DE_JOURNAL (Dharmas Education Journal)
http://ejournal.undhari.ac.id/index.php/de_journal
E-ISSN : 2722-7839, P-ISSN : 2746-7732
Vol. 7 No. 2 (2026), 372-384

The Effectiveness of Duolingo Mobile Application in Enhancing Vocabulary Acquisition (A Quasi Experimental Study at Seventh Grade of SMP 20 Kota Bengkulu)

Sulendra Priandita[1], Risnawati[2], Endang Haryanto[3]
Email: lennn628@gmail.com,
[123]English Language Education, The State Islamic University of Fatmawati Sukarno, Bengkulu

Abstract

Vocabulary mastery is an important aspect of English language learning, yet many seventh-grade students still experience difficulties in acquiring English vocabulary. This study aimed to determine whether there was a significant difference in English vocabulary acquisition between students who learned using the Duolingo mobile application and those who learned through conventional instruction at SMP Negeri 20 Kota Bengkulu. This study employed a quantitative approach with a quasi-experimental design involving two classes. Class VII-A, consisting of 30 students, was assigned as the experimental group and learned vocabulary using Duolingo, while class VII-B, consisting of 30 students, served as the control group and received conventional instruction. Data were collected through pre-tests and post-tests and analyzed using descriptive statistics and an Independent Samples t-test. The results showed that the mean score of the experimental group increased from 51.80 to 80.33, while the control group increased from 51.90 to 73.33. The t-test results showed a t-value of 2.468 with 58 degrees of freedom and a significance value of 0.017, which was lower than 0.05. Therefore, the alternative hypothesis was accepted, indicating a statistically significant difference in vocabulary achievement between the two groups.

Keywords: ***Duolingo, Vocabulary Acquisition, Mobile-Assisted Language Learning, Quasi Experimental Study, Junior High School Students.***

Abstrak

Penguasaan kosakata merupakan aspek penting dalam pembelajaran bahasa Inggris, namun masih banyak siswa kelas VII yang mengalami kesulitan dalam memperoleh kosakata bahasa Inggris. Penelitian ini bertujuan untuk mengetahui apakah terdapat perbedaan yang signifikan dalam pemerolehan kosakata bahasa Inggris antara siswa yang belajar menggunakan aplikasi seluler Duolingo dan siswa yang belajar melalui pembelajaran konvensional di SMP Negeri 20 Kota Bengkulu. Penelitian ini menggunakan pendekatan kuantitatif dengan desain kuasi-eksperimental yang melibatkan dua kelas. Kelas VII-A yang terdiri dari 30 siswa ditetapkan sebagai kelompok eksperimen dan belajar kosakata menggunakan Duolingo, sedangkan kelas VII-B yang terdiri dari 30 siswa ditetapkan sebagai kelompok kontrol dan menerima pembelajaran konvensional. Data dikumpulkan melalui pre-test dan post-test serta dianalisis menggunakan statistik deskriptif dan uji Independent Samples t-test. Hasil penelitian menunjukkan bahwa nilai rata-rata kelompok eksperimen meningkat dari 51,80 menjadi 80,33, sedangkan kelompok kontrol meningkat dari 51,90 menjadi 73,33. Hasil uji-t menunjukkan nilai t sebesar 2,468 dengan derajat kebebasan 58 dan nilai signifikansi sebesar 0,017, yang lebih kecil dari 0,05. Oleh karena itu, hipotesis alternatif diterima, yang menunjukkan adanya perbedaan yang signifikan secara statistik dalam pencapaian kosakata antara kedua kelompok.

Kata Kunci: Duolingo, Pemerolehan Kosakata, Pembelajaran Bahasa Berbantuan Perangkat Seluler, Penelitian Kuasi-Eksperimental, Siswa Sekolah Menengah Pertama.

Article Information: Received October 2026 | Accepted November 2026 | Published November 2026

Introduction

Vocabulary mastery plays an essential role in helping students comprehend texts, communicate their ideas, and engage in English-language interactions. Listening, speaking, reading, and writing are the four language skills that require sufficient vocabulary knowledge. When students have difficulty identifying word meanings, they may struggle to understand language input and express their thoughts appropriately. Sumitro et al. (2024) highlight the relationship between vocabulary breadth and successful English language learning. This suggests that vocabulary acquisition should be regarded as an important learning objective and evaluated based on students' ability to recognize and use words, rather than simply memorize vocabulary lists. In seventh-grade classrooms, vocabulary development deserves particular attention because some students continue to face challenges in acquiring English vocabulary.

Vocabulary knowledge consists of two important dimensions: breadth and depth. Vocabulary breadth refers to the number of words a student knows, while vocabulary depth involves a more comprehensive understanding of word meanings, forms, relationships, and usage. This distinction enables teachers to evaluate students' vocabulary knowledge more thoroughly instead of relying solely on the number of words they can remember. Recognizing the meaning of a word does not necessarily mean that students can use it correctly when forming sentences. Therefore, the assessment of vocabulary acquisition in this study should encompass both students' recognition of words and their ability to understand and apply word meanings appropriately.

Vocabulary mastery involves several interconnected dimensions that contribute to learners' overall understanding and use of words. Bao and Peng (2024) indicate that the receptive recognition of word meanings serves as a foundation for comprehending language input and contributes to language competence, particularly in reading and text comprehension. According to Panrat and Yanasugondh (2024), knowledge of synonyms enables learners to recognize differences in word usage across contexts, while understanding antonyms helps expand their semantic networks. Similarly, Atiqah et al. (2025) identify collocations as an important aspect related to fluency and accuracy in language use. These findings indicate that vocabulary acquisition should not be measured solely by students' ability to memorize word meanings. Accordingly, this study assesses vocabulary acquisition through three aspects: word meaning recognition, semantic relationships, and word usage.

Word meaning recognition is reflected in students' ability to match word forms with their appropriate meanings. Meanwhile, understanding semantic relationships involves identifying words that have similar or opposite meanings. Word usage is evaluated based on students' ability to select appropriate vocabulary according to the context of a sentence or a particular communicative situation. Although these three indicators are related, each requires a different aspect of understanding. Using these varied indicators allows for a more comprehensive assessment of changes in students' vocabulary acquisition after participating in learning activities.

The Duolingo application facilitates vocabulary learning through mobile devices and is used as the treatment investigated in this study. The constructivist perspective presented in the research review considers learning an active process through which students develop their own understanding. From this perspective, digital learning activities can offer opportunities for students to practice and engage with learning materials in a structured way. Nevertheless, the effectiveness of application-based learning should be evaluated based on students' learning outcomes rather than presumed simply because technology is used. This study examines the differences between Duolingo-assisted learning and conventional instruction in two seventh-grade classes at SMP Negeri 20 Kota Bengkulu. The comparison focuses on vocabulary acquisition, which is assessed before and after the implementation of the treatment.

The use of learning applications can involve a sequence of exercises designed to expose students to vocabulary and provide opportunities for repeated practice. Through repetition, students can

encounter word forms multiple times and reinforce the associations between words and their meanings. Furthermore, feedback provided during the exercises can assist students in identifying and correcting their mistakes as they learn. However, the integration of digital technology does not automatically ensure sustained student engagement, since factors such as device availability, student attention, and instructional guidance may also affect the learning experience. Therefore, the implementation of the treatment should be considered in relation to the actual learning activities and experiences provided to the experimental group.

Comparing Duolingo-based instruction with conventional teaching provides a basis for examining the learning outcomes of both groups under the conditions established in this study. The comparison group makes it possible to distinguish changes resulting from regular classroom instruction from differences that may be associated with the use of Duolingo. In interpreting the results, it is important to consider the consistency of learning materials, instructional duration, and assessment procedures across the two groups. Variations in students' initial characteristics may also affect the interpretation of the findings, particularly when the research involves existing classes rather than randomly assigned students. In this regard, a quasi-experimental design offers an appropriate approach for investigating differences in a school environment without assuming that participants have been randomly allocated to each group.

The previous studies discussed in this research provide a theoretical and empirical basis for determining the vocabulary indicators used in the study. However, these studies do not directly establish the effectiveness of Duolingo in the specific classroom setting under investigation. Research addressing word meaning recognition, synonyms, antonyms, and collocations describes various dimensions of vocabulary knowledge, while the present study emphasizes a comparison between two instructional approaches. Thus, the research gap is both empirical and methodological, as evidence from seventh-grade students at SMP Negeri 20 Kota Bengkulu is examined using a quasi-experimental design with a comparison group. Unlike studies that focus on measuring only one particular aspect of vocabulary knowledge, this study evaluates the learning outcomes of two groups exposed to different instructional methods. The novelty of the research lies in its methodological approach, which applies a quasi-experimental comparison to investigate differences in vocabulary acquisition between Duolingo-assisted instruction and conventional teaching in the context of the school.

The educational setting plays a significant role in understanding the empirical research gap addressed in this study. Students' characteristics, study habits, and classroom teaching practices may differ between schools and classes. Consequently, findings obtained from a single school reflect the particular participants and research conditions rather than representing all students in general. Conducting this study among seventh-grade students at SMP Negeri 20 Kota Bengkulu establishes a specific educational context for interpreting the comparative outcomes of the two instructional approaches. This limitation in context also emphasizes the importance of interpreting the findings cautiously before applying them to other educational environments.

A total of 60 students participated in this study, comprising 30 students from class VII-A assigned to the experimental group and 30 students from class VII-B assigned to the control group. The pre-test and post-test results were examined using descriptive statistical analysis and an independent-samples t-test. The experimental group achieved an increase in its mean score from 51.80 to 80.33, whereas the control group's mean score rose from 51.90 to 73.33. The results of the t-test showed a value of 2.468, with 58 degrees of freedom and a significance level of 0.017, which was below the predetermined threshold of 0.05. These results demonstrate a statistically significant difference in vocabulary achievement between the two groups and offer empirical evidence for evaluating the application's use in English language instruction within the school setting.

The pre-test is used to identify students' initial levels of achievement and provides a reference for evaluating changes after the treatment. Meanwhile, the post-test measures students' performance after each group has completed its designated learning approach. Descriptive statistics are employed to

summarize the distribution and patterns of students' scores, whereas an independent-samples t-test is conducted to determine whether a difference exists between the two groups. The significance value must be interpreted in relation to the research design and the characteristics of the participants. Although the statistical findings provide support for identifying differences between the groups, the generalizability of the results remains limited because the study involves only two classes.

This study aims to investigate whether there is a statistically significant difference in vocabulary acquisition between students who use Duolingo for learning and those who receive conventional instruction. To address this objective, the research question is stated as follows: RQ1: Is there a statistically significant difference in vocabulary acquisition between the two groups of seventh-grade students at SMP Negeri 20 Kota Bengkulu? The research hypothesis is developed based on the constructivist framework, which emphasizes students' active participation in developing their understanding through learning activities. Accordingly, the hypothesis is formulated as follows: H1: Vocabulary acquisition differs significantly between students who learn through Duolingo and those who receive conventional instruction. This study is expected to make two main contributions. First, it provides empirical evidence regarding the comparison of Duolingo-assisted learning and conventional instruction among seventh-grade students. Second, the findings provide a reference for teachers when considering Duolingo as an alternative resource for vocabulary learning activities, with particular attention to the measurable learning outcomes demonstrated by the two groups.

The research focuses on identifying differences in learning outcomes rather than presuming that one instructional method is consistently more effective in every situation. The hypothesis concerning differences is aligned with the research question, as it examines vocabulary acquisition in two groups exposed to different instructional approaches. The constructivist perspective provides a theoretical basis for understanding how student engagement contributes to the development of knowledge and understanding. However, actual learning outcomes must be examined to determine whether participation in application-based learning activities is associated with differences in achievement. From a practical perspective, this study provides contextual evidence that may assist teachers in selecting appropriate learning activities.

The interpretation of the research findings must take into account the study's design and scope. Comparing two classes makes it possible to identify differences in the outcomes of the participating groups, but these findings cannot automatically be generalized to all schools and educational levels. The improvement in the control group's mean score also indicates that conventional instruction contributed to increased vocabulary achievement during the research period. The difference in outcomes between the two groups provides evidence supporting the use of Duolingo in the context examined, although its effectiveness in other educational settings requires further investigation. This study addresses the question of differences in vocabulary acquisition by analyzing pre-test and post-test scores and statistical results from seventh-grade students at SMP Negeri 20 Kota Bengkulu.

The extent to which the findings can be generalized depends on several factors, including the number of classes involved, the characteristics of the participants, and the duration of the instructional intervention. Changes in students' scores may also be affected by other learning experiences occurring throughout the research period. Further research in different school environments could help establish whether the pattern of differences identified in this study remains consistent across contexts. In addition, incorporating vocabulary retention assessments could provide further insight into whether the learning outcomes persist after the intervention concludes. Taking these limitations into account ensures that the conclusions remain appropriately grounded in the available evidence.

Vocabulary competence includes the ability to identify word forms, comprehend their meanings, and apply them correctly in different contexts. Nation (2001) explains that vocabulary connects the comprehension of language input with the ability to produce language. This suggests that insufficient vocabulary knowledge may create difficulties in understanding written texts and

communicating ideas effectively. In this study, vocabulary mastery is evaluated based on students' ability to identify word meanings, recognize synonyms and antonyms, and understand collocations. These aspects incorporate both receptive vocabulary knowledge and the appropriate use of words, aligning with the content of the vocabulary assessment designed for the students.

Receptive vocabulary knowledge refers to the recognition of words encountered through reading and listening, while productive vocabulary knowledge involves actively using words in communication. Rafa (2025) discusses the connection between these two types of vocabulary knowledge in the context of vocabulary practice with Duolingo. Receptive knowledge serves as a foundation for developing more active language use. Furthermore, Ummah (2024) emphasizes that high-frequency vocabulary provides an essential starting point for learning English. Therefore, this study concentrates on basic vocabulary appropriate for seventh-grade students, particularly vocabulary related to food and everyday activities that is included in the instructional materials during the treatment.

Duolingo is a mobile-based language learning application that integrates interactive exercises, gamification features, and learning activities designed to reinforce students' knowledge. Shaqila et al. (2024) relate the use of this application to the incorporation of gamification into vocabulary instruction. The present study is grounded in constructivism, which emphasizes that understanding develops through students' active participation in learning experiences. Through repeated practice and immediate feedback, the application may provide opportunities for students to process, strengthen, and retain vocabulary knowledge. Nevertheless, the effectiveness of Duolingo must be evaluated through measurable learning outcomes. For this reason, the study compares vocabulary test scores between the experimental group and the control group.

Previous studies have generally reported positive findings regarding the use of Duolingo, although they differ in their research settings and methodological approaches. Research conducted at SMPN 2 Pabelan associated the consistent use of Duolingo with improved vocabulary scores within a relatively short period. Raihany and Rahmawati (2025) reported progress in English language skills, particularly in basic vocabulary mastery, among students at a rural school in Madura. In addition, Fediyanto (2025) associated the application with increased student motivation and improved vocabulary learning outcomes among junior high school students compared with conventional teaching methods. These findings serve as a basis for further examining differences in vocabulary acquisition between the experimental and control groups at SMP Negeri 20 Kota Bengkulu.

A synthesis of previous studies reveals consistent findings regarding the potential benefits of Duolingo, although the schools investigated and the learning outcomes examined vary across studies. Research on vocabulary provides a conceptual foundation for assessment, while studies focusing on the application offer empirical support for comparing digital learning with conventional teaching methods. This study contributes context-specific evidence by administering pre-tests and post-tests to two seventh-grade classes, without assuming that the findings can automatically be generalized to other schools. Theoretically, the study examines the relationship between interactive application-based exercises and vocabulary achievement within a constructivist framework. Practically, the findings may serve as a reference for informing vocabulary instruction at the school where the research is conducted.

From a constructivist perspective, students' active participation in learning activities can contribute to the development of their vocabulary understanding. Duolingo incorporates repeated practice activities that differ from traditional lectures and written exercises. Therefore, it is necessary to investigate whether these different instructional approaches lead to differences in students' learning outcomes. Previous studies that have reported improvements in vocabulary mastery provide support for this expectation, although their findings do not necessarily apply equally to all learning contexts. Accordingly, this study investigates this possibility by comparing the post-test scores obtained by the experimental and control groups. H1: A statistically significant difference exists in vocabulary acquisition between students who use Duolingo for learning and those who participate in conventional instruction.

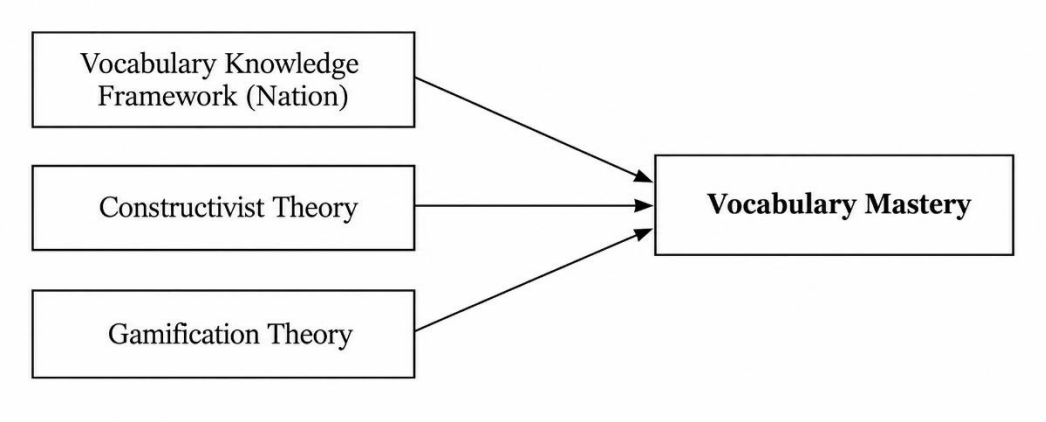


Figure 1. Conceptual Model of the Study

RESEARCH METHODOLOGY

This study employed a quantitative quasi-experimental approach using a pre-test and post-test control group design. This design was selected because the study involved two existing classes without random assignment of individual students, making it appropriate to use a quasi-experimental framework rather than a true experimental design. Capili and Anastasi (2024) explain that quasi-experimental designs are commonly used when randomization is not feasible, particularly in real-world educational settings, while the inclusion of pre-test and post-test assessments allows researchers to examine changes in outcomes following an intervention. The study involved 60 seventh-grade students at SMP Negeri 20 Kota Bengkulu, consisting of 30 students from class VII-A as the experimental group and 30 students from class VII-B as the control group. The experimental group received vocabulary instruction assisted by the Duolingo application, whereas the control group participated in conventional instruction through lectures and written exercises. Both groups completed vocabulary assessments before and after the instructional period to measure their vocabulary achievement. The use of Duolingo as a vocabulary-learning medium is supported by previous studies. Sari et al. (2023) reported the effectiveness of Duolingo in improving students' English vocabulary skills through a quasi-experimental study involving experimental and control classes. Similarly, Andini and Salmiah (2024) found improvements in vocabulary mastery among junior high school students following the use of Duolingo. Furthermore, Rouabhia and Kheder (2024), through a systematic review, identified measurable gains in receptive vocabulary knowledge in several studies involving Duolingo, although they emphasized the need to interpret these findings cautiously because of limitations such as small sample sizes and short intervention periods. In this study, the comparison of pre-test and post-test results between the two groups was intended to examine differences in vocabulary achievement associated with Duolingo-assisted instruction and conventional teaching. Since the participants were drawn from intact classes rather than randomly assigned, potential initial differences between the groups should also be considered when interpreting the findings.

Research Instrument

This study employed a written English vocabulary test developed in accordance with the seventh-grade syllabus for the second semester. The test assessed students' ability to recognize word meanings, understand synonyms and antonyms, and identify appropriate word collocations. The initial trial instrument consisted of 40 items, comprising 10 matching questions, 10 short-answer questions, and 20 multiple-choice questions. The researcher analyzed the test items using SPSS version 29 and determined their validity by comparing each item's calculated r-value with the r-table value of 0.468 at a 5% significance level. A total of 32 items were considered valid because their calculated r-values exceeded 0.468, while the remaining eight items were classified as invalid. Only the valid items were used in the research assessment, with scores converted to a scale ranging from 0 to 100.

Research Procedure

Before the instructional intervention commenced, students in both groups took a face-to-face pre-test to determine their initial vocabulary achievement. The researcher provided instructions on how to

complete the test and use the answer sheets, while supervising the assessment process in the classroom. The intervention for the experimental group was conducted over approximately four to six weeks, comprising eight face-to-face sessions. At the beginning of the intervention, students were introduced to the application, including how to log in, choose topics, and navigate the available levels and points. Each session started with an explanation from the teacher, followed by vocabulary practice using Duolingo. Throughout these activities, the teacher and researcher observed students' participation and addressed any technical problems that arose. The weekly learning topics in the application focused on food and everyday activities, while the control group continued learning through lectures and written exercises without the use of the application.

Upon completion of the instructional period, students took a post-test covering the same assessment content as the pre-test, although the questions were presented in a randomized order. The assessment lasted 45 minutes and was conducted under supervision. Students' answers were evaluated by the researcher using the same answer key applied to the pre-test. This procedure was intended to ensure consistency in testing conditions across both measurement periods and between the two groups. The researcher examined the validity of individual test items using the correlation results from the pilot test and the previously established criterion for the calculated r-value. In addition, the internal consistency of the instrument was examined using Cronbach's alpha, although the resulting coefficient was not reported. The study did not make claims about construct validity or reliability beyond the item analysis results that were presented.

Data Analysis

The data analysis focused on comparing students' vocabulary scores before and after the instruction, as well as examining differences in post-test performance between the experimental and control groups. Descriptive statistics were employed to summarize the scores obtained by each group in both the pre-test and post-test. An independent-samples t-test was conducted to determine whether there was a statistically significant difference between the two groups, using a significance level of 0.05. The three aspects of vocabulary were assessed collectively based on the total test scores. The findings from this analysis served as the statistical basis for evaluating differences in vocabulary achievement between students who used Duolingo as a learning tool and those who received conventional instruction.

Result and Discussion

The study involved 60 seventh-grade students from SMP Negeri 20 Kota Bengkulu, drawn from two intact classes. Class VII-A served as the experimental group, consisting of 30 students, while Class VII-B served as the control group, also consisting of 30 students. The experimental group received vocabulary instruction supported by Duolingo, whereas the control group received conventional instruction. Both groups completed vocabulary tests before and after the instructional period. The test scores were used to describe students' initial and final achievement and to compare the outcomes of the two groups following the intervention. Details regarding the number of participants in each group are presented in Table 2**.**

Table 2. Number of Students by Research Group

Kelompok	Kelas	Jumlah siswa
Eksperimen	VII-A	30
Kontrol	VII-B	30
Total		60

The vocabulary test instrument consisted of items designed to measure students' understanding of English vocabulary in accordance with the seventh-grade learning materials. The pilot-test analysis indicated that 32 of the 40 items met the validity criteria applied, while the remaining eight items did not meet those criteria. An r-table value of 0.468 was used as the reference value for item validity assessment. The study subsequently used only the items identified as valid, with test scores ranging from 0 to 100. This information explains the basis for selecting the items used to measure students'

achievement; however, it does not replace the students' actual scores from the pre-test and post-test assessments.

Table 3. Summary of Vocabulary Test Item Validity Results

Keterangan	Jumlah butir	Kriteria r-tabel
Valid	32	0,468
Tidak valid	8	0,468
Total butir uji coba	40	0,468

The experimental group obtained a mean score of 51.80 on the pre-test and 80.33 on the post-test. The minimum score in this group increased from 30 to 50, while the maximum score rose from 83 to 100. The recorded increase in the mean score was 28.53 points. A total of 29 students achieved higher scores on the post-test, while one student obtained the same score on both assessments. This distribution indicates that the improvement was not limited to students who achieved the highest scores, as the group's minimum score also increased after vocabulary instruction using Duolingo.

The control group obtained a mean score of 51.90 on the pre-test and 73.33 on the post-test. Thus, the initial mean scores of the two groups were nearly identical, whereas the experimental group recorded a higher mean score on the post-test than the control group. The improvement in the control group indicates that students' vocabulary achievement also increased during conventional instruction. Comparing the pre-test and post-test scores of each group provides a descriptive overview of the changes that occurred throughout the study. Details of the available mean scores and score ranges are presented in Table 4.

Table 4. Descriptive Statistics of Vocabulary Test Scores

Kelompok	N	Rata-rata tes awal	Rata-rata tes akhir	Minimum awal	Minimum akhir	Maksimum awal	Maksimum akhir
Eksperimen	30	51,80	80,33	30	50	83	100
Kontrol	30	51,90	73,33				

The post-test scores of the experimental and control groups were compared using an independent-samples t-test. The analysis produced a t-value of 2.468, with 58 degrees of freedom and a significance value of 0.017. Since the significance value was below the established threshold of 0.05, the results indicate a statistically significant difference in post-test scores between the two groups. The detailed findings of this statistical analysis are presented in Table 5.

Table 5: Results of the Independent Samples t-Test

Perbandingan	t	Derajat kebebasan	Signifikansi	Taraf signifikansi
Skor tes akhir kelompok eksperimen dan kontrol	2,468	58	0,017	0,05

The experimental group's higher post-test mean score was consistent with the outcome of the statistical comparison between the groups. The test results provided sufficient statistical evidence to reject the null hypothesis that there was no difference in post-test scores, thereby supporting the research hypothesis regarding differences in learning outcomes. Nevertheless, this conclusion applies specifically to the two classes investigated and does not automatically imply that the same results would be obtained in other schools. The increase in the control group's mean score also demonstrates that vocabulary achievement improved under conventional instruction. Overall, the descriptive findings and inferential analysis indicate that students who received vocabulary instruction using Duolingo achieved higher post-test scores than those who participated in conventional learning activities.

The difference between the mean score increases of the two groups was 6.10 points, calculated from gains of 28.53 points in the experimental group and 21.43 points in the control group. Descriptively, these figures suggest a greater increase in the experimental group. However, this difference should not be regarded as an effect size that has been confirmed through inferential testing. The statistical analysis reported in this study compared the groups' post-test scores; it did not directly examine differences in individual score gains or statistically control for pre-test scores. Accordingly, the statistical test supports the conclusion that the groups differed in their final achievement, while the greater increase observed in the experimental group remains a descriptive finding. Maintaining this distinction ensures that the interpretation of the results accurately reflects the statistical procedures used in the study.

One pedagogical perspective that may help explain the findings is retrieval practice, which involves students repeatedly identifying and recalling vocabulary forms and meanings. When an application introduces vocabulary gradually, repeated exposure may contribute to stronger memory retention through spaced repetition. In addition, presenting vocabulary through audio, written text, and visual contexts may encourage learning through multiple channels, consistent with Mayer's theory of multimedia learning. These concepts provide a theoretical basis for explaining how digital activities may support vocabulary learning. However, because the present study did not directly assess students' cognitive processes, these mechanisms should be understood as theoretically plausible explanations rather than causal effects established by the test results.

Differences in learning outcomes may also reflect students' levels of engagement during instructional activities, rather than simply the use of digital technology. According to self-determination theory, opportunities to choose activities, receive immediate feedback, and observe personal progress may enhance motivation when students' needs for autonomy, competence, and relatedness are supported. Nevertheless, the effectiveness of these features may depend on appropriate task difficulty, teacher support, and consistent application use in the classroom. Future studies could investigate the contributions of individual features by recording practice frequency, time spent using the application, activity completion rates, and students' reactions to feedback. Collecting such data would help researchers move beyond comparing test scores and develop a more detailed understanding of the learning conditions that facilitate vocabulary acquisition.

The research hypothesis predicted a difference in vocabulary acquisition between students learning with Duolingo and those receiving conventional instruction. In the experimental group, the mean score rose from 51.80 on the pre-test to 80.33 on the post-test. Meanwhile, the control group's mean score increased from 51.90 to 73.33. An independent-samples t-test of the post-test results produced a t-value of 2.468, with 58 degrees of freedom and a significance value of 0.017. Since the significance value was below the 0.05 threshold, the findings supported rejecting the null hypothesis that no difference existed between the two groups. The experimental group's higher post-test mean indicated that the observed difference favored students who used Duolingo. These findings provide an answer to the research question within the two classes included in the study, but they do not establish that the same results apply to all seventh-grade students.

The constructivist framework employed in this study helps explain the possible relationship between learning activities and vocabulary achievement. Students develop their understanding through engagement in practice, word recognition, and the use of feedback when answering questions in the application. Repeated practice provides opportunities for students to associate word forms with their meanings, while interactive activities allow them to check their answers throughout the learning process. Shaqila et al. (2024) associated Duolingo with gamification and interactive vocabulary practice. Their findings provide context for interpreting how the structure of the application's activities may encourage student engagement in vocabulary learning. However, this study did not measure engagement or feedback separately; therefore, these mechanisms represent plausible explanations rather than effects that were directly tested.

Vocabulary achievement should be interpreted as knowledge that supports both language comprehension and language use. Nation (2001) explained the relationship between vocabulary knowledge and receptive and productive language skills, suggesting that test scores reflect part of the ability students need to understand and communicate ideas. The research instrument covered meaning recognition, synonyms, antonyms, and collocations, extending beyond the memorization of isolated word meanings. Bao and Peng (2024) positioned receptive meaning recognition as a foundation for understanding language input, while Panrat and Yanasugondh (2024) associated synonyms and antonyms with the understanding of semantic relationships. This framework explains the relevance of the test content to vocabulary acquisition among seventh-grade students. However, the analysis relied on total scores, so the study did not determine which specific aspects of vocabulary knowledge changed the most.

The findings of this study are consistent with previous research on Duolingo, although the supporting evidence comes from contexts and research designs that are not entirely identical. Raihany and Rahmawati (2025) reported improvements in English language skills, particularly basic vocabulary, among students in rural schools in Madura. Fediyanto (2025) also associated Duolingo use with increased motivation and improved vocabulary learning outcomes among junior high school students compared with conventional methods. The similarity in the direction of these findings strengthens the rationale for examining the application in different school contexts, while differences in location limit the assumption that outcomes will be uniform. The present study at SMP Negeri 20 Kota Bengkulu adds a comparison between two seventh-grade classes using pre-test and post-test measurements. Its contextual contribution lies in examining differences in test scores within this particular school, rather than demonstrating that the application is always more effective in every learning environment.

The differences in students' achievement may be explained by the contrasting instructional approaches implemented in the two groups. The experimental group engaged in vocabulary exercises supported by Duolingo, whereas the control group received instruction through lectures and written exercises. The Duolingo materials covered several thematic areas, including clothing and shopping, the home, food and beverages, and workplace communication. Organizing vocabulary around specific themes may have helped students connect words with their relevant contexts, while the application's exercises allowed them to revisit the material. These differences in instructional activities provide a possible explanation for the higher post-test scores obtained by the experimental group. Nevertheless, the study did not examine the individual effects of thematic content, practice frequency, gamification features, or feedback intensity. Classroom conditions and student engagement may also have influenced the outcomes, but these factors were not analyzed independently. Consequently, any explanation of the mechanisms underlying the results should be interpreted cautiously.

The nearly identical pre-test mean scores of the experimental group (51.80) and the control group (51.90) provide a basis for understanding the comparison of their final achievement. The experimental group recorded a mean increase of 28.53 points, compared with 21.43 points in the control group, indicating descriptive improvements under both instructional methods. The difference between the mean score gains was 6.10 points. However, the inferential analysis reported in the study focused on post-test scores rather than changes in individual students' scores. Although the significance value of 0.017 indicates a statistically significant difference in final scores, it does not independently establish whether the difference in score improvements between the groups was statistically significant. In addition, because the study did not report an effect size or confidence interval, the practical magnitude and precision of the observed difference cannot be evaluated using these measures. The findings should therefore be interpreted by clearly distinguishing the statistical evidence provided from conclusions that would require further analysis.

The use of a quasi-experimental design with intact classes made it possible to compare instructional approaches under actual school conditions, although students were not randomly assigned

to the groups. As a result, unmeasured differences between the classes may have contributed to the outcomes, despite their similar pre-test mean scores. The study also relied exclusively on assessments conducted before and after the intervention, meaning that the long-term retention of vocabulary knowledge remains unknown. Furthermore, the total test scores do not reveal whether particular vocabulary components, such as word meanings, semantic relationships, or collocations, contributed differently to students' achievement. Theoretically, the study provides contextual support for the constructivist perspective, as participation in application-assisted vocabulary practice was associated with higher final achievement among the students examined. From a practical perspective, teachers may use Duolingo as a supplementary resource for vocabulary instruction while continuing to evaluate the relevance of the materials, the level of classroom support, and students' learning progress.

Measurement Validity and Vocabulary Transfer

Although improvements in test scores may indicate better vocabulary knowledge, they do not necessarily demonstrate that students can use the words spontaneously in real-life communication. Vocabulary recognition and vocabulary production represent different aspects of language proficiency. Therefore, future assessments should include productive tasks, such as sentence construction or short writing activities, to examine students' ability to use vocabulary appropriately in context. In addition, delayed retention tests should be administered to determine whether students can remember and apply the vocabulary after a certain period.

Implementation Conditions and Learning Accessibility

The use of applications in classroom instruction is influenced by technical resources and students' learning habits, which may affect their opportunities to engage in practice activities. Differences in device availability, internet connectivity, and familiarity with application features may result in varying learning experiences among students. To minimize disparities in learning opportunities, teachers can allocate classroom time for practice and provide alternative activities of equivalent value when digital access is unavailable or disrupted. Furthermore, monitoring students' learning processes, including their practice frequency and the challenges they encounter, can help teachers provide more appropriate guidance. Although these factors do not change the results of the comparison, they remain important considerations when evaluating the feasibility of implementing the approach in classrooms with different levels of available resources.

Recommendations for Future Research

Further studies could involve multiple classes from different schools to obtain a broader representation of diverse learning contexts. Analyzing score changes between the experimental and control groups, together with effect sizes and confidence intervals, would offer a clearer picture of the magnitude of the observed differences and the precision of the estimates. Qualitative data obtained through interviews or learning journals could provide additional insights into how students engage with practice activities and manage difficulties throughout the learning process. Researchers could also examine different levels of application use to investigate whether more frequent practice is associated with improved vocabulary retention. This approach would support a more comprehensive evaluation of learning outcomes, instructional processes, and implementation conditions without assuming that the same method is equally appropriate for every classroom.

CONLUSSION

The study involving 60 seventh-grade students at SMP Negeri 20 Kota Bengkulu revealed a difference in vocabulary achievement between the group using Duolingo and the group receiving conventional instruction. The mean score of the experimental group increased from 51.80 on the pre-test to 80.33 on the post-test, whereas the control group's mean score rose from 51.90 to 73.33. An independent-samples t-test conducted on the post-test scores yielded a t-value of 2.468, 58 degrees of freedom, and a significance value of 0.017. Since this value was below the 0.05 significance level, the research hypothesis concerning differences in achievement between the groups received statistical support. The higher post-test mean in the experimental group indicates that students who learned using Duolingo achieved higher vocabulary scores within the context of this study.

The improvement observed in both groups indicates that conventional instruction and application-assisted instruction were accompanied by increases in scores during the study. The mean score increase of 28.53 points in the experimental group and 21.43 points in the control group illustrates the pattern of change descriptively. The inferential test employed compared the post-test scores; therefore, the difference in score increases does not constitute statistical evidence from a direct test of individual-level changes. Theoretically, the findings provide contextual support for the constructivist framework, as vocabulary practice assisted by an application was associated with higher final achievement in the experimental group. However, this conclusion is limited to the two classes examined and does not establish that similar patterns would occur in other schools or educational levels.

In practical terms, the findings provide a basis for teachers to consider Duolingo as a supplementary option for vocabulary instruction among seventh-grade students. Its use should be aligned with the learning materials and monitored through assessments of student achievement. The quasi-experimental design involving intact classes limits the certainty that the score differences were attributable solely to the intervention, as students were not randomly assigned to the groups. Furthermore, this study measured achievement using total test scores, meaning that the findings do not reveal differences in achievement across individual vocabulary aspects separately. Future research could involve a wider variety of classes or schools, assess each vocabulary aspect independently, and employ score-change analyses and delayed assessments to examine the durability of learning outcomes**.**

REFERENCES

Atiqah, F., Radzuan, A., & Arif, M. M. (2025). Vocabulary Learning Research in ESL / EFL : A Bibliometric Review of Trends , Innovations , and Gaps. 13(2), 46–67.

Andini, J., & Salmiah, M. (2024). The effect of Duolingo on eighth-grade students' vocabulary mastery. Indonesian EFL Journal, 10(2), 247–254. https://doi.org/10.25134/ieflj.v10i2.10089

Bao, Z., & Peng, C. (2024). The effects of EFL wordlist and proficiency on vocabulary knowledge. June, 1–9. https://doi.org/10.3389/fpsyg.2024.1289106

Capili, B., & Anastasi, J. K. (2024). An introduction to the quasi-experimental design (nonrandomized design). The American Journal of Nursing, 124(11), 50–52. https://doi.org/10.1097/01.NAJ.0001081740.74815.20

Maharani, I. M. A. L. (2024). PENGARUH PEMANFAATAN PLATFORM DIGITAL INTERAKTIF TERHADAP PENGUASAAN KOSAKATA BAHASA INGGRIS SISWA SEKOLAH MENENGAH. Jurnal Ilmu Pendidikan, Bahasa Dan Sastra, 1(2), 29–34.

Nabzifa, N., & Fediyanto, N. (2025). The Use of Duolingo as A Digital Tool for Enchaching Junior High School Student'Vocabulary Mastery: Penggunaan Duolingo sebagai Alat Digital untuk Meningkatkan Penguasaan Kosakata Siswa Sekolah Menengah Pertama.

Nation, I. S. P., & Nation, I. S. P. (2001). Learning vocabulary in another language (Vol. 10). Cambridge university press Cambridge.

Panrat, T., & Yanasugondha, V. (2024). A Corpus-based Study of English Synonyms Clear , Obvious , Apparent , and Evident : Implications for. 0672(June).

Rafa, R. A. (2025). Vocabulary Development Strategies for Language Learners : A Review. May. https://doi.org/10.13140/RG.2.2.25183.37287

Raihany, A., & Rahmawati, F. (2025). mproving EFL students' skills using Duolingo in a rural Madurese school. Social, Humanities, and Educational Studies (SHES) Conference Proceedings.

Rizqi, H., & Sumitro, E. A. (2024). SMP : Media Pembelajaran Berbasis Teknologi Duolingo. 4(3), 2319–2330.

Rouabhia, R., & Kheder, K. (2024). Using Duolingo in teaching and learning vocabulary: A systematic review. Indonesian Journal of English Language Studies, 10(2). https://doi.org/10.24071/ijels.v10i2.8873

Shaqila, C., Ilham, I., Rahmaniah, R., Irwandi, I., Bafadal, F., & Hudri, M. (2024). Enhancing English Speaking Skills through Mobile Learning Applications: A Systematic Review. Seminar Nasional Paedagoria, 4(1), 59–71.

Sari, L. P., Hastomo, T., & Nurchurifiani, E. (2023). Assessing the efficacy of Duolingo for acquiring English vocabulary skills: Experimental research. Journal of English Teaching, Applied Linguistics and Literatures (JETALL), 6(2). https://ppjp.ulm.ac.id/journal/index.php/jetall/article/view/17145

Ummah, S. K. (2024). ENHANCING THE VOCABULARY MASTERY OF THE EIGHTH GRADE STUDENTS BY USING PICTIONARY GAME AT SMPN 1 KABAT BANYUWANGI IN THE ACADEMIC YEAR 2024/2025. STATE ISLAMIC UNIVERSITY.