KURAS Institute
Journal of Arabic Education, Linguistics and Literature Studies
journal homepage: https://journal.kurasinstitute.com/index.php/jicallss

The Effect of Zep Quiz on Improving Arabic Vocabulary Mastery Among Seventh-Grade Students at Central Jakarta Islamic Junior High School

Aṡaru Istikhdāmi Taṭbīqi Zep Quiz fī Taḥsīni Itqāni al-Mufradāti al-'Arabiyyati Ladā Tilāmīżi aṣ-Ṣaffi as-Sābi'i bil-Madrasati al-Mutawassiṭati al-Islāmiyyati al-Ḥukūmiyyati bi Jākartā al-Markaziyyah

Abdulloh Faiz Al Falah*, Muhammad Qadhafi, M. Rafif
Universitas Islam Jakarta, Indonesia
*Corresponding E-mail: alfalahfaiz76@gmail.com
DOI: https://doi.org/10.51214/jicalls.v4i1.2126

Abstract

Mastery of Arabic vocabulary is a foundational component of language learning, since lexical knowledge supplies the resources learners need to comprehend and produce meaningful language. Despite this importance, vocabulary instruction in many classrooms still relies heavily on repetition and rote memorization. Such practices tend to position students as passive recipients and offer limited opportunities for repeated retrieval, which is a process that strengthens long-term retention. Gamification-based digital platforms may address this limitation by turning practice into an interactive and goal-oriented activity. Against this background, this study examined the effect of Zep Quiz, a gamification-based digital learning platform, on Arabic vocabulary mastery among seventh-grade students at an Islamic junior high school (Madrasah Tsanawiyah) in Central Jakarta. A quantitative approach with a pre-experimental One-Group Pretest-Posttest Design was adopted. Participants comprised all 36 students of Class VII-A, who were selected through total sampling. Vocabulary mastery was measured with a 20-item Arabic vocabulary achievement test administered before and after the treatment. The resulting data were analyzed descriptively and then inferentially with a paired-samples t-test. The descriptive findings showed that the mean score rose from 60.14 on the pretest to 78.75 on the posttest, a gain of 18.61 points. Inferential analysis confirmed that this gain was unlikely to be due to chance: the paired-samples t-test yielded $|t| = 7.132$ with $df = 35$ and Sig. (2-tailed) = 0.000, and the calculated value exceeded the critical t-value of 2.030 at $\alpha = 0.05$. Because $|t| = 7.132 > 2.030$ and $p < 0.05$, the null hypothesis was rejected, indicating a statistically significant improvement in students' Arabic vocabulary mastery following instruction with Zep Quiz. These findings are consistent with the broader literature reporting that gamified and digital game-based learning can support vocabulary achievement, learner engagement, and repeated retrieval. Beyond confirming this pattern, the study contributes contextual evidence from a seventh-grade madrasah classroom in Central Jakarta. It also suggests that Zep Quiz can serve as a complementary medium alongside conventional Arabic vocabulary instruction.

Keywords: Arabic Vocabulary; Digital Learning; Gamification; Pre-Experimental Design; Vocabulary Mastery; Zep Quiz.

ملخّص

يُعَدُّ إتقانُ المفردات العربية مكوِّنًا أساسيًّا في تعلُّم اللغة، إذ توفّر المعرفةُ المعجميةُ الموارِدَ اللغويةَ التي يحتاج إليها المتعلّمون لفهم اللغة وإنتاجها بصورة ذات معنى. وعلى الرغم من هذه الأهمية، لا يزال تعليم المفردات في كثير من الفصول الدراسية يعتمد

© 2026 Abdulloh Faiz Al Falah, Muhammad Qodhafi, M. Rafif
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

اعتمادا كبيرا على التكرار والحفظ الآلي. وتميل هذه الممارسات إلى جعل الطلاب متلقّين سلبيين للمعلومات، كما أنها توفّر فرصا محدودة للاسترجاع المتكرر للمعلومات، وهي عملية تسهم في تعزيز الاحتفاظ بالمعلومات على المدى الطويل. ويمكن للمنصات الرقمية القائمة على التلعيب أن تسهم في معالجة هذا القصور من خلال تحويل التدريب إلى نشاط تفاعلي موجَّه نحو تحقيق الأهداف. وانطلاقًا من هذه الخلفية، هدفت هذه الدراسة إلى الكشف عن أثر استخدام منصة Zep Quiz ، وهي منصة تعليمية رقمية قائمة على التلعيب، في إتقان المفردات العربية لدى تلاميذ الصف السابع في إحدى المدارس المتوسطة الإسلامية الحكومية بجاكرتا المركزية .اعتمدت هذه الدراسة على المنهج الكمي باستخدام التصميم شبه التجريبي من نوع المجموعة الواحدة ذات الاختبار القبلي والاختبار البعدي .(One-Group Pretest-Posttest Design) وتكونت عينة الدراسة من جميع تلاميذ الصف السابع (أ)، وعددهم ٣٦ تلميذًا، تم اختيارهم باستخدام أسلوب العينة الشاملة .(Total Sampling) وتم قياس مستوى إتقان المفردات العربية من خلال اختبار تحصيلي مكون من ٢٠ سؤالًا، أُجري قبل المعالجة التعليمية وبعدها. وحلّلت البيانات الناتجة باستخدام التحليل الإحصائي الوصفي، ثم التحليل الاستدلالي من خلال اختبار (ت) للعينات المترابطة-(Paired Samples T-Test) وأظهرت النتائج الوصفية ارتفاع متوسط درجات التلاميذ من ٦٠,١٤ في الاختبار القبلي إلى ٧٨,٧٥ في الاختبار البعدي، بزيادة قدرها ١٨,٦١ درجة. وأكد التحليل الاستدلالي أن هذه الزيادة ذات دلالة إحصائية، حيث بلغت قيمة (ت) المحسوبة ٧,١٣٢- بدرجات حرية (df) تساوي ٣٥، وبلغت قيمة الدلالة الإحصائية (Sig. 2-tailed) مقدار ٠,٠٠٠، كما تجاوزت القيمة المحسوبة قيمة (ت) الجدولية البالغة ٢,٠٣٠ عند مستوى الدلالة .(α = 0.05) وبما أن القيمة المطلقة لـ (ت) المحسوبة (٧,١٣٢) أكبر من قيمة (ت) الجدولية (٢,٠٣٠)، وأن قيمة (p) أقل من ٠,٠٥، فقد تم رفض الفرضية الصفرية، مما يشير إلى وجود تحسن ذي دلالة إحصائية في مستوى إتقان المفردات العربية لدى التلاميذ بعد استخدام منصة Zep Quiz في عملية التعليم. وتتفق هذه النتائج مع الأدبيات العلمية السابقة التي تشير إلى أن التلعيب والتعلّم القائم على الألعاب الرقمية يمكن أن يساهما في تحسين التحصيل في المفردات، وتعزيز مشاركة المتعلمين، وتوفير فرص الاسترجاع المتكرر للمعلومات. وإلى جانب تأكيد هذا الاتجاه، تقدّم الدراسة أدلة سياقية من بيئة تعليمية في إحدى المدارس المتوسطة الإسلامية الحكومية للصف السابع بجاكرتا المركزية. كما تشير نتائجها إلى إمكانية توظيف منصة Zep Quiz بوصفها وسيلة تعليمية مساندة إلى جانب الأساليب التقليدية في تعليم المفردات العربية.

الكلمات المفتاحية: التلعيب; المفردات العربية; إتقان المفردات; التعلّم الرقمي; التصميم شبه التجريبي; منصة Zep Quiz

INTRODUCTION

Lexical knowledge underpins every dimension of language proficiency. Listening, speaking, reading, and writing all draw on a shared repertoire of words, and learners with limited vocabulary often find that comprehension stalls and self-expression remains halting, regardless of how well they have mastered grammar Vocabulary is therefore regarded not merely as one component of language competence but as its foundation, since a learner cannot decode a text, follow a conversation, or construct a coherent sentence without sufficient command of words and their meanings. In Arabic, this dependence is reinforced by rich morphology: each new lexical item *(mufradat)* often opens onto a family of related forms derived from a shared root, so early accumulation of core vocabulary shapes how efficiently learners can absorb later material. This dependence is sharper still in Indonesian madrasahs, where Arabic functions as a foreign language, exposure outside school is minimal, and classrooms are often the main setting where learners regularly encounter new words.

Vocabulary teaching methods therefore have a lasting influence on what students can later accomplish with Arabic. Field evidence, however, suggests that this responsibility is not being met satisfactorily. A preliminary test at MTs Jakarta Pusat showed that 36% of seventh-grade students scored within the failing range (0–59), while only 19% reached a satisfactory level. Classroom observation and a teacher interview identified instructional practice as a principal cause: teachers delivered word lists, which students then committed to memory through repetition, a teacher-centered routine that offers little room for active participation, tends to reduce interest and motivation, and yields knowledge that fades quickly once tests are over. Meanwhile, students already carried smartphones and had internet connectivity, yet these tools were rarely drawn into lessons. A pedagogical opportunity was available but remained unused.

Educational technology offers a plausible route out of this impasse, and gamification has attracted particular attention.[1] The term refers to the deliberate embedding of game mechanics, such as points, leaderboards, challenges, and instant feedback, within activities that are not games in themselves, with the aim of sustaining learner engagement The empirical record is encouraging.[2] Meta-analytic and systematic evidence reports a positive effect on cognitive learning outcomes, although the magnitude varies with the way game elements are designed and combined.[3] This qualification matters, because it implies that the benefits of gamification are not automatic and must be examined for each platform and each instructional setting.[4]

Studies in Arabic language teaching have begun to test this proposition. Wordwall was associated with better learning outcomes among eighth-grade students and Kahoot was reported to strengthen vocabulary by encouraging friendly rivalry within the class More broadly, interactive media have been found to capture attention and deepen understanding Read together, these findings point toward the potential of gamified tools to surpass conventional delivery. Yet the literature is concentrated on a small group of familiar platforms, namely Quizizz, Kahoot, and Wordwall, which leaves more recent tools largely unexplored.

Zep Quiz is one such recent tool. It allows teachers to host quizzes in virtual rooms through several game modes, among them Sequencing, Freestyle, and Round Match, and it combines these with a point system, rankings, and immediate feedback.[5] Scholarly attention

[1] Devita Meli, Arten H Mobonggi, and Alfian Erwinsyah, "Pengaruh Kompetensi Pedagogik Terhadap Minat Belajar Siswa" 7 (2019): 71–85.

[2] Abdullah Syarofi, Abdul Mujib Murtadho, and Khulafaur Rosyidin, "The Effectiveness of ZEP Quiz-Based Educational Games in Enhancing Students ' Motivation to Learn Arabic" 4, no. 1 (2026): 621–30, https://doi.org/10.37680/ssa.9454.

[3] Abdullah Al Hejaili and Paul Newbury, "LAA : Learn the Arabic Alphabet : Integrating Gamification Elements with Touchscreen Based Application to Enhance the Understanding of the Arabic Letters Forms" 21, no. 4 (2023): 353–65.

[4] Muhammad Fahmi Ardiansyah, Nur Sjamsuarini, and Pudji Astutik, "Implementasi Media Pembelajaran Platform Blooket Terhadap Pemahaman Pembelajaran Bahasa Arab" 5, no. 3 (2025): 1008–20.

[5] Afiyatul Maulidiyah et al., "Efektivitas Gamifikasi Dengan Zep Quiz Dalam Meningkatkan Penguasaan Kosakata Bahasa Arab" 8 (2025): 12038–44.

to the platform remains scarce, and what exists has gravitated toward affective outcomes. Found that Zep Quiz heightened students' motivation to learn Arabic at a vocational high school, and reported a comparable motivational effect in fiqh lessons at the primary level.[6] The closest precedent is, who applied a quasi-experimental one-group pretest-posttest design at an Islamic junior high school in Sidoarjo and recorded a significant improvement in eighth-grade students' Arabic vocabulary.[7] To date, this is the strongest evidence supporting the use of Zep Quiz for Arabic vocabulary.

Even so, three gaps are evident. The first is conceptual. Because most work on Zep Quiz has examined motivation, evidence on cognitive outcomes is limited and, in the Arabic context, rests on a single investigation. The second concerns the depth of that investigation. acknowledged a small sample, a brief intervention, and the lack of a control group, and they did not explain how the individual game modes might be ordered within vocabulary teaching, which limits the usefulness of their findings for practitioners.[8] The third is contextual. The available evidence originates from provincial, vocational, and primary settings, whereas no study has focused on seventh-grade students at an urban Islamic junior high school (MTs) in Central Jakarta, where learners come from varied socioeconomic backgrounds and enjoy digital access that is seldom directed toward academic ends. These gaps justify further empirical work.

This study addresses them in three ways. First, it assesses vocabulary through a test, so that the outcome reflects cognitive attainment rather than attitude. Second, it broadens the evidence base to a group not previously examined, namely seventh-grade MTs students in an urban Central Jakarta setting. Third, it documents the use of three Zep Quiz modes (Sequencing, Freestyle, and Round Match) as a structured instructional procedure that other teachers can replicate. By doing so, it contributes to the thin body of research on Zep Quiz in Arabic language teaching and offers concrete guidance for teachers who wish to introduce gamified media into their vocabulary lessons. Accordingly, the objective of this study is to determine the effectiveness of Zep Quiz in improving Arabic vocabulary mastery among seventh-grade students at MTs Jakarta Pusat.

METHODS

A quantitative approach with a pre-experimental design[9] was used to examine changes in students' Arabic vocabulary mastery following Zep Quiz implementation. A One-Group Pretest-Posttest Design was applied, involving one class without a control group. Participants consisted of 36 seventh-grade students from Class VII-A at MTs Jakarta Pusat, selected

[6] Rini Musrifa and Ida Ayu, "7 Th Grade Students ' Perceptions of Zep Quiz for Gamified Vocabulary Learning" 11, no. 4 (2026): 371–78.

[7] Maulidiyah et al., "Efektivitas Gamifikasi Dengan Zep Quiz Dalam Meningkatkan Penguasaan Kosakata Bahasa Arab."

[8] Moch. Imam Machfudi and Anissa'ul Afidah, "Students' Difficulties in Vocabulary Mastery," *CREW Journal* 1, no. 1 (2022): 01–12, https://doi.org/10.35719/crewjournal.v1i1.1359.

[9] TingLan Ma and Yen Lee, "Overview of Quantitative Research," *Family Medicine* 58, no. 2 (2026): 81–87, https://doi.org/10.22454/FamMed.2026.406133.

through total sampling. Data were collected through a vocabulary test administered before and after instructional treatment.

The test consisted of 20 items covering vocabulary recognition, meaning recall, and understanding of Arabic words. Both tests used identical assessment materials to measure changes in students' performance.[10] During treatment, students attended nine instructional meetings, each lasting approximately one hour. Every meeting introduced five to ten new Arabic vocabulary items taken from learning materials used in class. Zep Quiz supported vocabulary practice through Sequencing, Freestyle, and Round Match modes. Students accessed activities using smartphones or computers, completed vocabulary exercises, received immediate feedback, and viewed their scores and rankings. Data analysis included descriptive statistics (mean scores and standard deviations), followed by a paired-samples t-test to assess differences between pretest and posttest results. Statistical analysis was conducted using SPSS, with a significance level of 0.05. Results from both tests were compared to identify changes in students' Arabic vocabulary mastery after participation in Zep Quiz activities.

RESULT AND DISCUSSION

1. The Effectiveness of Zep Quiz in Improving Arabic Vocabulary Mastery

Hypothesis testing and main findings**.** results support the primary hypothesis: the use of Zep Quiz was associated with a statistically significant improvement in Arabic vocabulary mastery among seventh-grade students. Mean scores rose from 60.14 (SD = 18.88) on the pretest to 78.75 (SD = 16.36) on the posttest, a gain of 18.61 points. The paired-samples t-test indicated that this difference was significant, $|t| = 7.132$, $df = 35$, Sig. (2-tailed) = 0.000, with the absolute t-value exceeding the critical value of 2.030 at $\alpha = 0.05$. Accordingly, the null hypothesis (H_o) was rejected and the alternative hypothesis (H_1) was accepted. Further evidence came from the 95% confidence interval of the mean difference (−23.91 to −13.31), which excluded zero and indicated that the direction of change was consistent and unlikely to result from sampling error alone. Effect size analysis also showed a large standardized effect (Cohen's $d_z \approx 1.19$), exceeding the conventional threshold of 0.80.

Table 1. Descriptive Statistics of Students' Arabic Vocabulary Scores Before and After the Intervention

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	VAR00001	**60.1389**	**36**	**18.87879**	**3.14647**
	VAR00002	**78.7500**	**36**	**16.36089**	**2.72681**

[10] Laura M. O'Dwyer and James A. Bernauer, *Quantitative Research for the Qualitative Researcher* (SAGE Publications, 2013).

Table 2. Results of the Paired-Samples t-Test for Arabic Vocabulary Achievement

			Paired Samples Test		
			Paired Differences		
		Mean	Std. Deviation	Std. Error Mean	95% ConfidenceInterval of the Difference
					Lower
Pair 1	VAR00001 - VAR00002	-18.61111	15.65755	2.60959	-23.90886

Table 3. Comparison of Students' Mean Arabic Vocabulary Scores Before and After the Intervention

Paired Samples Test					
		Paired Differences			
		95% Confidence Interval of the Difference			
		Upper	t	df	Sig. (2-tailed)
Pair 1	VAR00001 - VAR00002	-13.31336	-7.132	35	.000

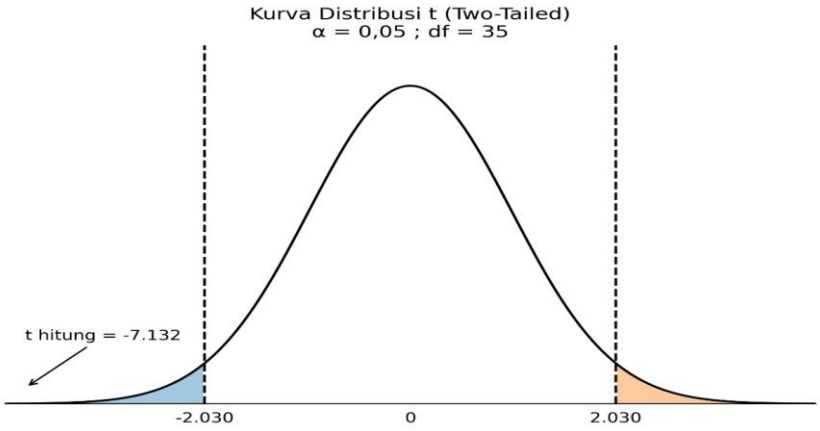


Figure 1. Changes in the Distribution of Students' Arabic Vocabulary Achievement Before and After the Intervention

Statistical significance alone says little about practical meaning, so the distribution of scores is more informative. The proportion of students scoring below 60 fell from [39%] to 11%, while the proportion in the highest band (90–100) rose from [3%] to 36%. This pattern suggests that the gain was not driven by a small group of high performers, but reflected movement across the score range. Students at the lower end of the distribution also benefited, which is particularly relevant in vocabulary learning, where weak initial mastery often hinders progress in reading and comprehension. One plausible explanation is that Zep Quiz provides repeated retrieval practice, immediate feedback, and game-based engagement, all of which can sustain attention and motivation during vocabulary drills. Because the present study did not measure these mechanisms directly, this explanation is best treated as an interpretation rather than a demonstrated cause.

Caution is also needed in interpreting the magnitude of improvement. A pretest–posttest comparison cannot fully rule out other influences, such as maturation, familiarity with test items, or concurrent classroom instruction. For this reason, the findings are best interpreted as evidence of a positive association between Zep Quiz use and vocabulary gains, rather than proof of a causal effect. Taken together, these findings address the research question and indicate that Zep Quiz has the potential to be an effective tool for improving Arabic vocabulary mastery in this setting. Future studies using a control group or a quasi-experimental design would strengthen the basis for generalizing these results.

2. Discussion of Findings in Relation to Previous Studies

Several features of the intervention offer a plausible, though untested, explanation. First, each practice session functioned as a retrieval task with immediate feedback, which allowed students to detect and correct errors while the new words were still fresh in memory. Retrieval of vocabulary items from memory, as opposed to passive rereading or listening, is likely to have been especially useful for Arabic words that carry unfamiliar phonological and morphological forms, and this kind of tends to strengthen long-term retention more than passive re-reading.[11] Second, the three game modes, Sequencing (التسلسل), Freestyle (اللعب الحر), and Round Match (مباراة الجولة), required students to recall the same lexical items under different task formats. Such variation in encoding conditions is likely to deepen word knowledge by linking each item to more than one retrieval context.[12] Third, nine weekly meetings, each introducing about ten textbook words, created a pattern of spaced and cumulative exposure, a schedule that tends to favor retention over massed practice.[13] The game and competitive elements may also have helped sustain attention during otherwise repetitive drills, although this study did not measure engagement or motivation directly. These mechanisms are therefore best read as likely pathways rather than demonstrated

[11] Baiq Tuhfatul, "Media Gambar Dalam Pembelajaran Kosakata Bahasa Arab 2, no. 1 (n.d.): 26–44.

[12] Rizal Wahid et al., "Zep Quiz in Action : Empowering Students ' Vocabulary through Gamified Learning" 4778 (2025): 5461–73, https://doi.org/10.24256/ideas.

[13] Rahmah Fitriah, "Improving The Eighth Grade Students' Ability In Writing Narrative Text By Using Unfinished Picture Series At Mtsn Model Amuntai," N.D.

causes. A further point deserves caution. The standard deviation declined only slightly (18.88 to 16.36), which might be taken as evidence of evenly distributed gains, but this difference was not tested. Moreover, the standard deviation of the individual gains, which can be back-calculated from the reported t-value and sample size at roughly 15.7 points [verify in SPSS output], is close to the pretest spread. This suggests that students benefited to quite different degrees, so the intervention may have worked better for some learners than for others. Individual-level analysis, such as N-Gain by initial proficiency, would be needed to clarify this.

Comparison with previous studies. Results obtained in this study are consistent with previous studies reporting that game-based learning platforms, such as Quizizz, Kahoot, and Wordwall, can contribute to improvements in vocabulary acquisition and retention. Previous research has demonstrated that integrating these platforms into vocabulary instruction can positively influence learners' ability to recall and retain new vocabulary. Such findings indicate that game-based activities provide opportunities for students to engage with vocabulary through interactive exercises, repeated practice, and immediate feedback. These learning processes are particularly relevant to vocabulary mastery because acquiring new words requires continuous exposure, active recall, and meaningful associations between lexical forms and their corresponding meanings.

In relation to findings of present research, improvement in students' Arabic vocabulary mastery following implementation of Zep Quiz suggests that interactive digital activities can support vocabulary instruction. Through quiz-based exercises, students are encouraged to identify, recall, and understand Arabic vocabulary while participating in structured learning activities. Game elements, including scores, challenges, and leaderboards, may also stimulate student participation and maintain attention during classroom instruction. Such features are relevant to Arabic vocabulary learning because students need repeated opportunities to recognize unfamiliar lexical forms, remember their meanings, and understand their use in appropriate contexts. Accordingly, results indicate that Zep Quiz has potential as an instructional medium for supporting Arabic vocabulary mastery among seventh-grade students. Nevertheless, interpretation of these results should consider instructional conditions, selection of vocabulary materials, alignment between quiz activities and learning objectives, and opportunities for repeated practice.[14] Two aspects distinguish present study from earlier research on Zep Quiz. Previous studies concentrated mainly on learner motivation, examining how game-based digital platforms influenced students' interest, enthusiasm, and participation during learning activities. Although motivation represents an important factor in the learning process, its development does not necessarily demonstrate improvements in students' actual academic achievement, particularly in vocabulary mastery. Consequently, further investigation is needed to examine how Zep Quiz contributes to measurable learning outcomes, especially in Arabic language instruction. This distinction is important because vocabulary mastery involves several cognitive processes, including

[14] Desal Putri Ramdhani et al., "Kalamuna : Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban Game-Assisted Multimodal Learning Efficacy on Arabic Vocabulary Mastery : A Pre-Experimental Study in Indonesian Islamic Elementary Education" 07, no. 1 (2026): 139–52.

recognizing lexical forms, remembering word meanings, understanding relationships between words and their meanings, and applying vocabulary appropriately within simple linguistic contexts.

Present study addresses this research gap by focusing specifically on Arabic vocabulary mastery among seventh-grade students at MTs Jakarta Pusat. Unlike previous research that primarily investigated motivational aspects, this study examines changes in students' vocabulary achievement through pretest and posttest measurements following implementation of Zep Quiz. This approach provides empirical evidence concerning students' vocabulary development by comparing their achievement before and after participating in technology-assisted learning activities. Furthermore, focusing on seventh-grade students provides a different instructional context in which learners are developing foundational Arabic vocabulary knowledge as part of their initial experience with more structured language learning at junior secondary school level. In this context, Zep Quiz offers opportunities for students to practice vocabulary recognition and recall through interactive questions and game-based exercises. Accordingly, the present study extends previous discussions of Zep Quiz by shifting attention from motivational outcomes toward measurable improvements in Arabic vocabulary mastery. This contribution provides additional insight into potential applications of game-based digital learning media in supporting vocabulary instruction within Islamic junior secondary education. This study examined cognitive vocabulary mastery through a performance test, which addresses the conceptual gap identified in the introduction. In addition, to our knowledge, it is the first to examine Zep Quiz with seventh-grade students in an urban Islamic junior high school (MTs) in Central Jakarta, which addresses the contextual gap. Unlike platforms built around a single quiz format, Zep Quiz combines quiz questions with several game modes in one environment, a feature that may allow more varied practice[15]Where the size of the gain departs from that of other studies , differences in test difficulty, intervention length, and learner characteristics are the most probable explanations, and direct comparison should remain tentative until instruments and designs are better aligned.[16]

3. Implications, Limitations, and Future Research

Several factors qualify these conclusions. Because the one-group pretest–posttest design does not involve a control group, potential influences of maturation, history, and testing effects cannot be completely ruled out. Maturation refers to changes in students' abilities, knowledge, or learning development that may occur naturally over time, regardless of instructional intervention. In this study, students may have improved their Arabic vocabulary knowledge through regular classroom instruction, independent learning, or increased familiarity with previously introduced vocabulary. History effects may also have influenced research outcomes through external learning experiences occurring during the

[15] Ida Laila and Ema Pariati, "Pengembangan Tes-Tes Hasil Belajar" 2, no. 2 (2024).

[16] Azmy Ali Muchtar et al., "ةيبرعلا ةغللا تادرفم ملعت يف بلاطلا تابوعص دلع بلغتلا يف ملعلما تايجيتارتسا ةراطنب ةيملاسلإا ةئيادتبلإا ةمكحلا ةسردلماب" ٥, no. 1 (2026): 55–73.

intervention period, such as additional learning activities, assignments, or exposure to Arabic vocabulary outside classroom sessions. Furthermore, testing effects may have contributed to differences between pretest and posttest scores because students had previously encountered similar questions, vocabulary items, or assessment formats. Familiarity with testing procedures may have enabled students to respond more effectively during posttest administration.

These considerations are particularly important when interpreting improvements in students' Arabic vocabulary mastery following implementation of Zep Quiz. Although statistical analysis demonstrated a significant difference between pretest and posttest scores, observed improvement cannot be attributed exclusively to Zep Quiz because other educational and contextual factors may have contributed to learning outcomes. In addition, repeated exposure to vocabulary materials during regular instruction may have strengthened students' ability to recognize and recall previously learned words. Consequently, results should be interpreted as evidence of improvement following implementation of Zep Quiz within a particular instructional context, rather than definitive evidence of an independent causal effect. Future research could address these limitations by employing a quasi-experimental design involving an experimental group and a control group, preferably with comparable initial characteristics. Such an approach would provide a stronger basis for distinguishing improvements associated with Zep Quiz from changes resulting from regular instruction, external learning experiences, and repeated assessment.[17]

Part of the observed gain may therefore reflect regular classroom instruction or students' growing familiarity with assessment procedures, and effect size should be interpreted with these considerations in mind. During intervention period, students continued to participate in regular Arabic language instruction, which may have contributed to their developing vocabulary knowledge independently of Zep Quiz activities. Repeated exposure to Arabic vocabulary through classroom explanations, learning materials, exercises, and teacher-guided activities may have strengthened students' ability to recognize and recall previously introduced words. In addition, familiarity with pretest questions and assessment procedures may have influenced posttest performance, particularly when students encountered similar vocabulary items or question formats during both assessments. Such familiarity could have reduced difficulties associated with understanding instructions and responding to test items, potentially contributing to higher posttest scores.

These possibilities are relevant when interpreting magnitude of improvement observed in present study. Although calculated effect size provides an indication of standardized difference between pretest and posttest scores, it does not independently establish how much improvement can be attributed exclusively to Zep Quiz. In a one-group pretest–posttest design, effect size reflects overall changes in measured performance following intervention, which may also incorporate influences from regular instruction, repeated assessment, and

[17] Ramdhani et al., "Kalamuna : Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban Game-Assisted Multimodal Learning Efficacy on Arabic Vocabulary Mastery : A Pre-Experimental Study in Indonesian Islamic Elementary Education."

other learning experiences during research period. Therefore, effect size should be understood as an indicator of magnitude of observed improvement within specific classroom conditions rather than a definitive estimate of isolated treatment effect. Nevertheless, substantial improvement in students' scores provides meaningful evidence that implementation of Zep Quiz was accompanied by positive changes in Arabic vocabulary achievement. Future studies involving control groups and more comprehensive assessment procedures would help clarify extent to which observed improvements are specifically associated with Zep Quiz and distinguish its contribution from other instructional influences.[18] Intervention was delivered directly by researcher, which may have introduced potential expectancy or experimenter bias during implementation of research activities. As both researcher and instructor, researcher was involved in organizing learning activities, explaining vocabulary materials, guiding students through Zep Quiz exercises, and monitoring classroom participation. Such dual involvement may have influenced classroom interactions and students' responses throughout intervention period. For instance, students may have demonstrated greater enthusiasm or participation because they were aware that their learning activities were part of a research project. Similarly, researcher expectations regarding effectiveness of Zep Quiz may have unintentionally influenced instructional emphasis, feedback delivery, or classroom encouragement. Although these actions may have supported students' engagement, they could also have introduced differences between intended instructional procedures and actual classroom implementation.

Furthermore, direct involvement of researcher in delivering intervention may have affected consistency of instructional practices and interpretation of students' learning progress. Researcher expectations could potentially influence how students' difficulties were addressed, how additional explanations were provided, and how learning activities were managed during classroom sessions. Such possibilities are relevant because students' vocabulary achievement may be influenced by instructional attention, teacher encouragement, and classroom interaction in addition to characteristics of digital learning media. Consequently, observed improvements following implementation of Zep Quiz should be interpreted with consideration of possible researcher-related influences. Future research could minimize these potential biases by involving an independent teacher to deliver intervention while researcher focuses on data collection and observation. Additionally, standardized instructional procedures, clearly defined lesson plans, and consistent assessment administration could improve implementation fidelity and strengthen objectivity in evaluating contributions of Zep Quiz to Arabic vocabulary mastery.[19] The 20-item test consisted of the most difficult and most frequently missed items, so it may not represent the full range of vocabulary taught. This could cut both ways: such items leave more room for

[18] S Beth Bierer et al., "Moving Beyond Simplistic Research Design in Health Professions Education : What a One-Group Pretest-Posttest Design Will Not Prove," 2025, 1–9.

[19] Emma Marsden and Carole J Torgerson, "Oxford Review of Education Research Designs : Some Methodological Concerns Single Group , Pre- and Post-Test Research Designs : Some Methodological Concerns," no. December (2012): 37–41.

improvement, but they may also underestimate or overestimate mastery of the broader syllabus. The reliability of the instrument should be reported. The paired-samples t-test also assumes normally distributed score differences. As for implementation, Implementation largely followed the plan of nine one-hour meetings, with students using smartphones or computers. However, the number of new words varied between five and ten per meeting, so exposure was not identical across sessions. Sessions with more new words placed a heavier learning load on students within the same time allocation, which may have reduced the amount of practice available for each word. Because gains were measured only at the level of the total test score, it is not possible to determine whether words introduced in particular sessions were learned better than others. Differences in device access may also have affected participation, although no such barrier was recorded. Variation in connection quality and in the type of device used could likewise have influenced how smoothly students completed the activities, even when these factors were not documented.

Generalizability. Findings come from a single class of 36 students selected through total sampling, and outcomes were measured immediately after the intervention. They should therefore not be generalized to other schools, grade levels, or longer time frames without further evidence. A single class shares the same teacher, schedule, and classroom climate, so features specific to this setting may have shaped the results. Immediate posttesting also leaves open whether the vocabulary gains were retained after the intervention ended. Zep Quiz seems most applicable to classes where students have smartphones or computers and where vocabulary is drawn from a fixed textbook syllabus. Replication in schools with different resources and student backgrounds would show whether similar gains can be expected elsewhere.

Implications and significance**.** Despite these limitations, results offer preliminary support for the view that gamified digital tools such as Zep Quiz can make Arabic vocabulary instruction more effective in Indonesian Islamic junior high schools, where a limited lexicon often constrains the development of other language skills. Because vocabulary underpins reading comprehension, listening, and productive use of the language, even modest gains at this stage may have cumulative benefits for later learning. For practice, findings suggest that teachers may use the platform as a structured, feedback-rich supplement to textbook vocabulary rather than as a replacement for instruction. In this study, the words practiced on the platform were drawn directly from the school's textbook, which indicates that such tools can be aligned with the existing curriculum without requiring additional content or major changes to lesson planning.[20] For future research, these findings indicate several important methodological directions that could strengthen evidence regarding the contribution of digital learning platforms to Arabic vocabulary mastery.

First, researchers should consider employing a control or comparison group to distinguish improvements associated with Zep Quiz from changes that may result from

[20] Mgambi Msambwa, Msafiri Daniel, and Kangwa Lianyu, "A Systematic Literature Review of ICT Integration in Secondary Education : What Works , What Does Not , and What Next ?," *Discover Education*, 2023, https://doi.org/10.1007/s44217-023-00070-x.

regular classroom instruction, students' previous learning experiences, or other external factors. An experimental design involving both experimental and control groups would provide a stronger basis for evaluating whether observed differences in vocabulary achievement are attributable to the instructional intervention. Where possible, random assignment should also be considered to reduce potential selection bias and improve the internal validity of research findings.

Second, future studies should incorporate delayed posttests to examine whether vocabulary gains remain stable after the intervention has ended. Immediate posttest results provide information about short-term learning achievement, whereas delayed assessments can reveal whether students retain newly acquired vocabulary over an extended period. This distinction is particularly important in Arabic vocabulary learning because successful mastery involves not only recognizing word meanings immediately after instruction but also remembering, retrieving, and applying vocabulary in subsequent learning activities. Repeated assessments conducted several weeks or months after treatment would therefore provide a more comprehensive understanding of vocabulary retention and long-term learning outcomes.

Third, future investigations should include direct measurements of students' motivation and engagement to identify the learning processes that may contribute to improvements in vocabulary achievement. Although game-based platforms commonly incorporate interactive activities, immediate feedback, scores, and competitive elements, their influence on students' psychological and behavioral responses should be examined through validated questionnaires, classroom observation, interviews, or engagement scales.

Measuring these dimensions alongside vocabulary test scores would help researchers determine whether increased participation, learning motivation, task persistence, or opportunities for repeated vocabulary practice are associated with improved achievement. Furthermore, examining differences in students' prior vocabulary knowledge, digital literacy, learning preferences, and access to technological devices could clarify which groups of learners benefit most from this instructional approach. Taken together, these methodological improvements would enable future studies to move beyond identifying whether Zep Quiz is associated with improved vocabulary achievement and investigate the mechanisms underlying its potential benefits, the characteristics of learners who respond most positively, and the extent to which learning gains can be maintained over time. Such evidence would provide a stronger empirical foundation for evaluating the pedagogical value and wider implementation of Zep Quiz in Arabic language education[21]. Studies in other regions and school types would also help determine whether the benefits observed here depend on local conditions or hold more broadly.[22]

[21] Akmal Walad, Nurtina Irsad Rusdiani, and Hassan Saleh Mahdi, “Mobile-Assisted Language Learning (MALL) in Arabic Education : Impacts on Vocabulary Acquisition” 3 (2026): 65–74.

[22] Asep Sopian et al., “Arabic Language Learning in a Multicultural Context at Pesantren” 11, no. 1 (2025): 77–89, https://doi.org/10.15575/jpi.v11i1.44104.

CONCLUSION

For classroom practice, Zep Quiz can serve as a feedback-rich supplement to textbook vocabulary instruction rather than a replacement for teacher-led learning. Its three game modes—Sequencing, Freestyle, and Round Match—provide flexible opportunities for repeated vocabulary practice, which teachers can integrate with explicit instruction and productive tasks such as sentence building and short dialogues. The observed gains also suggest that vocabulary learning may benefit from repeated retrieval, immediate feedback, and varied practice contexts, although these mechanisms were not directly measured. Future studies should employ stronger experimental designs, delayed posttests, and process measures to examine learning retention, engagement, and the mechanisms underlying vocabulary gains. Research across different grade levels, school contexts, and digital platforms is also needed to establish the generalizability of these findings. Overall, this study provides preliminary evidence of Zep Quiz's potential to support Arabic vocabulary mastery and highlights the importance of evaluating digital learning tools based on measurable learning outcomes rather than novelty or engagement alone.

REFERENCES

Aljabri, Sameer. “Timing of Feedback and Retrieval Practice: A Laboratory Study with EFL Students,” no. 2024 (n.d.): 1–10.

Ardiansyah, Muhammad Fahmi, Nur Sjamsuarini, and Pudji Astutik. “Implementasi Media Pembelajaran Platform Blooket Terhadap Pemahaman Pembelajaran Bahasa Arab” 5, no. 3 (2025): 1008–20.

Bierer, S Beth, Gary Beck Dallaghan, Nicole J Borges, Sam Brondfield, Cha Chi Fung, N Huggett, Cayla R Teal, Satid Thammasitboon, and Colleen Y Colbert. “Moving Beyond Simplistic Research Design in Health Professions Education : What a One-Group Pretest-Posttest Design Will Not Prove,” 2025, 1–9.

Ferlinna, Vera Sherli. “Efektivitas Penggunaan Media Gambar Diam Dalam Upaya Meningkatkan Penguasaan Kosa Kata Bahasa Arab Siswa Kelas VIII MTs Ma’arif Mandiraja Banjarnegara Jawa Tengah Tahun Ajaran 2014/2015,” 2015, 1–107.

Fitriah, Rahmah. “Mproving The Eight Grade Students’ Ability In Writing Narrative Text By Using Unfinished Picture Series At Mtsn Model Amuntai,” n.d.

Ghafar, Liyana Abdul, and Norah Noor. “Gamification Learning Environment (GLE) for Supporting Self- Regulated Learning in Arabic Vocabulary : A Quasi-Experimental Study .” X, no. 2454 (2026). https://doi.org/10.47772/IJRISS.

Hejaili, Abdullah Al, and Paul Newbury. “LAA : Learn the Arabic Alphabet : Integrating Gamification Elements with Touchscreen Based Application to Enhance the Understanding of the Arabic Letters Forms” 21, no. 4 (2023): 353–65.

Laila, Ida, and Ema Pariati. “Pengembangan Tes-Tes Hasil Belajar” 2, no. 2 (2024).

Machfudi, Moch. Imam, and Anissa’ul Afidah. “Students’ Difficulties in Vocabulary Mastery.” *CREW Journal* 1, no. 1 (2022): 01–12. https://doi.org/10.35719/crewjournal.v1i1.1359.

Ma, TingLan, and Yen Lee. “Overview of Quantitative Research.” *Family Medicine* 58, no. 2 (2026): 81–87. https://doi.org/10.22454/FamMed.2026.406133.

Marsden, Emma, and Carole J Torgerson. “Oxford Review of Education Research Designs : Some Methodological Concerns Single Group , Pre- and Post-Test Research Designs : Some Methodological Concerns,” no. December (2012): 37–41.

Maulidiyah, Afiyatul, M Baihaqi, Hisbullah Huda, Nur Sjamsuarini, and Pudji Astutik. “Efektivitas Gamifikasi Dengan Zep Quiz Dalam Meningkatkan Penguasaan Kosakata Bahasa Arab” 8 (2025): 12038–44.

Meli, Devita, Arten H Mobonggi, and Alfian Erwinsyah. “Pengaruh Kompetensi Pedagogik Terhadap Minat Belajar Siswa” 7 (2019): 71–85.

Msambwa, Mgambi, Msafiri Daniel, and Kangwa Lianyu. “A Systematic Literature Review of ICT Integration in Secondary Education : What Works , What Does Not , and What Next ?” *Discover Education*, 2023. https://doi.org/10.1007/s44217-023-00070-x.

Muchtar, Azmy Ali, Farah Nursaidah, Muhammad Qadhafi, and Islam Jakarta. “ ةيبرعلا ةغللا تادرفم ملعت يف بلاطلا تابوعص بلع بلغتلا يف ملعلما تايجيتارتسا ةراطنب ةيملاسلإا ةيئادتبلإا ةمكحلا ةسردلماب” ٥, no. 1 (2026): 55–73.

Musrifa, Rini, and Ida Ayu. “7 Th Grade Students ’ Perceptions of Zep Quiz for Gamified Vocabulary Learning” 11, no. 4 (2026): 371–78.

O’Dwyer, Laura M., and James A. Bernauer. *Quantitative Research for the Qualitative Researcher*. SAGE Publications, 2013.

Ramdhani, Desal Putri, Agus Yasin, Siti Nurfadlilah, and Andi Nurizzah Aqila. “Kalamuna : Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban Game-Assisted Multimodal Learning Efficacy on Arabic Vocabulary Mastery : A Pre-Experimental Study in Indonesian Islamic Elementary Education” 07, no. 1 (2026): 139–52.

Sopian, Asep, Maman Abdurahman, Yusuf Ali Tantowi, Ani Nur Aeni, and Universitas Pendidikan Indonesia. “Arabic Language Learning in a Multicultural Context at Pesantren” 11, no. 1 (2025): 77–89. https://doi.org/10.15575/jpi.v11i1.44104.

Sucipto, Ahmad, Mulyo Utomo, and Robiyah Nur. “Integrating Active Learning within an AI-Supported Gamified Environment through Zep Quiz to Improve Arabic Language Achievement” 01 (2026).

Syarofi, Abdullah, Abdul Mujib Murtadho, and Khulafaur Rosyidin. “The Effectiveness of ZEP Quiz-Based Educational Games in Enhancing Students ’ Motivation to Learn Arabic” 4, no. 1 (2026): 621–30. https://doi.org/10.37680/ssa.9454.

Tuhfatul, Baiq “Media Gambar Dalam Pembelajaran Kosakata Bahasa Arab no. 1 (n.d.): 26–44.

Wahid, Rizal, Permana Putra, Imam Wahyudi Karimullah, and Mutmainnah Mustofa. “Zep Quiz in Action : Empowering Students ’ Vocabulary through Gamified Learning” 4778 (2025): 5461–73. https://doi.org/10.24256/ideas.

Walad, Akmal, Nurtina Irsad Rusdiani, and Hassan Saleh Mahdi. “Mobile-Assisted Language Learning (MALL) in Arabic Education : Impacts on Vocabulary Acquisition” 3 (2026): 65–74.