IDEAS
Journal of Language Teaching and Learning, Linguistics and Literature

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Volume 14, Number 1, June 2026
pp. 4462 - 4481

Issued by English study program of UIN Palopo

Deep Learning-Based Adaptive Practice Improves Vocabulary Mastery Among Indonesian EFL Learners

Henryani[1], Zulfah[2], Abdul Haris Sunubi[3], Mujahidah[4], Magdahalena Tjalla[5]

[1,2,3,4,5] Pendidikan Bahasa Inggris, Institut Agama Islam Negeri Parepare, Parepare, Sulawesi Selatan

Article Info

Received: 2026 - 06- 01
Revised: 2026 06-18
Accepted: 2026 06-30

Keywords*:*
Adaptive Learning; Deep Learning-Based Practice; English Vocabulary Mastery; Retrieval Practice; Technology-Enhanced Learning.

DOI*:*
10.24256/ideasv14i1.10788

Corresponding Author:
Henryani
henryanihenryani@gmail.com
Pendidikan Bahasa Inggris, Institut Agama Islam Negeri Parepare, Parepare, Sulawesi Selatan

Abstract

This study investigated the effectiveness of Deep Learning-Based Adaptive Practice (DLBAP), an instructional approach grounded in meaningful learning, retrieval practice, spaced repetition, and adaptive reinforcement rather than artificial intelligence-based deep learning, in improving English vocabulary mastery among Indonesian junior secondary EFL learners. The research addressed whether DLBAP was more effective than conventional vocabulary instruction in enhancing students' vocabulary breadth, vocabulary depth, retention, and productive vocabulary use. The study employed a quantitative quasi-experimental design with a pre-test, post-test, and delayed post-test control group design. The participants consisted of 30 eighth-grade students of MTsS PP Nurul Haq Benteng Lewo, South Sulawesi, Indonesia, divided into an experimental group and a control group. Data were collected through vocabulary tests, classroom observations, and field notes. The experimental group received DLBAP through contextual exposure activities using CoSpaces Edu and adaptive reinforcement using Wordwall, while the control group received conventional instruction. Data were analyzed using descriptive and inferential statistics, including independent and paired samples t-tests. The findings revealed that the experimental group achieved substantially higher vocabulary mastery than the control group, with mean scores increasing from 68.53 in the pre-test to 93.53 in the post-test and remaining relatively high at 89.20 in the delayed post-test. Independent-samples t-test analysis indicated a statistically significant difference between the experimental and control groups in post-test performance ($p < .05$), with a large practical effect size. The experimental group improved from a mean score of 68.53 in the pre-

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

test to 93.53 in the post-test and maintained a relatively high score of 89.20 in the delayed post-test. The findings indicate that integrating contextual learning, retrieval practice, spaced repetition, and adaptive reinforcement can significantly strengthen sustainable vocabulary learning and support technology-enhanced EFL instruction.

1. Introduction

Vocabulary mastery is widely recognized as a fundamental component of second and foreign language learning because it directly influences learners' ability to understand, produce, and communicate meaning effectively. In English as a Foreign Language (EFL) context, inadequate vocabulary knowledge often prevents students from comprehending texts, participating in classroom interaction, and expressing ideas accurately in speaking and writing. Globally, researchers have emphasized that vocabulary acquisition is not merely the memorization of isolated words but a multidimensional process involving retention, contextual understanding, and productive use (Nation, 2013). In the digital era, the increasing integration of technology into education has also transformed vocabulary instruction by introducing adaptive and personalized learning systems that support long-term retention and learner engagement (Webb & Nation, 2017).

In Indonesia, vocabulary learning remains a persistent challenge, particularly in junior secondary schools where English exposure outside the classroom is limited. Many EFL learners still rely heavily on translation-based and memorization-oriented instruction, which often leads to short-term recall rather than sustainable vocabulary retention. This issue becomes more critical in the context of Kurikulum Merdeka, which encourages meaningful, reflective, and student-centered learning practices supported by technology integration in language learning environments (Rahmaniar et al., 2026). Consequently, there is an urgent need for innovative instructional approaches capable of integrating contextual learning, adaptive reinforcement, and technology-assisted review to improve vocabulary mastery among Indonesian EFL learners.

Vocabulary mastery in the present study refers to students' ability to recognize, understand, retain, and use English vocabulary appropriately in communicative contexts. This construct includes vocabulary breadth, vocabulary depth, vocabulary retention, and productive vocabulary use (Qian, 1999). Meanwhile, Deep Learning-Based Adaptive Practice (DLBAP) refers to an instructional approach integrating meaningful learning, mindful learning, joyful learning, retrieval practice, spaced repetition, and personalized feedback through digital learning platforms. For example, students may encounter vocabulary related to school activities in immersive virtual environments and later review the same words through adaptive quizzes adjusted to their learning performance. Such repeated contextual interaction is expected to strengthen both memory retention

and productive language use.

The DLBAP framework is informed by four complementary theoretical perspectives. First, Cognitive Theory of Multimedia Learning (Mayer, 2021) suggests that learning is enhanced when verbal and visual information are processed through multiple cognitive channels. Second, Retrieval Practice Theory argues that active recall strengthens memory consolidation and improves long-term retention. Third, Spaced Repetition Theory emphasizes that learning distributed across time promotes durable memory formation. Finally, Adaptive Learning Theory proposes that personalized feedback and differentiated learning pathways increase learning effectiveness by responding to individual learner needs. The integration of these four principles forms the conceptual foundation of DLBAP.

Previous studies have investigated various approaches to improving vocabulary learning and retention. In one study, Teng (2023) examined whether multimedia input enhanced vocabulary retention among EFL learners using an experimental design. The findings revealed that students who learned vocabulary through multimedia resources demonstrated better long-term retention than those using conventional instruction. The study implied that contextual and visual exposure can facilitate deeper lexical processing. Similarly, Mohammadi (2024) investigated the impact of mobile-assisted vocabulary learning among university students through a quasi-experimental design. The results showed that digital flashcards significantly improved both receptive and productive vocabulary knowledge, suggesting that technology-assisted learning environments can enhance vocabulary acquisition efficiency.

Research has also explored the role of spaced repetition and retrieval practice in vocabulary learning. In another study, Muqaibal et al., (2023) examined the effects of spaced practice on vocabulary retention using computer-assisted language learning. Their findings indicated that repeated review activities distributed over time strengthened long-term retention more effectively than massed learning activities. Likewise, Terai et al., (2021) investigated retrieval practice in EFL vocabulary instruction and found that active recall tasks significantly improved learners' vocabulary retention and productive use. These findings imply that repeated retrieval strengthens memory consolidation and supports sustainable vocabulary learning.

Similarly, studies concerning adaptive learning systems have reported promising outcomes. Xu, (2022) investigated an adaptive vocabulary learning system based on machine learning algorithms and found that personalized reinforcement and immediate feedback improved students' learning efficiency and vocabulary performance. Zhao (2024) further reported that adaptive learning systems adjusted instructional difficulty according to learners' progress, increasing both motivation and vocabulary retention. These findings suggest that adaptive mechanisms can optimize learning experiences by responding to individual learner

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

needs.

In contrast, several studies have highlighted limitations in conventional vocabulary instruction. Alzahrani & Madini (2023) reported that rote memorization methods without adaptive feedback often produce superficial learning outcomes and poor long-term retention. Similarly, Nguyen & Tsai (2024) argued that many vocabulary studies focus only on immediate post-test performance while neglecting delayed post-test measurement, resulting in limited understanding of long-term retention outcomes. These studies imply that vocabulary instruction should emphasize sustainable learning processes rather than short-term achievement alone.

Taken together, previous research demonstrates that contextual learning, retrieval practice, adaptive reinforcement, and technology integration positively influence vocabulary acquisition and retention. Existing studies consistently show that meaningful interaction with vocabulary and repeated review activities contribute to stronger lexical understanding and long-term memory consolidation. Moreover, adaptive systems appear capable of increasing learner engagement by providing personalized feedback and practice opportunities.

Despite these contributions, several important gaps remain. First, most previous studies examined multimedia learning, retrieval practice, adaptive learning, or spaced repetition separately rather than integrating them into a comprehensive instructional framework. Second, empirical studies combining deep learning principles with adaptive vocabulary practice in junior secondary EFL classrooms remain limited, particularly in Indonesian educational contexts. Third, many studies relied solely on immediate post-tests without evaluating delayed retention, making it difficult to determine whether vocabulary gains were sustainable over time. These limitations are significant because vocabulary mastery requires long-term retention and productive application rather than temporary memorization. Junior secondary learners constitute an important population for adaptive vocabulary interventions because they are undergoing a critical stage of language development characterized by increasing cognitive capacity, growing digital engagement, and expanding academic vocabulary demands. At this stage, students are generally receptive to technology-enhanced learning environments and can benefit substantially from personalized instructional support. Consequently, examining the effectiveness of DLBAP among junior secondary EFL learners is both theoretically and pedagogically relevant.

Therefore, the present study aims to examine the effectiveness of Deep Learning-Based Adaptive Practice in improving students' English vocabulary mastery among junior secondary EFL learners in Indonesia. Specifically, the study addresses the following research question: *Is Deep Learning-Based Adaptive Practice effective in improving students' English vocabulary mastery compared to conventional vocabulary instruction?* Based on this question, the study hypothesizes that students receiving DLBAP will demonstrate significantly higher

vocabulary mastery, including breadth, depth, retention, and productive use, than students receiving conventional instruction.

To address this objective, the study employs a quasi-experimental design involving an experimental group and a control group with pre-test, post-test, and delayed post-test procedures. The experimental group receives DLBAP through contextual exposure activities using CoSpaces Edu and adaptive reinforcement activities using Wordwall, while the control group receives conventional vocabulary instruction. This design allows the study to compare immediate vocabulary gains and long-term retention outcomes between the two groups. It is predicted that students exposed to contextual, adaptive, and retrieval-based learning experiences will achieve stronger vocabulary mastery and retention than students taught through traditional memorization-based approaches.

This study contributes theoretically and practically to the field of technology-enhanced language learning. Theoretically, it integrates deep learning principles, cognitive memory theory, and adaptive learning mechanisms into a unified instructional framework for vocabulary learning. Practically, the study offers evidence-based guidance for English teachers seeking effective strategies to improve vocabulary retention and productive vocabulary use in EFL classrooms. Furthermore, the study contributes novelty by combining meaningful learning, mindful learning, joyful learning, contextual exposure, adaptive reinforcement, retrieval practice, and spaced repetition within one integrated instructional model evaluated through delayed retention measurement. Although previous studies have demonstrated the effectiveness of multimedia learning, retrieval practice, spaced repetition, and adaptive learning individually, empirical evidence examining the synergistic integration of these principles within a single instructional framework remains scarce. This gap is particularly evident in Indonesian junior secondary EFL contexts, where vocabulary retention and productive vocabulary use continue to present significant instructional challenges.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

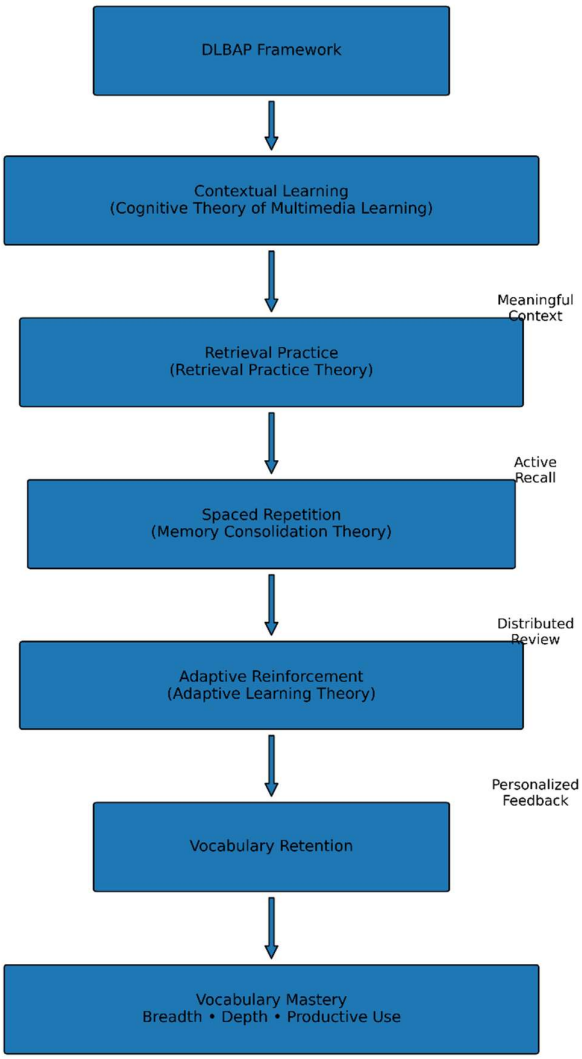

Figure 1. Conceptual Framework of Deep Learning-Based Adaptive Practice (DLBAP) for Vocabulary Mastery Development

Figure 1 presents the conceptual framework underpinning Deep Learning-Based Adaptive Practice (DLBAP). The framework integrates four complementary instructional principles: contextual learning, retrieval practice, spaced repetition, and adaptive reinforcement. Drawing upon Cognitive Theory of Multimedia Learning, Retrieval Practice Theory, Spaced Repetition Theory, and Adaptive Learning Theory, the model proposes that meaningful contextual exposure and repeated retrieval opportunities strengthen memory consolidation. Personalized adaptive reinforcement further supports learners by targeting vocabulary items requiring additional practice. These processes collectively promote vocabulary retention, which functions as the immediate learning outcome and mediating mechanism leading to higher levels of vocabulary mastery, including vocabulary

breadth, vocabulary depth, and productive vocabulary use.

2. Method

Research Design

This study employed a quantitative approach using a quasi-experimental design with a pre-test, post-test, and delayed post-test control group design. The design was selected because the research aimed to examine the effectiveness of Deep Learning-Based Adaptive Practice (DLBAP) on students' English vocabulary mastery in a natural classroom setting where random assignment was not feasible. The experimental group received the DLBAP treatment integrating CoSpaces Edu and Wordwall, while the control group received conventional vocabulary instruction.

The pre-test was administered before the treatment to measure students' initial vocabulary knowledge and to ensure the equivalence of both groups. After six treatment sessions, a post-test was conducted to evaluate immediate vocabulary improvement. A delayed post-test was administered two weeks later to assess long-term vocabulary retention and productive vocabulary use. This design enabled the researcher to compare learning gains and retention outcomes between the experimental and control groups.

Population and Research Participants

The population of this study consisted of all eighth-grade students of MTsS PP Nurul Haq Benteng Lewo, Kabupaten Sidenreng Rappang, South Sulawesi, Indonesia, during the 2025/2026 academic year. The population included two intact classes with a total of 30 students. Because the population size was relatively small and accessible, the study employed a total sampling technique in which all students participated in the research.

The participants were divided into two groups based on their existing classroom membership. Class VIII.A, consisting of 15 students, served as the experimental group receiving Deep Learning-Based Adaptive Practice. Meanwhile, Class VIII.B, consisting of 15 students, functioned as the control group receiving conventional vocabulary instruction. The participants were selected because they were studying English vocabulary as part of the junior secondary school curriculum and represented EFL learners with relatively similar educational backgrounds.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

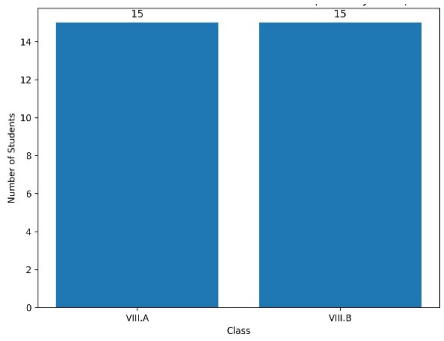

Figure 2. Distribution of Research Participants by Group

Figure 2 presents the demographic distribution of the research participants involved in the quasi-experimental study. The total population consisted of 30 eighth-grade students divided equally into two intact classes. Class VIII.A was assigned as the experimental group receiving Deep Learning-Based Adaptive Practice, while Class VIII.B functioned as the control group receiving conventional vocabulary instruction. The equal allocation of participants created balanced research conditions and minimized potential sampling bias during the treatment process. Furthermore, the use of total sampling strengthened the representativeness and credibility of the data because all members of the accessible population participated directly in the research activities and assessments.

The study involved 30 participants, consisting of 15 students in the experimental group and 15 students in the control group. Although this sample size was sufficient for the scope of the classroom-based intervention, it may limit the generalizability of the findings. Therefore, the results should be interpreted with caution.

Data Collection Techniques and Instrument Development

Data were collected through vocabulary tests and classroom observation. Three vocabulary tests were employed, namely the pre-test, post-test, and delayed post-test. These instruments were designed to measure vocabulary breadth, vocabulary depth, vocabulary retention, and productive vocabulary use. The tests consisted of 100 target vocabulary items including nouns, verbs, adjectives, and adverbs aligned with the treatment materials.

The test items were presented in multiple formats such as multiple-choice questions, fill-in-the-blank tasks, sentence completion activities, and contextual vocabulary use exercises. The pre-test measured students' baseline vocabulary knowledge before the treatment. The post-test evaluated immediate learning outcomes after the intervention, while the delayed post-test measured long-term retention two weeks after the treatment period.

The experimental treatment was conducted in six instructional sessions, each

lasting approximately 90 minutes. During the treatment, students in the experimental group participated in contextual exposure activities using CoSpaces Edu and adaptive reinforcement activities using Wordwall. The instructional procedures integrated meaningful learning, mindful learning, joyful learning, retrieval practice, spaced repetition, and personalized review activities. In contrast, the control group learned vocabulary through conventional instructional methods such as vocabulary lists, translation activities, and teacher-centered exercises.

To ensure content validity, the vocabulary instruments were reviewed by experts in English Language Teaching (ELT). The experts evaluated the relevance, clarity, and appropriateness of the vocabulary items according to the research objectives and students' proficiency level. In addition, a pilot test was conducted on students outside the research sample to examine instrument reliability. The reliability analysis showed a Cronbach's alpha coefficient higher than 0.80, indicating that the instruments had high internal consistency and were appropriate for data collection. The content validity of the research instrument was examined using the Content Validity Index (CVI). The validation process involved three experts with relevant expertise in early childhood education, language learning, and educational assessment. The experts evaluated each item based on its relevance, clarity, appropriateness, and suitability with the learning objectives and the developmental level of the participants.

Each item was assessed using a four-point rating scale, consisting of 1 = not relevant, 2 = somewhat relevant, 3 = quite relevant, and 4 = highly relevant. For the CVI calculation, ratings of 3 and 4 were categorized as relevant, while ratings of 1 and 2 were categorized as not relevant. The Item-Level Content Validity Index (I-CVI) was calculated by dividing the number of experts who rated an item as relevant by the total number of experts.

Since this study involved three experts, an item was considered valid when it obtained an I-CVI value of 1.00. Items with an I-CVI value below 1.00 were revised based on the experts' comments and suggestions. The Scale-Level Content Validity Index using the average method (S-CVI/Ave) was calculated by averaging the I-CVI values across all items. The result showed that the instrument obtained an S-CVI/Ave value of 0.93, indicating that the instrument had acceptable content validity before being used in the pre-test and post-test.

Classroom observations and field notes were also conducted during the treatment sessions to document students' participation, engagement, interaction with digital tools, and responses toward adaptive learning activities. These observational data supported the interpretation of quantitative findings and provided additional insights into the implementation process.

Data Analysis Techniques

The collected data were analyzed quantitatively using descriptive and inferential statistics. Descriptive statistics were used to calculate the mean scores,

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

standard deviations, minimum scores, and maximum scores of the pre-test, post-test, and delayed post-test results for both groups. These analyses provided an overview of students' vocabulary performance before and after the treatment.

Inferential statistical analysis was conducted to examine whether significant differences existed between the experimental and control groups. Prior to hypothesis testing, the data were tested for normality and homogeneity to ensure that the assumptions for parametric analysis were fulfilled. After meeting these assumptions, an independent samples t-test was employed to compare the vocabulary mastery scores of both groups.

In addition, paired samples t-tests were used to examine the improvement within each group between pre-test and post-test scores as well as between post-test and delayed post-test scores. The level of significance was set at 0.05. The findings from these analyses were used to determine whether Deep Learning-Based Adaptive Practice significantly improved students' vocabulary mastery, including vocabulary breadth, depth, retention, and productive vocabulary use.

3. Result

Descriptive Statistics of Students' Vocabulary Mastery

The descriptive statistics revealed differences in vocabulary achievement between the experimental and control groups across the research stages. The experimental group obtained a mean pre-test score of 68.53, while the control group recorded a mean score of 57.33. Following the implementation of the treatment, the experimental group demonstrated a steady increase across the six instructional meetings, reaching a mean score of 93.53 on the post-test and 89.20 on the delayed post-test.

In contrast, the control group showed more gradual progress throughout the instructional sessions. The mean score increased from 57.33 in the pre-test to 70.73 in the post-test and slightly decreased to 68.73 in the delayed post-test. The score progression of both groups is presented in Table 1.

Table 1. Descriptive Statistics, Standard Deviations, and Gain Scores Across Research Stages

Stage	Experimental M	Experimental SD	Experimental Gain	Control M	Control SD	Control Gain
Pre-Test	68.53	7.55	0.00	57.33	13.83	0.00
Meeting 1	73.20	7.70	4.67	59.13	14.18	1.80
Meeting 2	77.20	7.70	8.67	60.93	14.53	3.60
Meeting 3	81.20	7.70	12.67	62.73	14.88	5.40
Meeting 4	85.53	7.55	17.00	65.33	15.60	8.00
Meeting 5	89.53	7.55	21.00	67.13	15.96	9.80
Meeting 6	93.53	7.55	25.00	70.73	16.70	13.40
Post-Test	93.53	7.55	25.00	70.73	16.70	13.40

Delayed Post-Test	89.20	6.61	20.67	68.73	16.70	11.40

The table shows that the experimental group achieved greater vocabulary improvement than the control group across all research stages. The experimental group increased from 68.53 in the pre-test to 93.53 in Meeting 6 and the post-test, with a total gain of 25.00 points. In contrast, the control group improved from 57.33 to 70.73, with a gain of 13.40 points. Although both groups showed progress, the experimental group demonstrated a faster and stronger increase. In the delayed post-test, the experimental group maintained a high mean score of 89.20, indicating strong retention after the treatment. This suggests that DLBAP was more effective than conventional instruction.

Improvement Across Instructional Meetings

The experimental group demonstrated continuous score increases during each instructional meeting. The mean score increased by approximately four points between consecutive meetings from Meeting 1 to Meeting 6. The largest cumulative increase occurred between the pre-test and Meeting 6 stages.

The control group also showed gradual improvement during the instructional period, although the rate of increase was smaller than that observed in the experimental group. The progression pattern indicated incremental gains across meetings without substantial fluctuations.

A slight decrease appeared in both groups during the delayed post-test phase. The experimental group decreased from 93.53 to 89.20, while the control group decreased from 70.73 to 68.73.

Results of Vocabulary Mastery Components

The analysis of vocabulary mastery components in the experimental group demonstrated improvement across vocabulary breadth, vocabulary depth, retention, and productive vocabulary use. During Meeting 1, several students obtained relatively moderate scores in productive vocabulary use compared to vocabulary breadth and depth. By Meeting 6, the majority of students demonstrated higher total scores across all vocabulary components.

Vocabulary breadth consistently recorded the highest component scores during the treatment sessions. Productive vocabulary use remained the lowest scoring component in several meetings despite showing gradual improvement over time. Retention scores increased steadily from Meeting 1 to Meeting 6 without major score reductions.

The highest individual score recorded in the experimental group was 100, achieved by several students during Meeting 6 and the post-test stage. The lowest score recorded during the treatment sessions was 62 in Meeting 1.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

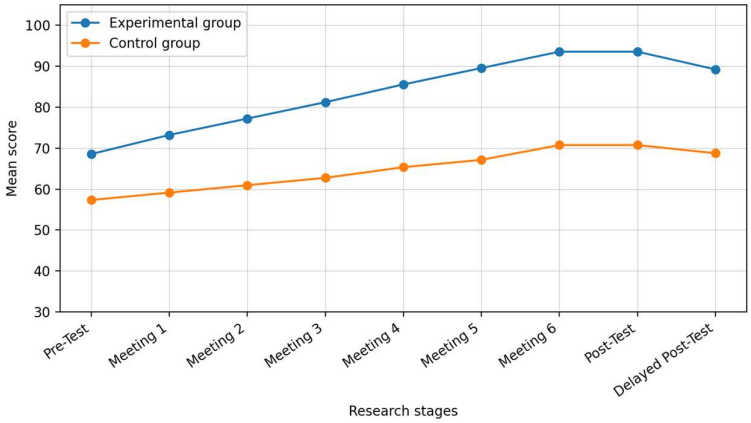

Figure 3. Score Progression Across Research Stages

The line graph illustrates the score progression of the experimental and control groups across the research stages. Both groups showed improvement from the pre-test to the post-test; however, the experimental group demonstrated a sharper and more consistent increase. The experimental group rose from 68.53 in the pre-test to 93.53 in Meeting 6 and the post-test, while the control group increased from 57.33 to 70.73. After the delayed post-test, both groups experienced a slight decline, but the experimental group maintained a higher mean score of 89.20. This pattern indicates that DLBAP supported stronger vocabulary development and retention.

Post-Test and Delayed Post-Test Findings

The post-test results indicated that students in the experimental group achieved higher vocabulary scores than students in the control group. Several students in the experimental group reached the "Excellent" category, with post-test scores ranging from 95 to 100. The delayed post-test scores showed that most students in the experimental group maintained relatively high achievement levels after the two-week interval.

In the control group, post-test scores increased compared to pre-test performance; however, the majority of students remained within moderate achievement categories. The delayed post-test results in the control group also showed a slight reduction in mean scores after the instructional period.

Unexpectedly, a small number of students in the experimental group experienced minor score decreases during the delayed post-test despite obtaining high post-test scores. Similarly, some students in the control group demonstrated stable delayed post-test scores without substantial decline.

Table 2. ANCOVA Results for Post-Test and Delayed Post-Test Scores

Outcome	Source	df effect	df error	F	p-value	Partial eta

						squared
Post-Test	Group	1.000	27.000	218.987	< .001	0.890
Post-Test	Pre-Test covariate	1.000	27.000	1762.430	< .001	0.985
Delayed Post-Test	Group	1.000	27.000	87.837	< .001	0.765
Delayed Post-Test	Pre-Test covariate	1.000	27.000	1060.905	< .001	0.975

The ANCOVA results show that group membership had a significant effect on both post-test and delayed post-test scores after controlling for pre-test scores. For the post-test, the group effect was significant, $F(1, 27) = 218.987$, $p < .001$, with a partial eta squared value of .890, indicating a very large effect. The pre-test covariate was also significant, suggesting that initial vocabulary ability strongly influenced post-test performance. Similarly, the group effect on delayed post-test scores was significant, $F(1, 27) = 87.837$, $p < .001$, partial $\eta^2 = .765$. These findings confirm the strong effect of DLBAP.

Comparative Results Between Experimental and Control Groups

The comparison between the experimental and control groups demonstrated differences in vocabulary achievement across all research stages. The experimental group consistently obtained higher mean scores than the control group from Meeting 1 through the delayed post-test stage.

The difference between groups became more visible after the fourth instructional meeting. During the post-test phase, the experimental group achieved a mean score approximately 22.80 points higher than the control group. During the delayed post-test phase, the mean difference between groups remained above 20 points.

Overall, the experimental group showed higher scores in vocabulary breadth, vocabulary depth, vocabulary retention, and productive vocabulary use throughout the instructional process.

Table 3. Vocabulary Mastery Component Analysis Between Experimental and Control Groups

Component	Experimental M	Experimental SD	Control M	Control SD	Mean Difference	t	df	p-value	95% CI Lower	95% CI Upper	Cohen's d
Breadth	28.000	2.204	21.400	4.881	6.600	4.773	28.000	< .001	3.767	9.433	1.743
Depth	23.46	1.846	17.7	4.06	5.73	4.97	28.0	< .00	3.37	8.09	1.81

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

	7		33	1	3	7	00	1	4	3	7
Retentio	23.40	1.957	17.5	4.42	5.86	4.69	28.0	< .00	3.30	8.42	1.71
n	0		33	2	7	9	00	1	9	4	6
Product	18.66	1.589	14.0	3.36	4.60	4.78	28.0	< .00	2.63	6.57	1.74
ive Use	7		67	9	0	3	00	1	0	0	6

The component analysis shows that the experimental group outperformed the control group in all vocabulary mastery components. The largest mean score was found in vocabulary breadth, where the experimental group achieved 28.00 compared to 21.40 in the control group. Significant differences were also found in depth, retention, and productive use, with all p-values below .001. The t-test results indicate that these differences were statistically significant across all components. Cohen's d values ranged from 1.716 to 1.817, showing large effect sizes. These findings suggest that DLBAP improved not only vocabulary recognition but also deeper understanding, retention, and productive vocabulary use.

4. Discussion

The present study examined the effectiveness of Deep Learning-Based Adaptive Practice (DLBAP) in improving English vocabulary mastery among Indonesian junior secondary EFL learners. The findings showed that the experimental group achieved higher vocabulary scores than the control group across the post-test and delayed post-test stages. The experimental group also demonstrated stronger performance in vocabulary breadth, vocabulary depth, retention, and productive vocabulary use. These findings suggest that DLBAP can support vocabulary learning more effectively than conventional vocabulary instruction. However, the results should be interpreted critically because the experimental group had a higher pre-test mean score than the control group, indicating that the two groups did not begin from completely equal baseline conditions.

The improvement observed in the experimental group may be explained by the interaction of four instructional principles: contextual learning, retrieval practice, spaced repetition, and adaptive reinforcement. These principles did not function separately but appeared to support one another within the DLBAP framework. Contextual learning through CoSpaces Edu provided meaningful exposure to vocabulary in visual and situational contexts, helping learners connect words with concrete meanings. Retrieval practice then required students to actively recall the vocabulary rather than simply recognize or reread it. Spaced repetition strengthened this process by distributing review activities across several meetings, which may have reduced forgetting and supported memory consolidation. Finally, adaptive reinforcement through Wordwall provided repeated practice and feedback based on students' learning needs. The combination of these principles likely created a more active and sustained learning

process than conventional vocabulary instruction.

The findings are consistent with Cognitive Theory of Multimedia Learning, which suggests that students learn more effectively when verbal and visual information are integrated meaningfully. In this study, vocabulary items were not presented only as isolated word lists but were embedded in contextual and interactive learning activities. This may have helped students develop stronger semantic associations and improved their ability to understand vocabulary in context. The results also support Retrieval Practice Theory, as students repeatedly recalled target vocabulary through adaptive tasks. Such repeated recall may have strengthened memory traces and improved long-term access to vocabulary knowledge. In addition, the use of spaced review activities is consistent with the principle that distributed learning is more effective for retention than massed practice.

The results also support previous studies showing that technology-enhanced vocabulary instruction can improve learners' vocabulary achievement, retention, and productive use. Studies on multimedia learning, spaced retrieval, and adaptive feedback have generally reported positive effects on vocabulary acquisition. The present study extends these findings by integrating these principles into one instructional model rather than applying them separately. This integration is important because vocabulary mastery requires more than word recognition. Learners need to understand word meanings, retain them over time, and use them productively in appropriate contexts. The component analysis in this study showed improvement across vocabulary breadth, depth, retention, and productive use, suggesting that DLBAP contributed to both receptive and productive vocabulary development.

Nevertheless, the findings should not be interpreted as evidence that technology alone caused the improvement. Several alternative explanations need to be considered. First, the novelty effect may have influenced students' motivation and engagement. Because CoSpaces Edu and Wordwall were more interactive than conventional instruction, students may have been more enthusiastic simply because the learning tools were new to them. Second, teacher enthusiasm during the implementation of the treatment may have contributed to students' participation and performance. Third, increased exposure to digital learning activities may have provided the experimental group with more varied learning experiences than the control group. This finding is also supported by Rahmaniar et al, (2024), who reported that differentiated instructional practices help educators address diverse learner needs and increase student participation in English language learning. In the present study, students received adaptive reinforcement activities adjusted to their learning performance.Therefore, while DLBAP appears to be effective, its success may have been influenced by a combination of instructional design, learner engagement, teacher support, and technology exposure.

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Another important issue concerns the initial difference between the experimental and control groups. The experimental group started with a higher pre-test mean score than the control group. This baseline difference may have influenced the final outcomes because students with stronger initial vocabulary knowledge may be better prepared to benefit from contextual exposure, retrieval practice, and adaptive reinforcement. For this reason, the post-test comparison should not be interpreted only from the difference in final mean scores. The use of ANCOVA was important because it helped control the initial pre-test difference and provided a more reliable estimate of the treatment effect. Although the adjusted analysis still indicated a strong treatment effect, the baseline difference remains a methodological limitation that should be acknowledged.

The delayed post-test result also deserves careful interpretation. The experimental group decreased from 93.53 in the post-test to 89.20 in the delayed post-test. Although the score remained high, the decline suggests that some vocabulary knowledge was not fully stabilized after the six instructional meetings. This finding is important because it shows that short-term improvement does not automatically guarantee complete long-term retention. The decrease may have occurred because the treatment duration was relatively brief, the delayed post-test was conducted after a two-week interval, and students may have had limited English exposure outside the classroom. It is also possible that productive vocabulary use requires longer practice than receptive recognition. Therefore, future implementation of DLBAP should include longer spaced review cycles and follow-up vocabulary practice after the main intervention.

The retention pattern also suggests that adaptive vocabulary learning should be sustained beyond the immediate treatment period. For instance, Chen & Hsu (2021) found that interactive digital vocabulary platforms significantly improved EFL learners' retention and productive vocabulary use because learners engaged more actively with contextualized lexical tasks. Similarly, Li & Hafner (2022) reported that immersive digital learning environments enhanced vocabulary acquisition by increasing learner interaction, contextual exposure, and cognitive engagement. While DLBAP helped students maintain relatively high delayed post-test performance, the decrease indicates that vocabulary retention requires continuous reinforcement. Teachers should not stop review activities after students achieve high post-test scores. Instead, vocabulary should be recycled across different tasks, contexts, and time intervals. This is particularly important in EFL contexts where students have limited opportunities to encounter and use English outside the classroom.

The pedagogical implications of this study are significant but should be considered realistically. DLBAP can be a useful model for English teachers because it combines meaningful context, repeated recall, spaced review, and adaptive feedback. However, successful implementation requires teacher readiness. Teachers need sufficient digital pedagogical competence to design contextual

vocabulary activities, monitor students' progress, and provide appropriate feedback. Infrastructure is also important. Schools need reliable internet access, adequate devices, and technical support to implement platforms such as CoSpaces Edu and Wordwall effectively. In addition, students' digital literacy should not be assumed. Some learners may need guidance in using digital platforms before they can fully benefit from adaptive vocabulary practice.

The findings also have implications for curriculum implementation. The results also align with Rahman et al. (2023), who found that spaced retrieval activities improved long-term vocabulary retention among secondary EFL learners. In the context of student-centered and technology-enhanced learning, DLBAP supports the goal of making vocabulary instruction more meaningful, interactive, and personalized. However, schools should avoid treating digital tools as automatic solutions. Technology should be integrated with clear pedagogical goals, structured learning procedures, and continuous assessment. Without careful planning, digital vocabulary activities may become entertaining but not necessarily effective for long-term learning. Furthermore, the findings are supported by Wang & Chen (2024), who demonstrated that gamified adaptive vocabulary learning environments increased learner motivation and participation in EFL classrooms. Similar findings were also reported by Busran et al., (2024), who found that digital snake-and-ladder game activities improved young learners' English vocabulary mastery through interactive and engaging learning experiences.

This study has several limitations. First, the sample size was small and limited to one school, which reduces the generalizability of the findings. Second, the study used intact classroom groups rather than random assignment, creating the possibility of selection bias. Third, the experimental group had a higher pre-test mean score than the control group, which may have influenced the learning outcomes despite the use of statistical control. Fourth, the treatment lasted only six meetings, which may not be sufficient to determine the long-term effect of DLBAP. Fifth, the study focused mainly on quantitative test results, while students' perceptions, motivation, and learning experiences were not explored in depth.
Future research should involve larger and more diverse samples across different school contexts. Randomized or matched-group designs would help reduce baseline differences between experimental and control groups. Future studies should also apply DLBAP over a longer period and include longer delayed post-test intervals to examine whether vocabulary retention remains stable over several months. In addition, qualitative data from interviews, learner diaries, or classroom observations could provide deeper insight into how students experience adaptive vocabulary learning. Further research should also examine whether DLBAP improves other language skills, such as speaking, reading comprehension, and writing.

In conclusion, the findings indicate that DLBAP has strong potential to improve vocabulary mastery among Indonesian junior secondary EFL learners. Its

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

effectiveness appears to come from the synergy between contextual exposure, retrieval practice, spaced repetition, and adaptive reinforcement. At the same time, the results should be interpreted with caution because of the small sample size, baseline differences between groups, possible novelty effects, and the decline in delayed post-test scores. Therefore, DLBAP should be viewed not merely as a technology-based intervention but as a structured pedagogical approach that requires sustained practice, teacher readiness, adequate infrastructure, and careful implementation.

5. Conclusion

This study investigated the effectiveness of Deep Learning-Based Adaptive Practice (DLBAP) in improving English vocabulary mastery among Indonesian junior secondary EFL learners. Specifically, the research aimed to determine whether the integration of contextual multimedia exposure, adaptive reinforcement, retrieval practice, and spaced repetition could enhance students' vocabulary breadth, depth, retention, and productive vocabulary use compared to conventional instruction.

The findings demonstrated that DLBAP was more effective than traditional vocabulary learning approaches across all stages of assessment. First, students in the experimental group consistently achieved higher vocabulary scores than those in the control group. Second, the implementation of adaptive and retrieval-based learning activities contributed to significant improvement in vocabulary breadth and depth. Third, the delayed post-test findings indicated that students receiving DLBAP maintained stronger long-term vocabulary retention. Fourth, the integration of contextual and interactive digital learning environments enhanced students' productive vocabulary use and learning engagement. Overall, the research hypothesis was supported, confirming that DLBAP positively influenced sustainable vocabulary mastery among EFL learners.

The study contributes theoretically by integrating deep learning principles, adaptive learning mechanisms, retrieval practice, and multimedia learning into a unified framework for vocabulary instruction. Practically, the findings provide evidence-based guidance for English teachers, curriculum developers, and policymakers seeking more effective technology-enhanced vocabulary learning strategies in EFL classrooms. The study highlights the importance of student-centered, adaptive, and meaningful learning experiences in promoting long-term lexical development.

Based on these findings, educators are encouraged to incorporate adaptive digital platforms, retrieval-based review activities, and contextual vocabulary exposure into classroom instruction. Policymakers should also support teacher training programs and digital learning integration to strengthen vocabulary instruction in secondary education. Future research is recommended to involve larger samples, longer intervention periods, and mixed-method approaches to

explore the long-term impact of adaptive vocabulary learning across different educational contexts and language skills.

6. References

Ahmed, S., & Pawar, N. (2025). Technology-assisted vocabulary instruction and productive vocabulary development among adolescent EFL learners. *Journal of Language and Education*, *11*(1), 45–61. https://doi.org/10.17323/jle.2025

Alzahrani, M., & Madini, A. A. (2023). Traditional Memorization Versus Adaptive Vocabulary Instruction in EFL Classrooms. *Journal of Language Teaching and Research*, *14*(3), 744–753. https://doi.org/10.17507/jltr.1403.18

Busran, N. S., Surahmat, Z., Jamaliah, M. N., & Majid, A. R. (n.d.). *PENERAPAN PERMAINAN ULAR TANGGA DIGITAL UNTUK MENINGKATKAN PENGUASAAN KOSAKATA BAHASA INGGRIS ANAK USIA DINI*.

Chen, Y., & Hsu, L. (2021). Interactive digital vocabulary learning and lexical retention among EFL students. *Computer Assisted Language Learning*, *34*(8), 1045–1065. https://doi.org/10.1080/09588221.2020

Fitri, E., & Sinaga, L. (n.d.). A Systematic Examination of Digital Technology Integration in English Language Teaching at Vocational High Schools: Challenges and Opportunities Rahmaniar IAI DDI Sidenreng Rappang. *Jurnal Pendidikan Bahasa*, *13*(1).

Husnaini, H. (2022). Development of Self Esteem-Oriented Micro Teaching Materials for IAIN Palopo English Education Students. IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature, 10(1), 538-560.

Ismayanti, D., & Syam, A. T. (2022). The Community Manipulation through Big Brotherâ€™ s Tyranny in George Orwellâ€™ s Nineteen Eighty-Four. IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature, 10(2), 1556â-1569.

Kim, H., & Lee, J. (2021). Adaptive feedback systems and vocabulary learning efficiency in EFL classrooms. *Educational Technology Research and Development*, *69*(5), 2761–2780. https://doi.org/10.1007/s11423-021

Li, R., & Hafner, C. A. (2022). Immersive digital environments and vocabulary acquisition in second language learning. *Language Learning & Technology*, *26*(3), 78–95.

Mohammadi, M. (2024). Mobile-Assisted Vocabulary Learning and Productive Vocabulary Development Among EFL Learners. *Education and Information Technologies*, *29*(2), 2145–2163. https://doi.org/10.1007/s10639-023-11874-6

Muqaibal, M. H., Kasprowicz, R., & Tissot, C. (2023). Evaluating the impact of spaced practice using computer-assisted language learning (CALL) on vocabulary learning in the classroom. *Language Teaching Research*. https://doi.org/10.1177/13621688221146146

ISSN 2338-4778 (Print)
ISSN 2548-4192 (Online)

Nation, I. S. P. (2013). *Learning Vocabulary in Another Language* (2nd ed.). Cambridge University Press.

Nguyen, T. H., & Tsai, Y. R. (2024). Measuring Delayed Vocabulary Retention in Technology-Enhanced EFL Learning. *Language Learning and Technology*, *28*(1), 55–74.

Qian, D. D. (1999). Assessing the Roles of Depth and Breadth of Vocabulary Knowledge in Reading Comprehension. *Canadian Modern Language Review*, *56*(2), 282–308. https://doi.org/10.3138/cmlr.56.2.282

Rahman, F., Suryani, D., & Akbar, M. (2023). Spaced retrieval practice and long-term vocabulary retention among secondary EFL learners. *Journal of Educational Computing Research*, *61*(4), 913–931.

Rahmaniar, R., Surahmat, Z., Sardi, A., & Nurnaifah, I. I. (2024). Challenge and Opportunities: A Qualitative Exploration of Junior High School English Language Educators' Perspectives on Implementing Differentiated Instruction. *JELITA*, *5*(1), 28-40. https://doi.org/10.56185/jelita.v5i1.597.

Teng, M. F. (2023). The Effects of Multimedia-Assisted Vocabulary Learning on EFL Learners' Retention and Productive Knowledge. *Computer Assisted Language Learning*, *36*(5–6), 901–923. https://doi.org/10.1080/09588221.2021.1933327

Terai, M., Yamashita, J., & Pasich, K. E. (2021). Effects of Learning Direction in Retrieval Practice on EFL Vocabulary Learning. *Studies in Second Language Acquisition*, *43*(5), 1023–1045. https://doi.org/10.1017/S0272263121000124

Wang, X., & Chen, Z. (2024). Gamified adaptive vocabulary learning and student engagement in EFL classrooms. *Education and Information Technologies*, *29*(2), 2331–2350. https://doi.org/10.1007/s10639-024

Webb, S., & Nation, I. S. P. (2017). *How Vocabulary Is Learned*. Oxford University Press.

Xu, Y. (2022). An Adaptive Learning System for English Vocabulary Using Machine Learning. *Mobile Information Systems*, *2022*, 1–11. https://doi.org/10.1155/2022/4562187

Zhao, D. (2024). The impact of AI-enhanced natural language processing tools on writing proficiency: an analysis of language precision, content summarization, and creative writing facilitation. *Education and Information Technologies*, *30*(6), 8055–8086. https://doi.org/10.1007/s10639-024-13145-5