ISELT 76
Interdisciplinary Studies in English Language Teaching

Pages 76 to 89

The Effect of a Gamified Mobile Learning Environment on Vocabulary Learning and Retention among EFL Students

Mohsen Bagheri [1]*, Ali Shahsavan [2], Sirus Mansoori [1]
[1] Associate Professor, Educational Science Department, Arak University, Arak, Iran
[2] New Educational Technology Center, Arak University, Arak, Iran

10.22080/iselt.2026.32041.1199

Received
June 6, 2026
Accepted
September 9, 2026
Available online
September 16, 2026

Keywords:
Gamification, Game-Based Learning, English Language Learning, Retention

Abstract

This research was carried out to examine how gamified mobile learning environment influences both the acquisition and long-term retention of vocabulary among university students. The study employed a quasi-experimental design featuring a pretest and posttest alongside a control group. The participants were 34 students from the educational sciences department at Arak University, selected through convenience sampling and randomly divided into an experimental group and a control group. The experimental group received vocabulary instruction through a researcher-developed gamified mobile learning environment, TabTest, which incorporated interactive learning materials, immediate feedback, progress monitoring, points, leaderboards, and course-based competitions, whereas the control group received conventional classroom instruction A researcher-developed test containing 30 questions was used to measure the students' vocabulary of specialized academic texts. The results of an analysis of covariance (ANCOVA) at the 0.05 significance level revealed that gamified mobile learning environment had a meaningful positive effect on both vocabulary learning and retention. After accounting for the pretest scores, significant differences were observed between the experimental and control groups regarding both learning and retention. Additionally, paired-samples t-test results showed that while both groups made progress from the pretest to the posttest and retention test, the improvement in the experimental group was significantly greater. These findings indicate that the incorporation of gamification elements not only enhances immediate vocabulary learning outcomes, but also supports the preservation of learning over time.

1. INTRODUCTION

Learning English continues to be shaped by persistent challenges, particularly in vocabulary acquisition, learner motivation, and the maintenance of learning over time (Ameen, 2026; Ayana et al., 2024). Although vocabulary is central to communicative competence, the large number of

* **Corresponding Author:** Mohsen Bagheri, Associate Professor, Educational Science Department, Arak University, Arak, Iran, **Email**: mohsenbag@gmail.com

words learners need to acquire and the limited amount of English exposure available in many educational settings make vocabulary learning difficult (Biseko, 2025).

Traditional approaches that rely heavily on memorization and isolated word lists have been criticized for producing limited engagement and insufficiently durable learning outcomes (Li et al., 2023; Panmei & Waluyo, 2023). Teacher-centered instruction and repetitive practice may also be associated with reduced motivation and weaker long-term vocabulary development (Chan & Lo, 2024; Zhang & Hasim, 2023). These concerns have increased interest in learner-centered and technology-enhanced approaches that address both cognitive and motivational dimensions of language learning.

Among contemporary educational approaches, gamification has gained attention as a strategy that incorporates game elements—such as points, badges, rankings, levels, challenges, and immediate feedback—into non-game learning contexts (Zeng et al., 2024). Drawing on perspectives such as Self-Determination Theory, Flow Theory, and Constructivist Learning Theory, gamification may support learners' needs for competence, autonomy, and sustained engagement (Laura-De La Cruz et al., 2023).

Gamification has become increasingly prominent in language education because it embeds game mechanics within instructional activities without requiring the entire learning process to become a game (Cheung & Ng, 2021). Advances in digital platforms and mobile-assisted language learning have further expanded opportunities for interactive vocabulary practice (Fithriani, 2021; Xu & Richardson, 2024). Feedback systems, competition, and goal-oriented activities may also influence learners' engagement and learning behaviors (Panmei & Waluyo, 2023).

The growth of gamification in language teaching reflects the broader digital transformation of education and the increasing availability of mobile technologies and learning management systems. Applications such as Duolingo, Kahoot!, Quizizz, and customized platforms can transform routine vocabulary practice into more interactive learning experiences (Laksanasut, 2025; Mleiki, 2026).

This development represents a shift from predominantly passive instruction toward more active and experiential learning environments, an especially relevant development in EFL contexts where learners often have limited opportunities to use English beyond the classroom (Ameen, 2026; Al-Khresheh, 2025).

Empirical evidence suggests that gamification can improve vocabulary learning, motivation, and learner confidence relative to conventional instruction (Li et al., 2023). Gamified environments have also been associated with greater learner satisfaction and sustained engagement with learning materials (Zhang & Yu, 2022). Despite the growing body of research on gamification, important methodological limitations remain. Many interventions are short and assess outcomes immediately after instruction, leaving limited evidence about delayed vocabulary retention (Pratiwi et al., 2024). Capatina et al. (2024) similarly note that although gamification may support learning and knowledge sharing, fewer studies have examined whether its benefits persist over time. In addition, much of the literature relies on general-purpose platforms rather than theoretically informed and context-sensitive gamified environments designed around learners' specific needs.

In light of these limitations, particularly the limited attention given to delayed vocabulary retention and the scarcity of researcher-developed gamified platforms designed for Iranian learners, the present study examined the effects of a mobile-based gamified learning environment on vocabulary learning and retention. A locally developed platform was used to investigate whether a structured gamified intervention could produce benefits beyond immediate post-instruction performance.

2. LITERATURE REVIEW

Growing empirical evidence suggests that gamification can positively influence vocabulary learning by increasing engagement, motivation, and active participation. Experimental studies across educational contexts have generally reported stronger vocabulary outcomes under gamified instruction than under conventional approaches. Li (2021) for example, found that gamified vocabulary instruction improved English vocabulary learning while also enhancing learner confidence and motivation. Similarly, Yu (2023) reported that gamified learning environments promoted sustained engagement and satisfaction, indicating that gamification may support both motivational and instructional dimensions of vocabulary development.

Evidence from technology-supported language learning environments further reinforces these findings. In higher education, gamified digital platforms have been associated with improved vocabulary acquisition, participation, and self-regulated learning. Al-Khresheh (2025) reported positive effects of digital gamification platforms on vocabulary learning and learner enthusiasm. Pratiwi et al. (2024) likewise found that gamified flipped-classroom instruction produced stronger vocabulary outcomes than comparison instruction, while emphasizing that the quality of instructional implementation is more important than the mere presence of game elements. Capatina et al. (2024) similarly argued that gamification supports retention and knowledge sharing when game mechanics are aligned with learning objectives and embedded in coherent instructional design.

Specific gamification features may also strengthen learning experiences. Laura-De La Cruz et al. (2023) reported that avatars, progress-monitoring systems, and structured feedback can support vocabulary learning and enhance learners' educational experiences. Taken together, these findings suggest that learning outcomes depend not simply on whether gamification is present, but on how game mechanics are pedagogically designed and integrated.

Systematic reviews generally support these empirical findings while emphasizing variation across contexts and intervention designs. Zhang and Hasim (2023), Dehghan-Zadeh et al. (2021), Fadzli and Hashim (2026), and Ameen (2026) reported generally positive effects of gamification on vocabulary learning, retention, motivation, and engagement, while also noting that outcomes are shaped by instructional design, learner characteristics, and contextual conditions.

Despite this generally positive evidence, several unresolved issues remain. Chan and Lo (2024) observed substantial methodological inconsistency across studies in intervention design, research methods, and outcome measurement. Cheung and Ng (2021) further argued that although gamification reliably increases engagement, its contribution to deeper learning processes, particularly delayed vocabulary retention, remains insufficiently understood. These observations point to the need for additional empirical studies with clearly defined outcomes and more rigorous designs.

Meta-analytic and bibliometric evidence further illustrates the rapid expansion of this field. Yu (2023) reported an overall positive effect of gamification on motivation and academic performance, while Liu et al. (2025) found positive effects of mobile game-based learning on language outcomes, particularly vocabulary acquisition. Shang (2025) documented substantial growth in gamification research between 2009 and 2024, especially in English language learning, motivation, and digital environments. Together, these syntheses indicate a growing evidence base that still requires theoretically grounded intervention studies with clearly specified learning outcomes.

Taken together, the literature indicates that gamification generally supports vocabulary learning, engagement, and motivation across educational settings. However, important gaps remain. First, many studies emphasize immediate learning outcomes rather than delayed retention.

Second, variation in intervention design, measurement, and instructional context limits direct comparison across studies. Third, relatively few quasi-experimental studies have examined researcher-developed mobile gamification platforms for vocabulary learning and retention in authentic university ESP classrooms. The present study addressed these gaps by examining a researcher-developed gamified mobile learning environment among undergraduate students enrolled in an English for Specific Purposes course.

3. METHODOLOGY AND DESIGN

This study employed a quasi-experimental pretest–posttest design with a control group. The design was selected to compare vocabulary learning and delayed retention under gamified and conventional instructional conditions while statistically adjusting for participants' initial performance.

Although participants were randomly assigned to the experimental and control groups, the study is classified as quasi-experimental rather than fully experimental because the initial pool of participants was obtained through convenience rather than random sampling, and several extraneous variables inherent to a classroom setting (e.g., prior exposure to English materials, individual study habits, and out-of-class practice) could not be experimentally controlled. Intact administrative constraints did not permit assignment from a larger randomly sampled population, which is the defining feature that distinguishes a quasi-experimental design from a true experimental one.

Participants

The participants were undergraduate students majoring in Educational Sciences who were enrolled in the Specialized Texts in Educational Sciences course during the first semester of the 2025–2026 academic year. Thirty-four students were selected through convenience sampling and randomly assigned to an experimental group (n = 17) and a control group (n = 17). All the participants were female and ranged in age from 20 to 22 years.

Instruments

The data were collected using a researcher-developed English for Educational Sciences achievement test designed to assess vocabulary learning and delayed retention. The test contained 30 multiple-choice items measuring knowledge of vocabulary appearing in specialized Educational Sciences texts. Five experienced instructors of specialized Educational Sciences texts reviewed the items for content relevance and essentiality. Content validity was assessed using the Content Validity Ratio (CVR), and all items exceeded a CVR of 0.80.

Reliability was examined separately for the three administrations of the test. Because the items were scored dichotomously, the Kuder–Richardson 20 coefficient was used. The coefficients reported in Table 1 indicate acceptable internal consistency for the pretest, posttest, and retention-test administrations.

Table 1. Examination of test reliability using the Kuder-Richardson 20 coefficient

Kuder-Richardson 20 Coefficient			Variable
Retention	Posttest (Learning)	Pretest	
0.81	0.79	0.83	Learning and retention test

The same content domain was assessed at three time points: pretest, posttest, and retention test. The versions differed slightly in item order and wording to reduce the likelihood of simple item memorization. The retention test was administered one month after the posttest, during which the participants did not have access to the instructional materials or the gamification platform.

Procedure

The participants in the experimental group received instruction through a gamified learning approach, whereas those in the control group received conventional instruction. The instructional content and course materials were identical across groups; the principal difference was the instructional approach.

In the experimental group, the instructor organized each session through a gamification platform. Text, audio, and video materials were uploaded to the platform, and multiple-choice and short-answer questions were created from the instructional content. The students studied the materials, completed the exercises, and answered questions to reinforce learning. The platform recorded the responses and generated performance rankings based on the students' achievement.

The gamification intervention was implemented using the TabTest platform, which was developed by a specialized team at Arak University on the basis of established gamification principles. The platform consisted of an administrative panel and a user application. The administrative panel allowed the instructor to create classes, add users, define course content, and design assessment questions. Each participant received a unique username and password, and the user application included sections for courses, profiles, performance, leagues, and two-player competitions.

Students accessed instructional materials through the course section. Correct answers to embedded questions earned points, whereas incorrect answers generated immediate corrective feedback. The performance section allowed students to monitor progress on a daily, weekly, and monthly basis. Leaderboards displayed relative performance, and two-player competitions enabled students to compete with classmates on course-related content. The league feature allowed the instructor to organize course-based competitions. The instructor could also monitor learning analytics, including accumulated points, correct and incorrect responses, and time spent using the application.

The students had access to the platform and its instructional content outside the classroom and were required to complete the assigned learning activities through the application. During face-to-face sessions, the instructor answered questions, addressed learning difficulties, and conducted a platform-based competition in each session. Competition results and performance information were displayed in the classroom. The experimental group received this instructional treatment for ten 90-minute sessions. The control group received conventional instruction for the same number and duration of sessions. The instructor followed the course syllabus, assigned homework, and conducted classroom discussions and question-and-answer activities related to the instructional content.

Figure 1. Parts of the research implementation steps in the Tap Test gamification environment

As shown in Figure 1, the three panels illustrate, from left to right, the instructor-facing administrative panel used to create classes and assessment questions, the student-facing course and instructional-content interface, and the performance and leaderboard interface through which students tracked points, rankings, and competition results.

The control group received traditional instruction. The instructor taught each lesson according to the course syllabus, assigned homework to be completed outside the classroom, and conducted classroom discussions and question-and-answer sessions related to the instructional content. The duration of instruction for this group was also ten 90-minute sessions

The data were analyzed using IBM SPSS Statistics version 26. An independent-samples t-test was used to examine the baseline differences between the groups on the pretest scores. Paired-samples t-tests were used to examine within-group changes from the pretest to the posttest and from the pretest to the delayed retention test. The primary between-group analyses were conducted using analysis of covariance (ANCOVA), with the pretest scores treated as the covariate. ANCOVA was selected to compare the adjusted posttest and retention scores while accounting for potential initial differences in vocabulary knowledge.

Before the ANCOVA analyses, the assumptions of normality, homogeneity of variances, and homogeneity of regression slopes were examined. The statistical significance was evaluated at the .05 level. The effect sizes were reported using eta squared where available.

4. RESULTS

The data were analyzed using the analysis of covariance (ANCOVA). Before conducting ANCOVA, its underlying assumptions were examined. The normality of score distributions was assessed using the Shapiro–Wilk test. the homogeneity of variances was evaluated using Levene's test, and the assumption of homogeneity of regression slopes was tested by examining the interaction between the covariate (pre-test scores) and the grouping variable. The assumptions were met, supporting the use of ANCOVA. Statistical significance was determined at the .05 level. The results of the assumption checks are shown in Tables 2 through 4.

Table 2: Shapiro-Wilk Test

	Group	Statistic	df	Sig.
Pre-test	Experimental	0.944	17	0.374
	Control	0.956	17	0.553
Post-test	Experimental	0.951	17	0.476
	Control	0.964	17	0.699

Table 3: Testing Homogeneity of Variances

Variable	F statistic	df1	df2	Sig.
Learning	1.547	1	32	0.223
Retention	2.289	1	32	0.140

Table 4: Homogeneity of Regression Slopes

Variable	F statistic	df1	Sig.
Learning	0.497	1	0.486
Retention	0.817	1	0.373

In this study, an independent t-test was used to ensure that there was no significant difference between the two groups at the beginning of the course.

Table 5: Independent-Samples T-Test for the Pretest

Group	N	Mean	Standard Deviation	Sig.	t	df	Sig.	Effect Size
Control	17	5.75	0.687	0.486	0.565	32	0.576	0.27
Experimental	17	5.41	1.651	0.373				

An independent-samples t-test was conducted to compare the pretest scores of the experimental and control groups. The findings show that the two groups did not differ significantly at the outset. Their mean scores suggest comparable baseline performance, and the t-test results support this: the t-value was 0.565 (df = 32, p = 0.576). Since the p-value exceeds 0.05, there is no statistically significant difference between the groups on the pretest. Additionally, the effect size of 0.27 reflects a very small, practically negligible difference at the outset of the study. Overall, we can conclude that the experimental and control groups were equivalent regarding their initial level of the dependent variable, which reinforces the credibility of the results from later analyses .

Table 6: ANCOVA Test for Learning Variable

Source of Variation	Sum of Squares	df	Mean Square	F Value	Sig.	Eta Squared
Corrected Model	56.338	2	28.169	10.426	0.001	0.402
Intercept	573.709	1	573.709	212.338	0.001	0.873
Covariate	3.213	1	3.213	1.189	0.284	0.037
Group	50.048	1	50.048	18.523	0.001	0.374
Error	83.758	31	2.702			

An analysis of covariance was performed on the learning variable. After accounting for the pretest scores, a statistically significant difference emerged between the experimental and control groups. The overall model was significant ($F = 10.426$, $p < 0.001$), suggesting a good model fit. The pretest covariate, however, was not significant ($F = 1.189$, $p = 0.284$), meaning that once included in the model, it no longer made a meaningful contribution to predicting posttest scores. The key finding relates to the group effect, which revealed a significant difference in learning outcomes between the experimental and control groups ($F = 18.523$, $p < 0.001$). The eta-squared value for this effect was 0.374, indicating that the gamification intervention had a relatively strong impact on language learning. Overall, these results demonstrate that gamification-based instruction led to significantly better language learning outcomes compared to traditional teaching methods.

Table 7: ANCOVA Test for Retention Variable

Source of Variation	Sum of Squares	df	Mean Square	F Value	Sig.	Eta Squared
Corrected Model	56.318	2	28.159	9.315	0.001	0.375
Intercept	469.792	1	469.792	155.409	0.001	0.834
Covariate	9.259	1	9.259	3.063	0.090	0.090
Group	42.557	1	42.557	14.078	0.001	0.312
Error	93.711	31	3.023			

An analysis of covariance was conducted on the retention variable. After adjusting for the pretest scores, a significant difference in language retention was found between the experimental and control groups. The overall model was significant ($F = 9.315$, $p = 0.001$), suggesting a good model fit. The pretest covariate, however, was not significant ($F = 3.063$, $p = 0.090$), meaning that once included in the model, it no longer contributed meaningfully to explaining retention scores. The key finding relates to the group effect, which revealed a significant difference in retention between the experimental and control groups ($F = 14.078$, $p = 0.001$). The eta-squared value was 0.312, reflecting a moderate to relatively strong effect of gamification on language retention. Overall, these results demonstrate that gamification-based learning not only enhanced immediate learning, but also produced a significant and lasting effect on language retention.

Table 8: Paired-Samples T-Test for Learning

Group	Pair	Pretest Mean	Posttest Mean	Mean Difference	SD of Difference	t	df	Sig.	Within-Group Correlation (r)
Experimental	Pre-Post	5.73	17.02	-11.29	1.785	-26.075	16	0.001	0.338
Control	Pre-Post	5.41	14.52	-9.11	2.388	-15.740	16	0.001	0.74

A paired-samples t-test was used to examine the within-group changes in learning. The results showed that both the experimental and control groups experienced significant differences between their pretest and posttest scores. In the experimental group, the mean scores increased substantially after the intervention. The t-test yielded a t-value of -26.075 (df = 16, p = 0.001), reflecting a highly significant improvement, with a within-group correlation of 0.338, indicating a moderate association between the two measurements. The control group also showed a significant increase from the pretest to the posttest, with a t-value of -15.740 (df = 16, p = 0.001). The within-group correlation for the control group was relatively high (0.741), suggesting that individual score patterns remained fairly stable over time. In summary, both groups improved significantly from the pretest to the posttest; however, the experimental group showed a much more pronounced gain, which is likely attributable to the gamification intervention.

Table 9: Paired T-Test for Retention

Group	Pair	Pretest Mean	Posttest Mean	Mean Difference	SD of Difference	t	df	Sig.	Within-Group Correlation (r)
Experimental	Pretest Retention	5.75	16.294	-10.55	1.801	-24.163	16	0.001	0.132
Control	Pretest-Retention	5.41	13.94	-8.52	2.273	-15.465	16	0.001	0.348

A paired-samples t-test was conducted to examine the within-group changes in retention. The results showed that both the experimental and control groups had significant differences between their pretest and retention test scores. In the experimental group, the mean scores rose substantially from the pretest to the retention. The t-test produced a t-value of -24.163 (df = 16, p = 0.000), reflecting a highly significant improvement, with a low within-group correlation of 0.132, indicating considerable change in individual score patterns over time. The control group also showed a significant increase from the pretest to retention, with a t-value of -15.465 (df = 16, p = 0.000) and a weak-to-moderate within-group correlation of 0.348, suggesting some relative change in score patterns. In summary, both groups demonstrated significantly higher retention performance compared to their pretest levels. However, the experimental group showed greater improvement, which is likely attributable to the gamification-based intervention, while the improvement observed in the control group may be partly due to practice or the mere passage of time.

Table 10: Statistical power estimates

Variable	Effect Size (f)	Approximate Power
Learning	0.773	> .99
Retention	0.673	≈ .98–.99

In addition to the statistically significant ANCOVA results, the intervention demonstrated a large effect on both outcome variables. According to Cohen's (1988) benchmarks, the obtained effect sizes were large for vocabulary learning ($\eta^2 = .374$) and vocabulary retention ($\eta^2 = .312$). Furthermore, a post hoc power analysis indicated that the statistical power exceeded .98 for both analyses, suggesting that the study had sufficient power to detect the observed group differences.

5. DISCUSSION AND CONCLUSION

The present study examined the effects of a gamified mobile learning environment on vocabulary learning and delayed retention among undergraduate EFL learners. The ANCOVA results showed statistically significant between-group differences in both posttest learning and delayed retention after adjustment for the pretest performance. Although both groups improved over time, the experimental group demonstrated stronger descriptive gains. These findings suggest that a structured gamified mobile learning environment can support vocabulary development and may help sustain learning beyond immediate post-instruction assessment in the context examined.

These findings are broadly consistent with previous research. They align with Li (2021), Yu (2023), Dehghanzadeh et al. (2021), Laura-de la Cruz et al. (2023), and Al-khresheh (2025), that generally reported positive effects of gamification on vocabulary learning, engagement, and motivation. The findings also accord with the systematic reviews of Zhang and Hasim (2023) and Fadzli and Hashim (2026), particularly in suggesting that repeated engagement and feedback can support learning. At the same time, the retention results contribute to the continuing discussion identified by Chan and Lo (2024) and Cheung and Ng (2021) concerning whether gamification can produce learning that persists beyond immediate instruction.

The contribution of the present study extends beyond confirming immediate learning effects. By including a delayed retention assessment, the study provides evidence that the advantage associated with the gamified condition was also observed after the intervention. However, because the retention assessment occurred one month after the posttest and the study used a relatively small, context-specific sample, the findings should be interpreted as evidence of short-term delayed retention rather than definitive evidence of long-term retention.

The present study employed TabTest, a gamified learning platform developed by the researchers for the Iranian educational context. The platform supported course-based competition, one-to-one competition, progress monitoring, performance comparison, immediate feedback, and repeated retrieval opportunities. These features may have contributed to the observed outcomes by increasing opportunities for practice and self-monitoring. Nevertheless, because the intervention combined several features, the present design cannot determine which individual game mechanics were the most influential.

The findings can be interpreted through several complementary theoretical perspectives. From the perspective of Self-Determination Theory, the immediate feedback and progress indicators may have supported learners' sense of competence, while interactive and competitive activities may have encouraged relatedness and sustained engagement. From a Flow Theory perspective, progressively challenging tasks, feedback, and clear performance goals may have supported focused involvement. From a Constructivist Learning Theory perspective, learners

actively constructed knowledge through engagement with vocabulary tasks, feedback, and self-monitoring. These interpretations remain theoretically plausible rather than directly tested in the present study, because motivation, flow, and learner perceptions were not measured as outcome variables.

Beyond these motivational accounts, the delayed retention advantage observed for the experimental group can also be considered from the perspective of memory and learning theory. The gamified environment embedded repeated retrieval opportunities, in that students answered questions on the same material multiple times across sessions and received immediate corrective feedback after each attempt. This pattern resembles retrieval-practice and testing-effect phenomena, whereby actively recalling information strengthens memory traces more effectively than passive re-exposure to content, and the immediate feedback further supports the correction of errors before they are consolidated. The distributed nature of practice across ten sessions may likewise have functioned similarly to spaced-practice arrangements, which are generally associated with more durable retention than massed study. Because the present design did not isolate retrieval frequency, feedback timing, or spacing as independent variables, these memory-based mechanisms remain a plausible rather than a directly tested explanation for the retention findings.

In light of the findings, the present study adds to the growing body of evidence that gamified learning environments can facilitate vocabulary development when game-based features are meaningfully embedded in instruction. Its contribution lies in examining both immediate learning and delayed retention within a researcher-developed mobile environment rather than focusing on a single isolated game mechanic. The results therefore reinforce the view that the educational value of gamification depends substantially on the quality of instructional design and the alignment between game mechanics and learning objectives.

Overall, the findings suggest that gamified mobile learning environments may be useful complementary tools for vocabulary instruction when their mechanics support pedagogically meaningful processes such as retrieval practice, timely feedback, repeated engagement, and progress monitoring. The results should not, however, be interpreted as evidence that gamification is universally effective across educational settings.

Conclusive Implications and Suggestions for Further Studies

The findings of the present study should be interpreted in light of several limitations. First, the participants were undergraduate Educational Sciences students enrolled in an ESP course at one university, which limits generalizability to other disciplines, educational levels, institutions, and cultural contexts. Second, although ANCOVA adjusted for the pretest performance, other potentially relevant variables—including prior English proficiency, motivation, digital literacy, access to technology, and individual study habits—were not directly controlled. These factors may have contributed to the observed outcomes.

Third, the retention test was administered one month after the posttest. This interval provides evidence of the delayed retention, but does not establish the long-term sustainability of the learning gains. Fourth, a novelty effect cannot be ruled out because the experimental group encountered a newly developed gamified environment. Finally, the achievement test focused on receptive vocabulary knowledge related to the instructional objectives and did not assess broader dimensions such as productive vocabulary use, speaking, or writing. Accordingly, the findings should be interpreted as evidence concerning the measured vocabulary outcomes in the specific instructional context examined.

Although the instruments were designed to measure the targeted outcomes, they do not capture broader dimensions of language proficiency, such as productive vocabulary use, speaking, or writing. Accordingly, the findings should be interpreted as evidence concerning receptive vocabulary learning and delayed test performance within the instructional context examined. The statistically significant group differences should, therefore, not be attributed exclusively to gamification because learner-related and contextual factors may also have contributed.

Despite these limitations, the findings offer several practical implications. First, well-designed gamified mobile learning environments may serve as useful complementary tools for vocabulary instruction in university EFL classrooms by providing additional opportunities for repeated practice, active participation, and immediate formative feedback. Second, the results suggest that the educational value of gamification depends more on instructional design than on the presence of isolated game elements. Instructors who adopt gamification should, therefore, align game mechanics with clearly defined learning objectives rather than using them primarily for entertainment or competition.

The findings may also be relevant for curriculum developers, particularly those responsible for ESP instruction in higher education. Carefully designed gamified activities may provide additional opportunities for vocabulary review and retrieval practice while remaining aligned with curricular objectives.

The study also offers implications for educational technology developers. Features such as timely feedback, progress monitoring, structured challenges, and meaningful opportunities for practice may contribute more directly to learning than reward systems or competition alone.

At the institutional level, universities may support instructors who integrate technology-enhanced language learning by providing appropriate infrastructure and professional development. Such support can help ensure that technology remains aligned with pedagogical objectives rather than becoming an additional activity disconnected from the curriculum.

Several directions for future research follow from these findings. First, replication studies with larger and more diverse samples across disciplines, educational levels, and institutions are needed to examine generalizability.

Second, future studies are encouraged to investigate the long-term effects of gamified learning environments by employing longer retention intervals, such as three- or six-month follow-up assessments. Such research would provide stronger evidence regarding the durability of learning gains beyond the short-term retention period examined in the present study. Third, comparative studies examining different gamification designs or specific game mechanics may help identify which instructional features contribute most effectively to vocabulary learning and retention. Because the present study evaluated a comprehensive gamified learning environment, the individual contribution of specific game elements remains unclear and warrants further investigation. Finally, future research may benefit from adopting mixed-methods designs that combine quantitative learning outcomes with qualitative evidence, such as learner interviews or classroom observations. Such approaches may provide a richer understanding of learners' experiences, perceptions, and patterns of engagement during gamified learning and help explain why particular instructional designs appear to be more effective than others.

References

Al-Khresheh, M. H. (2025). The cognitive and motivational benefits of gamification in English language learning: A systematic review. *The Open Psychology Journal*, *18*(1). https://doi.org/10.2174/0118743501359379250305083002

Ameen, F. (2026). Individual and group competitive digital gamification in ESL adult classrooms: A systematic review of effects on vocabulary retention, motivation, and engagement. *Smart Learning Environments*, *13*(1), 22. https://doi.org/10.1186/s40561-026-00447-z

Ayana, H., Mereba, T., & Alemu, A. (2024, August). Effect of vocabulary learning strategies on students' vocabulary knowledge achievement and motivation: the case of grade 11 high school students. In *Frontiers in education* (Vol. 9, p. 1399350). Frontiers Media SA. https://doi.org/10.3389/feduc.2024.1399350

Biseko, J. M. (2025). Vocabulary learning in EFL context: Do primary school English Subject textbooks provide structured support?. *Cogent Education*, *12*(1), 2455047. https://doi.org/10.1080/2331186X.2025.2455047

Capatina, A., Dinu, V., & Stoica, M. (2024). Gamification and knowledge retention in digital learning environments: A systematic review. *Education and Information Technologies*, *29*(2), 1453–1476.

Chan, S., & Lo, N. (2024, August). Enhancing EFL/ESL instruction through gamification: A comprehensive review of empirical evidence. In *Frontiers in education* (Vol. 9, p. 1395155). Frontiers Media SA. https://doi.org/10.3389/feduc.2024.1395155

Cheung, G., & Ng, K. (2021). Educational games and gamification in language learning: A critical review. *Computer Assisted Language Learning*, *34*(7), 1–24.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Dehghanzadeh, H., Fardanesh, H., Hatami, J., Talaee, E., & Noroozi, O. (2021). Using gamification to support learning English as a second language: a systematic review. *Computer Assisted Language Learning*, *34*(7), 934-957. https://doi.org/10.1080/09588221.2019.1648298

Fadzli, M. R., & Hashim, H. (2026). Gamification and vocabulary retention in EFL classrooms: A systematic review (2021–2025). *Language Teaching Research*, *30*(1), 88–110.

Fithriani, R. (2021). The utilization of mobile-assisted gamification for vocabulary learning: Its efficacy and perceived benefits. *Computer-Assisted Language Learning Electronic Journal*, *22*(3), 146-163.

Laksanasut, S. (2025). Gamification in ESL/EFL education: Transforming language learning and teaching through play. *Theory and Practice in Language Studies*, *15*(4), 1234–1248. https://doi.org/10.17507/tpls.1504.12

Laura-De La Cruz, K. M., Noa-Copaja, S. J., Turpo-Gebera, O., Montesinos-Valencia, C. C., Bazán-Velasquez, S. M., & Pérez-Postigo, G. S. (2023). Use of gamification in English learning in higher education: A systematic review. *Journal of Technology and Science Education*, *13*(2), 480-497. https://doi.org/10.3926/jotse.1740

Li, R. (2021). Does game-based vocabulary learning app influence Chinese EFL learners' vocabulary achievement, motivation, and self-confidence? *SAGE Open*, *11*(1), 21582440211003092. https://doi.org/10.1177/21582440211003092

Li, M., Ma, S., & Shi, Y. (2023). Examining the effectiveness of gamification as a tool promoting teaching and learning in educational settings: A meta-analysis. *Frontiers in Psychology*, *14*, 1253549.

Liu, S., Zhang, S., & Dai, Y. (2025). Do mobile games improve language learning? A meta-analysis. *Computer Assisted Language Learning*, 1-29. https://doi.org/10.1080/09588221.2025.2528786

Mleiki, A. K. (2026). The impact of game-based learning tools on vocabulary acquisition and motivation in undergraduate EFL classrooms. *Isagoge-Journal of Humanities and Social Sciences*, *6*(1), 44-60. https://doi.org/10.59079/isagoge.v6i1.286

Panmei, B., & Waluyo, B. (2023). Gamification in EFL learning: Motivation, engagement, and learning outcomes. *Education Sciences*, *13*(4), 412.

Pratiwi, D. I., Sari, R., & Hidayat, T. (2024). Flipped classroom integrated with gamification for vocabulary learning in EFL context. *TESOL Quarterly*, *58*(1), 221–240.

Shang, X. (2025). Mapping the integration of gamification into foreign language learning in E-learning settings: A bibliometric analysis. *Results in Engineering*, 100496. https://doi.org/10.1016/j.caeai.2025.100496

Xu, Q., & Richardson, J. C. (2024). The impact of social media and gamification of a mobile vocabulary learning app: Self-regulation and learning persistence. *Online Learning*, *28*(4), 1–25. https://doi.org/10.24059/olj.v28i4.4592

Yu, Z. (2023). Learning outcomes, motivation, and satisfaction in gamified English vocabulary learning. *SAGE Open*, *13*(2), 1–14. https://doi.org/10.1177/21582440231158332

Zeng, J., Sun, D., Looi, C. K., & Fan, A. C. W. (2024). Exploring the impact of gamification on students' academic performance: A comprehensive meta-analysis of studies from the year 2008 to 2023. *British Journal of Educational Technology*, *55*(6), 2478-2502. https://doi.org/10.1111/bjet.13471

Zhang, Q., & Yu, Z. (2022). Meta-analysis on investigating and comparing the effects on learning achievement and motivation for gamification and game-based learning. *Education Research International*, *2022*, 1519880. https://doi.org/10.1155/2022/1519880

Zhang, S., & Hasim, Z. (2023). Gamification in EFL/ESL instruction: A systematic review of empirical research. *Frontiers in Psychology, 13*, Article 1030790. https://doi.org/10.3389/fpsyg.2022.1030790