IOSR Journal of Research & Method in Education (IOSR-JRME)
e-ISSN: 2320–7388, p- ISSN: 2320-737x Volume 16, Issue 3 Ser. 2 (May. – June. 2026) 35-40
www.iosrjournals.org

The Effect Of Liveworksheets On English Vocabulary Retention Among Eighth-Grade Students

Vu Dinh Bac, Mai Van Can
(Faculty Of Foreign Languagues, Thai Nguyen University Of Education, Vietnam)

Abstract:
This study examined the longitudinal development of English vocabulary retention among 44 eighth-grade EFL learners at a lower secondary school in Vietnam over a 16-week instructional intervention. Grounded in the action research framework of Kemmis and McTaggart (1988), the study integrated Liveworksheets as a supplementary retrieval-based practice tool within the Practice stage of the Presentation-Practice-Production (PPP) pedagogical framework. Vocabulary knowledge was measured at five time points using a modified Vocabulary Knowledge Scale (VKS): progress tests at Weeks 4, 8, and 12, a post-test at Week 14, and a delayed post-test at Week 16. The results revealed a steady upward trajectory in progress test scores (Week 4: M = 29.7, SD = 4.2; Week 12: M = 43.6, SD = 2.4), followed by strong post-test performance (83.0% of students scoring 80 or above) and relative score stability in the delayed post-test (p = .118, Cohen's d = 0.241). These findings suggest that structured digital retrieval practice through Liveworksheets can support incremental vocabulary development and contribute to short-term retention stability in EFL lower secondary classrooms.

Key Word: *vocabulary retention, Liveworksheets, retrieval practice, EFL, action research, lower secondary*

Date of Submission: 12-05-2026 Date of Acceptance: 22-05-2026

I. Introduction

Vocabulary knowledge occupies a central position in language acquisition, serving as the foundation for communicative competence in all four skills (Nation, 2001). Despite years of English instruction, a recurring challenge among lower secondary learners in Vietnam is difficulty with long-term vocabulary retention, the ability to recall and use lexical items accurately beyond the immediate instructional period. This challenge is particularly pronounced in EFL contexts characterized by limited exposure outside the classroom, where words learned through rote memorization or passive reading are easily forgotten within days of initial presentation.

Research in cognitive psychology and applied linguistics has consistently highlighted the role of retrieval practice, spaced repetition, and immediate corrective feedback in deepening memory encoding and strengthening lexical retention (Nation, 2001; Webb, 2007). These mechanisms, however, are difficult to operationalize within the constraints of traditional paper-based instruction, where opportunities for structured repetition are limited and feedback is often delayed. Technology-enhanced learning environments offer promising alternatives. Among available digital platforms, Liveworksheets has emerged as a low-barrier, interactive tool that enables teachers to design vocabulary exercises with automatic correction and repeated practice across multiple sessions.

While existing studies have explored the general effectiveness of Liveworksheets in EFL contexts, most have relied on short-term, cross-sectional designs without delayed post-tests, limiting their capacity to address long-term retention. Furthermore, few have been conducted in the specific context of lower secondary education in Vietnam, where vocabulary demands increase substantially as learners transition from foundational to intermediate content. There is therefore a need for longitudinal, classroom-based evidence that tracks vocabulary development and retention across a full instructional semester.

This study addresses these gaps by reporting the findings of a 16-week action research intervention conducted with 44 eighth-grade students at a secondary school in Northen Vietnam, where Liveworksheets was integrated as a twice-weekly out-of-class vocabulary practice tool. The study focuses specifically on Research Question: How does students' vocabulary retention change over time when Liveworksheets is integrated into classroom instruction? Vocabulary knowledge was tracked across five measurement points, enabling a longitudinal view of development (progress tests) and retention (post-test and delayed post-test).

II. Literature Review

Vocabulary Retention: Short-term and Long-term Dimensions

Vocabulary retention refers to the capacity to store and subsequently retrieve lexical knowledge after a period of time has elapsed since initial learning. Nation (2001) distinguishes between the breadth of vocabulary knowledge (how many words a learner knows) and its depth (how well individual words are known across

dimensions of form, meaning, and use). Retention, in this framework, is not simply the passive preservation of information but an active process of consolidation that depends on the quality and frequency of engagement with target items.

Short-term retention, typically assessed within hours or days of instruction, is influenced by factors such as processing depth, encoding richness, and the degree of active engagement during practice. Long-term retention, measured weeks or months after instruction, requires that initial memory traces be sufficiently strengthened through repetition and retrieval to resist decay (Schmitt, 2000). Critically, the conditions that best support long-term retention are not always identical to those that produce the highest immediate scores. Pair-mediated tasks, for instance, may show lower immediate gains than individual work but yield superior retention over time due to deeper collaborative processing. Spaced practice has similarly been found to promote long-term over short-term retention compared to massed exposure (Schuetze, 2015).

Retrieval practice, the act of actively recalling information from memory rather than passively reviewing it, is among the most consistently supported mechanisms for long-term retention in both cognitive psychology and applied linguistics (Karpicke & Roediger, 2008). When learners retrieve a word under conditions of appropriate challenge, neural pathways associated with that item are strengthened, making future retrieval more efficient. This principle underpins the pedagogical value of exercises such as gap-fill tasks, matching activities, and sentence production within digital practice platforms.

Liveworksheets as a Digital Vocabulary Practice Platform

Liveworksheets is a web-based educational platform that enables educators to convert static printable worksheets into interactive, self-correcting digital exercises. Its functionality includes multiple-choice items, fill-in-the-blank tasks, drag-and-drop matching, listening-based activities, and multimedia embedding (Liveworksheets, n.d.; Prabjandee, 2023). From a pedagogical standpoint, several features of the platform align with established mechanisms of vocabulary retention.

First, the platform's interactive task formats, particularly gap-fill and matching exercises, require active retrieval from memory, operationalizing the retrieval practice effect. Second, immediate automated feedback following each response creates a corrective loop that enables learners to identify and revise inaccurate form-meaning associations in real time, reducing the reinforcement of incorrect knowledge. Third, repeated exposure to the same target vocabulary across multiple sessions over successive weeks replicates the conditions of spaced retrieval, supporting the gradual consolidation of lexical memory traces. These mechanisms are consistent with Nation's (2001) framework of vocabulary learning conditions, which emphasizes the importance of noticing, retrieval, and generative use.

Beyond Liveworksheets specifically, related research on computer-based flashcard platforms has produced convergent evidence: Phuong et al. (2026), in a quasi-experimental study with 84 university students in northern Vietnam, found that Flippity Flashcards significantly improved ESP multimedia vocabulary knowledge across receptive and productive dimensions, with a large effect size, a finding that broadens the empirical base for technology-mediated vocabulary practice in Vietnamese EFL contexts. Nevertheless, methodological limitations across this body of work, including small samples, short intervention durations, and absence of delayed post-tests, have prevented strong conclusions about long-term retention effects.

The PPP Framework and Liveworksheets in the Practice Stage

The Presentation-Practice-Production (PPP) framework provides a widely adopted model for sequencing vocabulary instruction. In the Presentation stage, new lexical items are introduced with contextual support, facilitating initial encoding and noticing. The Practice stage involves controlled, form-focused activities designed to reinforce form-meaning associations through repeated exposure and retrieval. The Production stage invites learners to deploy vocabulary in more open-ended communicative tasks, supporting deeper processing and contextualization (Harmer, 2007; Scrivener, 2011).

In the present study, Liveworksheets was positioned primarily within the Practice stage. Following classroom-based vocabulary presentation, students completed Liveworksheets activities twice weekly as out-of-class assignments, providing structured retrieval-based practice between successive lessons. This design aligns with the theoretical predictions of the PPP model: controlled and repeated Practice-stage engagement is expected to strengthen memory traces and reduce forgetting prior to summative assessment. While the Production stage was not systematically formalized, limited follow-up classroom tasks occasionally provided opportunities for contextualized vocabulary use.

III. Methodology

This study employed a single-cycle action research design based on the spiral model of Kemmis and McTaggart (1988), comprising planning, acting, observing, and reflecting phases. Action research was selected because of its suitability for examining and improving pedagogical practice within a specific classroom context,

and because the researcher served concurrently as the class teacher. The iterative design enabled continuous monitoring of student engagement and vocabulary development throughout the intervention, with reflective adjustments made to task delivery where necessary.

The study was conducted at a lower secondary school, during the first semester of the 2025-2026 academic year. Participants comprised 44 eighth-grade students (23 male, 21 female; mean age = 13.2 years) from a single intact class, selected through convenience sampling. Most students demonstrated lower-intermediate English proficiency based on previous semester grades and classroom observation. Twenty-seven students had personal digital devices (primarily smartphones), while the remaining 17 used family-shared devices. All Liveworksheets assignments were completed at home, outside scheduled class time, to ensure accessibility.

The school used Global Success 8 as the official English textbook, with three 45-minute English lessons scheduled weekly. Ethical clearance was obtained from the school administration, and informed written consent was secured from both students and parents prior to the study.

The intervention spanned 16 weeks. During Weeks 1 through 12, new vocabulary was introduced during regular classroom instruction (Presentation stage) and subsequently reinforced through two weekly Liveworksheets sessions assigned as homework (Practice stage). Each assignment targeted 10 vocabulary items drawn from the corresponding curricular unit. Activities included matching, fill-in-the-blank, multiple-choice, and drag-and-drop tasks. Students submitted completed assignments digitally via the platform's built-in submission function; the teacher monitored completion through the online gradebook. Weeks 13 and 14 served as a consolidation and post-test period, followed by the delayed measurement at Week 16.

Vocabulary knowledge was assessed using a modified Vocabulary Knowledge Scale (VKS) (Wesche & Paribakht, 1996), a five-level self-report instrument in which learners indicate their knowledge of each target word from Level I (no recognition) to Level V (accurate productive use). To enhance scoring objectivity, responses at Levels III, IV, and V required written evidence (a synonym, translation, definition, or original sentence), which was evaluated against predefined criteria. Responses that did not meet the criteria were downgraded. Inter-rater verification was conducted with a second English teacher from the same school, with consensus reached on all discrepancies.

Progress tests of 10 VKS items each were administered in-class at Weeks 4, 8, and 12 (maximum score per test: 50). The post-test and delayed post-test each comprised 25 target vocabulary items selected for coverage across instructional units and range of difficulty (maximum score: 125). The delayed post-test used the same items as the post-test, with question order randomized to reduce test-retest effects. The two-week interval between post-test and delayed post-test was chosen as a practical balance between ecological validity and participant retention.

Descriptive statistics (means and standard deviations) were computed for all test scores using SPSS (version 27). A paired-samples t-test was conducted to evaluate the statistical significance of the difference between post-test and delayed post-test scores. Effect size was reported using Cohen's d and Hedges' g. Score distributions for post-test and delayed post-test results were organized into frequency bands to illustrate performance profiles. The significance threshold was set at $p < .05$.

IV. Results

Progress Test Results

Table 1 presents descriptive statistics for the three progress tests administered at Weeks 4, 8, and 12 of the intervention. Each test assessed 10 target vocabulary items using the VKS format (maximum score: 50).

Table 1. Descriptive Statistics for Progress Tests (Weeks 4, 8, and 12)

Test Point	Week	N	Mean (max 50)	SD	Min	Max
Progress Test 1	Week 4	44	29.7	4.2	20	40
Progress Test 2	Week 8	44	36.4	3.6	26	46
Progress Test 3	Week 12	44	43.6	2.4	36	50

The progress test data reveal a consistent upward trend in students' vocabulary knowledge across the three measurement points. Mean scores increased from 29.7 (SD = 4.2) at Week 4 to 36.4 (SD = 3.6) at Week 8, and further to 43.6 (SD = 2.4) at Week 12. This represents a cumulative gain of 13.9 points (47% relative improvement) over the 12-week practice period. Notably, the standard deviation decreased progressively across the three tests, suggesting that lower-performing students made proportionately greater gains over time, narrowing the gap with higher-performing peers. These patterns are consistent with the expectations of the PPP framework, wherein regular Practice-stage engagement supports incremental lexical consolidation.

Post-test Results

The post-test, administered at Week 14, assessed 25 target vocabulary items (maximum score: 125) from across the intervention period. Table 2 presents the score distribution.

Table 2. Post-Test Score Distribution (Week 14, N = 44)

Score Range	Frequency	Percent	Cumulative Percent
50–74	3	6.8	6.8
75–99	19	43.2	50.0
100–124	19	43.2	93.2
125	3	6.8	100.0
Total	44	100.0	—

Post-test results indicate generally strong vocabulary performance at the end of the instructional period. Combining the two highest score bands, 83.0% of students achieved scores of 100 or above (out of 125), with 6.8% achieving the maximum. Only 6.8% of students fell below the 75-point threshold. These results suggest that, on average, students had developed sufficient command of the target vocabulary to score at VKS Levels IV or V on the majority of items, indicating accurate form-meaning knowledge and emerging productive competence. This profile is consistent with the cumulative gains observed in the progress tests and reflects the sustained retrieval-based practice across the 12-week intervention.

Delayed Post-test Results and Retention Analysis

The delayed post-test, using the same 25-item VKS instrument with randomized item order, was administered at Week 16, two weeks after the post-test. Table 3 presents the delayed post-test score distribution, and Table 4 summarizes the paired-samples t-test comparing the two occasions.

Table 3. Delayed Post-Test Score Distribution (Week 16, N = 44)

Score Range	Frequency	Percent	Cumulative Percent
50–74	5	11.4	11.4
75–99	20	45.5	56.9
100–124	17	38.6	95.5
125	2	4.5	100.0
Total	44	100.0	—

Comparing Tables 2 and 3, the score distribution shifted modestly downward: the proportion of students scoring 100 or above declined from 50.0% to 43.1%, and those in the 50–74 range increased from 6.8% to 11.4%. However, the majority of students (88.6%) continued to score at or above 75 points at the delayed measurement point.

Table 4. Paired-Samples T-Test: Post-Test vs. Delayed Post-Test

Comparison	Mean Diff.	SD	SE Mean	95% CI Lower	95% CI Upper	t	df	p
Post – Delayed	2.86	9.14	1.38	−0.12	5.84	2.07	43	.118

The paired-samples t-test revealed a mean score decline of 2.86 points from post-test to delayed post-test (SD = 9.14), with a 95% confidence interval spanning from −0.12 to 5.84. This difference was not statistically significant, $t(43) = 2.07$, $p = .118$, and the confidence interval crosses zero, indicating that the observed decline falls within the range of sampling variability. Table 5 presents the corresponding effect size estimates.

Table 5. Effect Size: Post-Test vs. Delayed Post-Test

Comparison	Standardizer	Cohen's d	95% CI Lower	95% CI Upper
Post – Delayed	9.14	0.241	−0.060	0.538
Post – Delayed (Hedges' g)	9.22	0.239	−0.059	0.534

Cohen's d of 0.241 and Hedges' g of 0.239 both indicate a small effect size according to conventional benchmarks (Cohen, 1988). The 95% confidence intervals for both estimates include zero, providing further evidence that the true effect may be negligible. Taken together, the statistical and practical significance analyses consistently indicate that the vocabulary knowledge demonstrated at the post-test was largely preserved at the delayed measurement two weeks later.

V. Discussion

The findings of this study address Research Questionby providing longitudinal evidence of vocabulary development and short-term retention stability in a classroom where Liveworksheets was integrated as a structured Practice-stage tool. Three main interpretive threads emerge.

First, the steady upward trajectory in progress test scores, from a mean of 29.7 at Week 4 to 43.6 at Week 12, suggests that repeated retrieval-based practice through Liveworksheets was associated with cumulative vocabulary learning during the intervention. The simultaneous decline in score variance indicates that improvement was not limited to higher-performing students but was distributed across the class. This pattern aligns with theoretical accounts of retrieval-strengthened memory consolidation (Karpicke & Roediger, 2008) and is consistent with prior studies reporting incremental vocabulary gains associated with digital practice tools. From the perspective of the PPP framework, these gains are interpretable as the product of sustained Practice-stage engagement, where repeated controlled retrieval progressively strengthens form-meaning associations initiated during classroom Presentation.

Second, the strong post-test performance, with 83.0% of students scoring at or above 100 of 125 possible points, suggests that most learners had developed functional command of the target vocabulary by the end of the instructional period. VKS scores at Levels IV and V indicate not merely recognition but the capacity to provide accurate meanings and produce the words in context, reflecting qualitative depth of lexical knowledge consistent with Nation's (2001) multidimensional framework. These findings are concordant with previous studies highlighting the positive effects of retrieval-based digital practice on summative vocabulary outcomes (Pradnyadari et al., 2025; Phuong et al., 2026). Notably, Phuong et al. (2026) reported similarly strong gains using Flippity Flashcards in a Vietnamese university EFL context, suggesting that technology-mediated retrieval practice can produce substantial vocabulary gains across educational levels and platform types.

Third, and most critically with respect to the retention dimension of Research Question, the non-significant difference between post-test and delayed post-test scores ($p = .118$, Cohen's $d = 0.241$) indicates that the vocabulary knowledge demonstrated immediately after instruction was largely maintained two weeks later, without additional intervention. The small effect size, with confidence intervals that include zero, provides quantitative evidence that the observed decline was not meaningfully different from chance variation. This finding extends prior cross-sectional evidence by demonstrating that retention stability can be observed within a real classroom context at the lower secondary level, and is consistent with studies of technology-enhanced vocabulary learning that have reported stable short-term retention following digital retrieval-based practice.

Several caveats warrant attention. The absence of a pre-test and control group limits causal attribution; observed improvements may reflect factors beyond the Liveworksheets intervention, including general maturation, increased familiarity with the VKS test format, or regular in-class vocabulary instruction. The relatively short two-week delay between post-test and delayed post-test cannot address questions about retention over longer intervals. Additionally, because Liveworksheets assignments were completed at home, the researcher could not fully control for external assistance or ensure equivalent practice conditions across all students. These limitations underscore the exploratory nature of the findings and the need for replication with more controlled designs.

VI. Conclusion

This study investigated how students' vocabulary retention changed over time when Liveworksheets was integrated into classroom instruction as a Practice-stage tool within a 16-week action research intervention. The results documented a consistent pattern: progress test scores improved steadily across the intervention (from $M = 29.7$ to $M = 43.6$), post-test performance indicated strong acquisition of the target vocabulary (83.0% scoring ≥ 100 out of 125), and delayed post-test results confirmed relative retention stability, with no statistically significant decline over a two-week interval ($p = .118$, $d = 0.241$).

These findings carry several pedagogical implications. For English teachers at the lower secondary level, the study supports the integration of structured digital retrieval practice, particularly through platforms such as Liveworksheets, as a complement to classroom vocabulary instruction. The twice-weekly out-of-class assignment format used in this study provides a replicable model that does not require additional class time and is compatible with existing PPP-based lesson structures. For curriculum designers, the results suggest that vocabulary review cycles incorporating spaced retrieval can be embedded into semester-long instructional sequences to support retention.

Future research should address the limitations of the current study by incorporating pre-tests, control groups, larger and more diverse samples, and extended delayed measurement intervals. Comparative studies examining differential effects across learner proficiency levels, and investigating which specific Liveworksheets task types contribute most to retention, would also enrich the evidence base for technology-enhanced vocabulary pedagogy in Vietnamese EFL contexts.

References

[1]. Cohen, J. (1988). Statistical Power Analysis For The Behavioral Sciences (2nd Ed.). Lawrence Erlbaum.
[2]. Harmer, J. (2007). The Practice Of English Language Teaching (4th Ed.). Pearson Longman.
[3]. Karpicke, J. D., & Roediger, H. L. (2008). The Critical Importance Of Retrieval For Learning. Science, 319(5865), 966–968.
[4]. Kemmis, S., & Mctaggart, R. (1988). The Action Research Planner (3rd Ed.). Deakin University Press.
[5]. Liveworksheets (N.D.). About Liveworksheets. Retrieved From Https://Www.Liveworksheets.Com
[6]. Nation, I. S. P. (2001). Learning Vocabulary In Another Language. Cambridge University Press.
[7]. Phuong, T. T. N., Ngoc, N. T. B., & Can, M. V. (2026). Enhancing Multimedia ESP Vocabulary Acquisition Through Flippity Flashcards. Journal Of Educational And Social Research, 16(3), 640. Https://Doi.Org/10.36941/Jesr-2026-0352
[8]. Schmitt, N. (2000). Vocabulary In Language Teaching. Cambridge University Press.
[9]. Schuetze, U. (2015). Spacing Techniques In Second Language Vocabulary Acquisition: Short-Term Gains Vs. Long-Term Retention. Language Teaching Research, 19(1), 28–42.
[10]. Scrivener, J. (2011). The Essential Guide To English Language Teaching (3rd Ed.). Macmillan.
[11]. Webb, S. (2007). The Effects Of Repetition On Vocabulary Knowledge. Applied Linguistics, 28(1), 46–65.
[12]. Wesche, M., & Paribakht, T. S. (1996). Assessing Second Language Vocabulary Knowledge: Depth Versus Breadth. Canadian Modern Language Review, 53(1), 13–40.