

AI Powered Gamified Tools and Vocabulary Retention Among Vietnamese Non-English Major Undergraduates: An Explanatory Sequential Mixed Methods Study at Nguyen Tat Thanh University

Authors Details

Do Duy Phuoc Thanh^{1*}

Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam.

Corresponding Author: *Do Duy Phuoc Thanh*

ABSTRACT: The current research aimed to examine the link between using the gamified vocabulary activities that utilized AI technology and vocabulary performance of undergraduate students from the Vietnamese university who did not pursue majors in English language studies. For the purposes of the current investigation, the researchers have applied an explanatory sequential mixed methods design. The quantitative data were collected from 60 students with B1 level of language competence via vocabulary pre-test, vocabulary post-test, and 47-item questionnaire. Next, the qualitative data were gathered through interviews of 10 students who were interviewed in order to reveal their experiences of participation in vocabulary activities. The intervention program included Quizlet, Bloocket, and Wayground activities coupled with repeated reviewing of vocabulary during seven weeks. It was found that the mean vocabulary score of participants rose from 5.06 (SD = 1.83) prior to the intervention to 7.79 (SD = 1.49) following the intervention. Paired-samples test revealed the significant result, $t(59) = 35.52$, $p < .001$, where the effect size was very large. Besides, students had positive perceptions regarding immediate recall, repeated review, and vocabulary use in speaking in class. Descriptive measures showed high values for ease of use, usefulness, relatedness, interest, and autonomy. The findings indicate that gamified digital tools can support vocabulary review when embedded in a

deliberate teaching sequence. Given the single-group design, however, the results should be interpreted as associations rather than causal effects.

Keywords: *AI-powered gamification, vocabulary retention, technology acceptance, self-determination theory, Vietnamese undergraduates.*

1. Introduction

Vocabulary knowledge is crucial for comprehension and use in second-language acquisition. It enables learners' engagement in classroom interactions, understanding of disciplinary texts, and communication in situations outside the language classroom (Nation, 2013). Vocabulary learning of non-English major Vietnamese undergraduates may be hampered by their limited exposure to the English language, preference for rote learning, and lack of opportunities for word retrieval and practice in context. Such limitations may be especially apparent in mandatory English classes that include students of various majors and various levels and different motivations regarding vocabulary learning.

There has been a widespread application of digital game-based learning to address those problems. Gamification is a process that involves the implementation of gaming design elements, including points, badges, leaderboards, challenges, and instant feedback in non-gaming contexts (Deterding et al., 2011). Game-based applications like Quizlet, Blooket, and Wayground enable teachers to transform vocabulary review into short gamified exercises in English-language classrooms. Reviews of gamification research nevertheless caution that outcomes vary according to design, learning context, user characteristics, and the relationship between game mechanics and learning goals (Hamari et al., 2014).

The increasing availability of AI features on educational platforms has allowed further development of options for content creation, adaptivity of practice, feedback, and personalized review. But the designation AI-powered must not be used as a certificate of educational efficiency. A user may appreciate the platform because it is accessible, entertaining, or social but the features responsible for vocabulary learning may still be repetition, feedback, and effective follow-up exercises. Research then has to analyze both learning achievements and students' reports on how they operate the tools.

There is research evidence in Vietnamese higher education that the attitudes of students towards the platforms and games for vocabulary learning are rather positive. However, there is a shortage of empirical studies of how non-English majors use these platforms and what is their effect on motivation and learning results of such students. The study attempts to fill this gap by exploring a seven-week package of activities with the help of Quizlet, Blooket, and Wayground at Nguyen Tat Thanh University.

2. Literature Review and Conceptual Framework

2.1 Gamification and vocabulary learning

Gamification does not mean that one uses game-based learning. It implies the inclusion of certain aspects of a game into a lesson, as opposed to using a whole educational game (Deterding et al., 2011; Kapp, 2012). The aspects to be used in vocabulary lessons include time-limited tests, progress bars, competition, rewards, and repetition. These aspects will be effective if the activity requires learners to pay attention to the form, meaning, and usage of words instead of just getting the right score.

For better vocabulary retention, it is important for learners to come across the target words repetitively and recall them after an interval, rather than just reread the list. Distributed learning works much better than massed learning for verbal memory; however, the optimal interval depends on the desired duration of memory (Cepeda et al., 2006). Flashcards and quick quizzes help learners to review the material repetitively more efficiently. The vocabulary test should take into account the differences between receptive knowledge and productive knowledge. Learners can recognize the right meaning of the word in a test but fail to recall the word in spontaneous speech (Nation, 2013; Webb, 2005).

It has been synthesized that practice using Quizlet and digital game-based methods in general are conducive to learning vocabularies and emotions when there is repeated exposure to the target vocabulary and immediate feedback is provided (Nguyen & Le, 2022; Özdemir & Seçkin, 2024). Nevertheless, learning through platforms is not always positive. The competitive aspect of the game may arouse learners' interests, yet the same can cause learners to experience more social comparison and lower satisfaction or intrinsic motivation if the rankings become more prominent (Hanus & Fox, 2015).

2.2 Technology acceptance and motivation

According to the Technology Acceptance Model, perceived usefulness and ease of use determine how learners accept information systems (Davis et al., 1989). In terms of learning vocabularies, perceived usefulness relates to how the learner perceives the usefulness of a platform in learning or reviewing words, while perceived ease of use refers to whether it is easy for them to use the platform without any difficulties. Behavioral intention to keep using the platform is often regarded as an outcome of these two factors, rather than as a separate learning process.

Motivation theory, which is called Self-Determination Theory, presents another viewpoint. It implies that people become engaged in learning if the conditions are favorable to autonomy, competence, and relatedness (Ryan & Deci, 2000). Gamified activities encourage learners' autonomy by providing them with choices of set or mode of the game, competence by means of feedback and progress tracking, and relatedness by allowing collaboration and social interaction. However, these activities will bring positive results only if they are implemented correctly.

2.3 Conceptual framework

The research viewed the activity package of seven weeks as an instructional condition which comprised computer-assisted review, game, retrieval practice and teaching assistance. Performance on the vocabulary test and perception of vocabulary learning were considered as outcome measures. Utility, ease of use, interest, autonomy, competence and relatedness were considered as dimensions that described students' experiences. Interviews with open-ended questions were conducted following the quantitative stage in order to understand the reasons for high scores and why recognition success was not transferred into productive usage of vocabulary.

3. Method

3.1 Research design and participants

An explanatory sequential mixed-methods approach was used in this study. First, quantitative data was collected and analyzed, then qualitative interviews were conducted in order to explain and expand on the quantitative findings (Creswell & Plano Clark, 2018). The quantitative part of the research had a one-group pre-test/post-test design. In other

words, the class went through the intervention and took a vocabulary test both before and one week after the intervention. No control group was used because of classroom limitations and, thus, no causality can be established.

60 students who were not majoring in English at Nguyen Tat Thanh University took part in the experiment. The participants were studying English at the approximate B1 level and were primarily taking English Course 3. The convenience sampling technique was used as the participants were taken from the available context. Then 10 students were selected for further interviews using maximum variation purposive sampling based on their post-test results – 5 high-scorers and 5 low-scorers.

Table 1. Participant profile (N = 60)

Characteristic	Category	n	%
Year of study	First year	19	31.7
	Second year	11	18.3
	Third year	18	30.0
	Fourth year	12	20.0
Experience using gamified tools	Less than 1 month	17	28.3
	1–3 months	14	23.3
	More than 3 months	29	48.3
Frequency of use	1–2 times per week	42	70.0
	3–4 times per week	14	23.3
	5 or more times per week	3	5.0

3.2 Intervention and instruments

Quizlet, Blooket, and Wayground were utilized for seven weeks in the intervention. About 135 words at the level of B1 in the English language general purpose were selected in such topic groups as personality, places, jobs, health, entertainment, socializing, and work. The weekly activities included flashcards revision, practice at one's own pace, quiz made by the teacher, games in a group, instant feedback, and speaking follow-up tasks. This research does not intend to differentiate the effectiveness of the platforms separately but rather treats them as a pedagogical bundle.

Vocabulary knowledge was tested using a pre- and post-test of 30 minutes each evaluated by the scale of 10 points. Each test included 35 questions about multiple-choice, matching, and fill-in-the-blanks items covering receptive and productive knowledge of vocabulary. Questionnaire included 47 items and was evaluated according to a scale of five points. It was used for measuring the interest, autonomy, competence, relatedness, usefulness, ease of use, short-term recall, short-delay retention, vocabulary use in speaking, and behavioral intention. The internal-consistency coefficients varied from .848 to .944 while the total questionnaire alpha was .982.

3.3 Data analysis

Descriptive statistics were applied for the description of participant characteristics, questionnaire data, and testing results. Paired samples t-test was run in order to examine the change before and after the test. Effect size d_z was calculated according to Cohen. Interviews were transcribed, and content analysis was performed according to the six-stage approach suggested by Braun & Clarke (2006). Coding included both deductive and inductive categories; deductive categories were based on the Theory of Technology Acceptance and Self-Determination Theory, while inductive categories were linked to technological problems, competition, repetition, and transfer from recognition into speaking.

4. Results

4.1 Vocabulary test performance

Table 2 presents descriptive statistics for the vocabulary tests. The mean post-test score was higher than the mean pre-test score, and score variability decreased after the intervention. The increase was observed across the score range, although the study does not establish that the intervention alone caused the change.

Table 2. Vocabulary test results

Test	N	M	SD	Minimum	Maximum
Pre-test	60	5.06	1.83	1.80	8.90
Post-test	60	7.79	1.49	3.80	10.00

The paired-samples analysis showed a mean difference of 2.74 points, $t(59) = 35.52$, $p < .001$, 95% CI [2.58, 2.89], $d_z = 4.58$. The effect estimate is very large and should be interpreted carefully in light of the one-group design, the relatively short delay before the post-test, and the need to establish equivalence between test forms. Nonetheless, the result is consistent with an improvement in vocabulary performance following the period of gamified review.

4.2 Students' perceptions of vocabulary learning

The students' perception about the vocabulary activities was positive. Flashcards and games were helpful in comprehending the meanings of words faster, remembering new words right away after practice, and remembering the vocabulary during classroom discussions. The highest mean score for an item was for the comprehension of the meaning of words faster ($M = 4.53$, $SD = 0.57$). The lowest mean score was for the second and fourth-week recall of the vocabulary ($M = 4.28$, $SD = 0.72$). However, due to the timing of the posttest, which was conducted one week later, the results cannot be considered objective measures of long-term retention.

Table 3. Selected perception items on vocabulary learning

Statement	M	SD
I understand the meanings of new words more quickly through flashcards and games.	4.53	0.57
Immediately after a lesson, I can remember most of the new vocabulary.	4.42	0.72
Repeated review prompted by the tools helps me retain vocabulary for longer.	4.40	0.67
I can use the words I have learned in classroom speaking activities or presentations.	4.38	0.74
After 2–4 weeks, I can still remember words learned through these tools.	4.28	0.72

4.3 Technology acceptance and motivational dimensions

Table 4 summarizes the construct-level results. All dimensions were rated positively. Perceived ease of use had the highest mean ($M = 4.49$), followed by behavioral intention and perceived usefulness (both $M = 4.47$). Relatedness, interest, autonomy, and perceived

competence were also rated highly. These findings describe students' positive experiences; they do not demonstrate that any one construct predicted vocabulary gain.

Table 4. Technology acceptance and motivational dimensions

Dimension	Items	M	SD
Perceived ease of use	4	4.49	0.72
Behavioral intention	4	4.47	0.64
Perceived usefulness	4	4.47	0.56
Relatedness	4	4.45	0.58
Interest	4	4.43	0.57
Autonomy	4	4.43	0.59
Perceived competence	4	4.30	0.62

4.4 Qualitative explanations and integration

The interview findings agreed with the quantitative results in a number of aspects. The students said that the applications were easy to use on mobile phones and easy to comprehend after minimal instruction. Flashcards, repetition, and immediate corrective feedback appeared to be associated with memorizing words and their meanings. Some students mentioned studying vocabulary independently, which helped explain the high levels of behavioral intentions. Group games, rankings, and game-like elements were mentioned as fun activities that facilitate communication with peers.

At the same time, the interview data provided some limitations of the positive quantitative results. Wi-Fi outages, frozen screens, difficulty with logging in, and variation in platform interfaces prevented some students from participating in the study. Though competition aroused excitement in most students, some participants experienced stress or were preoccupied with gaining points rather than studying vocabulary. The sense of autonomy was more evident during self-study than in teacher-led lessons. Finally, some students were able to recognize words in games but did not dare to pronounce or use them freely in oral communication.

Table 5. Integration of quantitative and qualitative findings

Dimension	Quantitative pattern	Qualitative explanation
Ease of use	Highest mean (M = 4.49)	Mobile access and simple interfaces were valued; Wi-Fi and login problems remained.
Usefulness	High mean (M = 4.47)	Flashcards, feedback, and repeated exposure supported recall of word forms and meanings.
Interest and relatedness	High means (M = 4.43–4.45)	Competition and group activities increased enjoyment and peer interaction, but could create pressure.
Autonomy and competence	Positive means (M = 4.30–4.43)	Self-paced review and feedback supported confidence; teacher-controlled activities reduced choice.
Vocabulary use	Positive self-reports	Recognition did not always transfer to spontaneous pronunciation or spoken use.

5. Discussion

The analysis reveals a significant improvement in vocabulary test results after using a set of gamified review exercises for seven weeks. Such a finding supports the idea that practice and retrieval opportunities could lead to better recall when these processes are spaced rather than crammed in one session (Cepeda et al., 2006). Gamified flashcard and quiz activities using Quizlet may have promoted practice and repetition of the target material, while Blooket and Wayground might have introduced feedback and motivation elements. In addition, the finding supports previous research showing how computer-assisted lexical instruction works if the target words are actively involved in the process (Nguyen & Le, 2022; Özdemir & Seçkin, 2024).

It is, however, essential to note that the finding cannot prove the cause-and-effect relationship between gamified tools and improved test results. Without an alternative condition to compare the effect of gamification with, it is not possible to distinguish between various factors, including regular instruction, instructor attention, testing itself, or

outside-of-class preparation on the part of the students. The effect size is enormous, which is why replication of the study is especially important with the delayed assessment and control conditions (Shadish et al., 2002).

Both constructs of the Technology Acceptance Model and Self-Determination Theory had high scores according to the students' responses, which confirmed their relevance. The perceived usefulness and ease of use mean that students viewed the platforms as something functional and useful rather than difficult to use, which is an essential prerequisite for sustained use (Davis et al., 1989). Motivation was well covered by the interest, relatedness, autonomy, and perception of competence as they were rated high, but interview data indicates that these characteristics were not equally experienced. Competition might have prompted people's engagement but at the same time, the focus on ranking could bring some pressure. Thus, this limited finding confirms earlier warnings that gamification might lead to increased social comparison and demotivation due to an unbalanced approach (Hanus & Fox, 2015).

The major implication for teaching relates to the differentiation between receptive and productive knowledge of vocabulary. Although students appreciated the opportunities offered by the game for memorizing words, some still faced problems with their pronunciation and spontaneous use. Hence, games may serve as a tool for introduction and review of new words through digital platforms, and further application of vocabulary in short dialogues, descriptions of pictures, role plays (Nation, 2013; Webb, 2005).

6. Conclusion

An explanatory sequential mixed methods study was conducted to investigate the effect of a seven-week intervention consisting of AI-driven gamified activities on the vocabulary acquisition of 60 undergraduates at Nguyen Tat Thanh University who are not majoring in English language. Participants performed better in the vocabulary test after the intervention and showed a positive attitude towards reviewing vocabulary through such activities, their usability, usefulness, motivation, and continued use intention. According to interviews, repetition, immediate feedback, learner choice, and peer interaction appeared as facilitating conditions.

The current study has revealed several constraints for gamified activities. They are technical issues, high competitiveness, teacher-paced activities, and difficulties in shifting from recognition to production of vocabulary. Thus, this research suggests balancing the integration of the gamification tools and a systematic sequence of vocabulary teaching and learning activities. However, it is important to mention that this study involved one experimental group with no control group and short-term testing. So, the findings cannot be generalized and viewed as causations. Further studies need to compare experimental and control groups, conduct a delayed testing, measure speaking proficiency directly, and investigate individual platforms separately.

References

1. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
2. Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354–380. <https://doi.org/10.1037/0033-2909.132.3.354>
3. Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.
4. Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982–1003. <https://doi.org/10.1287/mnsc.35.8.982>
5. Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification: Using game-design elements in non-gaming contexts. In *Proceedings of the 2011 annual conference extended abstracts on human factors in computing systems* (pp. 2425–2428). Association for Computing Machinery. <https://doi.org/10.1145/1979742.1979575>
6. Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In *2014 47th Hawaii international conference on system sciences* (pp. 3025–3034). IEEE. <https://doi.org/10.1109/HICSS.2014.377>

7. Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, 80, 152–161. <https://doi.org/10.1016/j.compedu.2014.08.019>
8. Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education*. Pfeiffer.
9. Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
10. Nguyen, L. Q., & Le, H. V. (2022). Quizlet as a learning tool for enhancing L2 learners' lexical retention: Should it be used in class or at home? *Human Behavior and Emerging Technologies*, 2022, Article 8683671. <https://doi.org/10.1155/2022/8683671>
11. Özdemir, O., & Seçkin, H. (2024). Quantifying cognitive and affective impacts of Quizlet on learning outcomes: A systematic review and comprehensive meta-analysis. *Frontiers in Psychology*, 15, Article 1349835. <https://doi.org/10.3389/fpsyg.2024.1349835>
12. Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
13. Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
14. Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, 27(1), 33–52. <https://doi.org/10.1017/S0272263105050023>