) Students' Initial Reading Comprehension Challenges, (2) Individualized Worksheet-Based Learning, (3) The Collaborative Role of

Tutors in Supporting Independent Learning, and (4) Increased Motivation and Reading Confidence.

3.1. Students' Initial Reading Comprehension Challenges

The analysis of interviews and classroom observations revealed that students entered the Kumon English Program with varying levels of reading comprehension ability. Although most students were able to decode individual English words, they encountered difficulties in constructing meaning from sentences and short passages. These initial challenges appeared consistently throughout the placement process and the first learning sessions, indicating that students' reading difficulties extended beyond pronunciation accuracy.

During the initial observations, several students frequently paused when encountering unfamiliar vocabulary, reread sentences multiple times, and waited for tutor confirmation before answering comprehension questions. Rather than attempting to infer meaning from context, many students focused on understanding individual words first, which interrupted the overall reading process.

a) Tutor 1 described this pattern as follows:

"Most new students are able to pronounce the words correctly, but they usually do not understand the meaning of the sentences. When they find unfamiliar vocabulary, they tend to stop reading and wait for assistance."

b) The students themselves expressed similar experiences. One student explained, *"Sometimes I know how to read the words, but I don't know what the passage is talking about."*

c) Another student admitted, *"When I find difficult words, I usually stop because I am afraid my answer will be wrong."*

The classroom observations further illustrated this pattern. During one placement activity, a student reread the same paragraph several times before attempting to answer a comprehension question. Instead of identifying the main idea, the student repeatedly searched for the meaning of unfamiliar vocabulary. Similar behaviors were observed among other students during the early meetings, suggesting that vocabulary uncertainty frequently disrupted meaning construction.

Taken together, the interview and observation data indicate that students' initial reading comprehension challenges followed a similar pattern. Difficulties in vocabulary understanding, sentence interpretation, and limited confidence interacted with one another, preventing students from reading texts fluently and comprehending them as complete units of meaning. This initial condition became the starting point for the individualized learning process implemented in Kumon.

3.2. Individualized Worksheet-Based Learning

Following the identification of students' initial reading difficulties, the learning process centered on Kumon's individualized worksheet system. The findings demonstrate that the worksheets functioned not merely as learning materials but as a structured pathway through which students gradually developed reading comprehension.

Classroom observations showed that every student worked independently according to his or her current achievement level. Students were not required to progress at the same pace; instead, advancement depended on successful completion of the previous worksheet. This individualized progression enabled students to experience gradual improvement without feeling pressured by classmates' performance.

- a) Tutor 2 explained,
"Students only move to the next worksheet after mastering the previous one. The repetition may look simple, but it helps them understand the text little by little."
- b) The owner also emphasized that the worksheets were intentionally designed to support incremental learning.
"The worksheets are arranged from simple to more complex levels. Students build their understanding step by step instead of learning everything at once."

Observation data further revealed that tutors consistently encouraged students to solve reading problems independently. During one learning session, a student asked about the meaning of a sentence. Rather than translating the sentence immediately, the tutor instructed the student to reread the previous paragraph and identify keywords that could help infer the meaning. Only after the student had attempted several strategies did the tutor provide additional guidance.

This instructional pattern was repeatedly observed throughout the learning sessions. Students were encouraged to revisit the text, reconsider their answers, and identify textual clues before receiving direct assistance. Consequently, learning became an active process in which students gradually developed strategies for constructing meaning independently.

- c) One student reflected on this experience by stating,
"The worksheets become more difficult slowly, so I don't feel overwhelmed."

These findings suggest that individualized worksheet-based learning did not simply increase reading practice. Instead, it gradually shifted students from relying on tutor explanations toward becoming more active readers who attempted to solve comprehension problems independently.

3.3. The Collaborative Role of Tutors in Supporting Independent Learning

As students progressed through increasingly challenging worksheets, tutors played an essential role in maintaining learning consistency while preserving the principle of

independent learning. Rather than functioning as knowledge transmitters, tutors acted as facilitators who continuously monitored students' progress and adjusted their guidance according to individual needs.

- a) The owner explained,
"Independent learning does not mean students learn alone. Tutors observe every student's progress and provide guidance only when it is needed."
- b) Similarly, Tutor 1 stated,
"We do not immediately tell students the answers. We first encourage them to think again, reread the passage, and identify the clues by themselves."

These instructional practices were consistently reflected in the classroom observations. During several learning sessions, tutors responded to students' questions with guiding prompts instead of direct explanations. Students were encouraged to reread passages, compare sentences, or identify keywords before discussing their answers. Such interactions allowed students to remain actively engaged in the comprehension process while still receiving sufficient support to prevent frustration.

- c) Students also recognized this learning approach.
"If I make mistakes, the tutor asks me to read again first. Usually I can find the answer by myself."

These findings demonstrate that tutor support was closely integrated with the individualized worksheet system. Rather than reducing students' independence, tutor guidance functioned as temporary assistance that encouraged learners to become increasingly responsible for their own reading comprehension.

3.4. Increased Motivation and Reading Confidence

The progressive learning process observed throughout the Kumon program was accompanied by noticeable improvements in students' motivation and reading confidence. Changes became increasingly evident across the observation period. Students who initially hesitated to read aloud gradually demonstrated greater willingness to participate in Oral Reading Check (ORC) activities and attempted to answer comprehension questions with less hesitation.

- a) One student explained,
"At first I was afraid of reading English because I thought I would pronounce the words incorrectly. Now I feel more confident because I practice every day."
- b) Another student stated,
"When I finish difficult worksheets by myself, I feel proud and want to continue learning."
- c) The tutors also observed similar changes.

"Students who were initially reluctant to read aloud eventually became much more confident after completing several worksheet levels."

Observation notes further supported these interview findings. During the early meetings, several students spoke quietly, paused frequently, and waited for tutor confirmation before continuing to read. By the later meetings, these students completed longer passages more fluently, volunteered to read aloud, and corrected minor mistakes independently without waiting for tutor intervention.

The findings indicate that increased confidence emerged as a gradual outcome of repeated learning experiences rather than as an immediate result of instruction. Students' successful completion of progressively challenging worksheets appeared to strengthen their willingness to engage with English texts, while continuous tutor guidance helped maintain persistence throughout the learning process.

Overall, the four themes illustrate a coherent developmental process. Students initially entered the program with limited reading comprehension, particularly in vocabulary understanding and meaning construction. Through individualized worksheet-based learning supported by structured tutor guidance, they gradually developed greater independence in reading. As these successful learning experiences accumulated, students demonstrated increased motivation, stronger reading confidence, and greater willingness to participate actively in English reading activities.

4. Discussion

The findings indicate that the Kumon Learning Method supports reading comprehension through a sequential learning process rather than through worksheet practice alone. Students' initial reading difficulties were addressed through individualized worksheets, reinforced by tutor-guided independent learning, and gradually transformed into increased motivation and reading confidence. These interconnected findings suggest that reading comprehension develops through the interaction of individualized instruction, continuous practice, and learner autonomy.

The first finding revealed that students' primary challenges involved constructing meaning from texts rather than simply decoding English words. Limited vocabulary knowledge and uncertainty in interpreting sentence relationships often interrupted comprehension, causing students to focus on individual words instead of the overall message. This finding is consistent with Panjaitan et al. (2022), who reported that vocabulary limitation remains a major obstacle to reading comprehension among Indonesian EFL learners. Likewise, Banditvilai (2020) emphasized that effective reading requires learners to employ comprehension strategies rather than relying solely on word-by-word translation. These findings suggest that students required instructional support that gradually developed both vocabulary understanding and meaning construction.

These initial differences explain why individualized worksheet-based learning became essential in the Kumon program. Allowing students to progress only after mastering each worksheet enabled them to build comprehension according to their own readiness and reduced the pressure of learning at a uniform pace. This finding supports

the principles of personalized learning proposed by Pane et al. (2017) and is consistent with Orcos et al. (2019), who found that Kumon's incremental progression promotes mastery learning through structured practice. Unlike previous studies, however, the present findings indicate that the worksheets served not only as learning materials but also as a mechanism for developing independent reading strategies through repeated interaction with texts.

The findings further demonstrate that individualized learning was strengthened by tutor facilitation. Rather than providing direct answers, tutors guided students to reread passages, identify contextual clues, and evaluate their own responses. Such guidance enabled students to solve reading problems independently while maintaining learning consistency. This finding agrees with Trihapsari (2021), who described tutors as facilitators of self-learning, and with Indah (2023), who found that learner autonomy develops progressively through guided practice. Therefore, learner independence in the Kumon Method emerged through continuous instructional support rather than through learning without teacher involvement.

As students experienced repeated success in completing progressively challenging worksheets, they demonstrated greater confidence and motivation to read English texts. Increased participation during Oral Reading Check (ORC), greater willingness to read aloud, and improved persistence reflected the cumulative effects of gradual learning. Similar findings were reported by Husna and Murtini (2019), who argued that achievable learning goals enhance students' motivation, and by Farani et al. (2024), who found that structured independent learning increases learners' confidence and engagement. The present study further suggests that confidence developed as a consequence of repeated mastery experiences supported by individualized instruction and tutor feedback.

Overall, this study extends previous research on the Kumon Learning Method (Yuliasih et al., 2018; Trihapsari, 2021; Indah, 2023; Farani et al., 2024) by explaining how reading comprehension develops rather than merely confirming that the method is effective. The findings propose a developmental learning process in which students' initial comprehension challenges are addressed through individualized worksheet progression, supported by tutor-guided independent learning, and reinforced through repeated mastery experiences that foster motivation and reading confidence. This integrated explanation contributes to a deeper understanding of individualized learning in EFL reading instruction and provides practical implications for implementing learner-centred reading programs.

5. Conclusion

This study concludes that the Kumon Learning Method supports students' reading comprehension through an integrated learning process consisting of four interconnected stages: initial reading comprehension challenges, individualized worksheet-based learning, tutor-guided independent learning, and increased reading motivation and confidence. The findings indicate that reading comprehension develops not merely through repeated worksheet completion but through the interaction of individualized progression, guided learning, and continuous practice that gradually promotes learner

autonomy. These findings contribute to the understanding of how individualized learning facilitates reading comprehension development in EFL contexts.

This study was conducted in a single Kumon learning center with a limited number of participants; therefore, the findings should be interpreted within this context. Future research is recommended to involve multiple learning centers and a larger number of participants or employ longitudinal and mixed-methods designs to examine the long-term impact of the Kumon Learning Method on students' reading comprehension and other English language skills.

References

- Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 255–291). Longman.
- Ardiansyah, A. P., Yuliatiningtyas, S., & Komunikasi, I. (2022). Analisis semiotika iklan Kumon “Brings Success.” *1*(2), 179–186.
- Asipi, L. S., Rosalina, U., & Nopiyadi, D. (2022). The analysis of reading habits using Miles and Huberman interactive model to empower students' literacy at IPB Cirebon. *2*(3), 117–125.
- Banditvilai, C. (2020). The effectiveness of reading strategies on reading comprehension. *International Journal of Social Science and Humanity*, *10*(2). <https://doi.org/10.18178/ijssh.2020.10.2.1012>
- Cameroz, N. A., Cuilan, J.T., & Chavez, J.V. (2024). Parental Discretionary Influence on Their Children's Manner of Learning English Language. *Forum for Linguistics Studies*, *6*(4).
- Farani, M. A., Farmasari, S., & Isnaini, M. (2024). An analysis of the influences of Kumon learning method toward the learning of English as a foreign language. *Journal of English Education Forum*, *4*(1). <https://doi.org/10.29303/jeef.v4i1.609>
- Halimah, S. N. (2019). *Peranan bimbingan belajar Nurul Fikri Pamulang dalam membantu siswa lulus seleksi masuk PTN* [Undergraduate thesis].
- Hastuti, T. (2020). *An analysis of students' reading comprehension in descriptive text at tenth grade students of SMK Muhammadiyah 2 Metro* [Undergraduate thesis].
- Husna, A. H., & Murtini, R. T. (2019). A study on students' motivation in learning English as a foreign language (EFL) at STIKES Cendekia Utama. *207–220*.
- Indah, A. (2023). *An analysis of students' autonomous learning to develop their English ability at Kumon Institution of Education* [Undergraduate thesis].
- Jumaira, V., Yulina, M., Wati, N. S., Ramadhani, N., Zazalmi, G. C., Berinal, F. A., Rizki, I., Umami, K., Irawan, Y., Zahara, I., Azizah, W., Dewi, K., & Amanda, D. (n.d.). *Solving young learners' problem in English learning*.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Sage Publications.
- Means, B., Toyama, Y., Murphy, R., & Baki, M. (2014). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, *115*(3), 1–47.

- Nunan, D., & Carter, R. (2001). Teaching English to speakers of other languages. *17*(1), 48–50.
- Orcos, L., Hernández, R. M., Espigares, M. J., & Magreñán, Á. A. (2019). The Kumon method: Its importance in the improvement of the teaching and learning of mathematics from the first levels of early childhood and primary education. *Mathematics*, *7*(1), Article 109. <https://doi.org/10.3390/math7010109>
- Pane, J. F., Steiner, E. D., Baird, M. D., & Hamilton, L. S. (2017). *Informing progress: Insights on personalized learning implementation and effects*. RAND Corporation.
- Panjaitan, L., Siahaan, N. P. W., Tampubolon, V. A., & Sipayung, N. (2022). The students' reading problems in private schools in Pematangsiantar. *2*(1), 24–28.
- Rao, P. S. (2019). The role of English as a global language. *Research Journal of English (RJOE)*, *4*(1), 65–79.
- Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI* (pp. 573–603). Erlbaum.
- Saldaña, J. (2021). *The Coding Manual for Qualitative Researchers* (4th ed.). Sage Publications.
- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, *16*(1), 32–71.
- Trihapsari, I. B. (2021). *Practices of teaching English in Kumon Course Tulungagung* [Undergraduate thesis, State Islamic Institute (IAIN) Tulungagung].
- Trilestari, K., Maya, R., Manurung, L., & Simangunsong, V. N. (2024). The benefits of learning English in early childhood: Positive and negative effects.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Yuliasih, R., Suparno, P., Widodo, E., Akbar, P. M., & Pd, M. (2018). Evaluation of implementation Kumon method. *International Journal of Advanced Engineering and Management Research*, *3*(3).
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64–70.