INTERACTION: Jurnal Pendidikan Bahasa Vol. 13, No. 3; September 2026
P-ISSN: 2406-9558; E-ISSN: 2406-9566

# The Effectiveness of ZEP QUIZ-Based Assessment on Students' Synonym Vocabulary Mastery

**Difa Amini Harliana**
difaamini024@gmail.com

**Sarlita D. Matra**
starlighta.unique11@gmail.com

**Universitas Pekalongan**

## ABSTRACT

This study investigated the effectiveness of ZEP QUIZ-based assessment in improving synonym vocabulary mastery among eleventh-grade EFL students at MAS Simbang Kulon, Indonesia. A quasi-experimental design with a non-equivalent control group was employed involving 62 students. Data were collected through pre-tests and post-tests and analyzed using the non-parametric Mann-Whitney U test on the students' normalized gain (N-gain) scores. The findings revealed a significant difference in vocabulary improvement rates between the experimental and control groups (Asymp. Sig. < .001), with the experimental group achieving significantly higher mean N-gain score (85.37%, categorized as Effective) than the control group (28.71% categorized as Ineffective). These results indicate that ZEP QUIZ-based assessment is more effective than conventional assessment in enhancing students' synonym vocabulary mastery. The study highlights the potential of gamified assessment as an engaging and effective tool for vocabulary learning in EFL classrooms.

**Keywords:** ZEP QUIZ; Gamified Assessment; Synonym Vocabulary; Vocabulary Mastery

## INTRODUCTION

Having a strong vocabulary is essential for learning English as it helps learners enhance the four language abilities: listening, speaking, reading, and writing. If students lack enough words, they might struggle to comprehend written materials, share their thoughts, or communicate clearly (Sun et al., 2023). One important aspect of vocabulary mastery is synonym vocabulary, which helps students vary word choices, avoid repetition, and better comprehend different meanings in texts (Boontam, 2022). Therefore, it is crucial for students to learn synonym vocabulary in order to improve their overall skill in English.

Nonetheless, evidence obtained from the researchers' firsthand classroom observations and an evaluation of students' daily task portfolios at MAS Simbang Kulon (Buaran Subdistrict, Pekalongan Regency) revealed that eleventh-grade students encounter significant difficulties in mastering vocabulary related to synonyms. This issue was concretely reflected in the classroom observation

records, which indicated that over 70% of the students consistently fell short of the expected target score during regular vocabulary exercises. Additional classroom observation notes indicated during their daily assignments, students often struggled to recognize near-synonyms, frequently confusing words that possess slight semantic distinctions depending on the context. Moreover, they faced challenges in selecting suitable synonyms in particular textual situations and consistently dependend on a limited, repetitive range of words while constructing their sentences. These findings were derived from the researchers' interactions while teaching the eleventh-grade Social Studies classes (P8 and P9) at MAS Simbang Kulon. This predicament may arise from traditional paper-based testing and teaching strategies that are passive and do not effectively engage students in active vocabulary learning (Muhtarom and Maghfiroh, 2022). As a result, students frequently feel unmotivated during vocabulary lessons, underscoring the pressing need for a more engaging and effective formative assessment method.

To address these challenges, recent educational studies have indicated that the implementation of game-based assessment (GBA) instruments like ZEP QUIZ has yieleded encouraging outcomes. Research conducted by Ratnawati and Wahyuni (2026) illustrated that ZEP QUIZ serves as a robust assessment mechanism that notably enhances students' comprehension abilities through interactive functionalities. Their results indicated that the competitive aspect of ZEP QUIZ fosters a more captivating assessment environment, which consequently leads to improved academic success compared to conventional paper-based evaluations. Additionally, the platform provides immediate feedback, which allows students to reflect on their errors in real time. This technological advancement demonstrates superior effectiveness in assessing students' genuine language skills while helping them maintain concentration throughout examinations.

To grasp the reasons behind the effectiveness of this platform, it is vital to outline its fundamental features. ZEP QUIZ is an online learning solution intentionally crafted to enhance the creation and execution of assessments through an engaging, game-creation and execution of assessments through an engaging, game-like approach (Adipat et al., 2021). This platform was developed by the South Korean entity ZEP, which is a partnership between NAVER Z and Supercat (Lee et al., 2024). In the context of local pedagogy, its structed framework has been officially influencing the Indonesian educational sector since its broader adaptation in early synchronous learning setups (Ratnawati & Wahyuni, 2026). Essentially, ZEP QUIZ turns the classroom into a live "game show" atmosphere that can be accessed using Information and Communication Technology (ICT) devices such as laptops and smartphones. Unlike traditional educational tools, ZEP QUIZ provides multi-player experiences centered around adventure themes, allowing learners to use customizable avatars while encountering real-time challenges paired with engaging audio components (Korkmaz, 2021; Lee et al., 2024). By incorporating these gamified elements, teachers can significantly boost student involvement and shift their attention toward constructive academic activities, which in turn promotes

notable advancements in vocabulary proficiency (Setyaningrum et al., 2025; Udeozor et al., 2023).

Moreover, the use of ZEP QUIZ addresses a prevalent issue within education, where numerous students regard assessments as a daunting experience. As noted by Ratnawati and Wahyuni (2026), the gamified features of this platform successfully support student engagement by turning a typically stressful assessment into an enjoyable experience. This supportive atmosphere enables students to express their knowledge more effortlessly and offers a more accurate representation of their understanding of synonyms. This is in agreement with Muryani and Yunus (2024), who argue that digital media contributes to a more adaptable and effective learning setting for contemporary students.

However, a distinct research gap persists in the current literature. Although earlier investigations concerning digital assessment and ZEP QUZ have thoroughly examined general student engagement, basic reading comprehension, and overall technical functionalities (e.g., Adipat et al., 2021; Korkmaz, 2021; Zhu et al., 2023) they have failed to consider how this gamified platform specifically influences the evaluation of nuanced lexical distinction, such as synonym vocabulary assessment. Most previous studies focused on general vocabulary size rather than the targeted measurement of semantic variations under real-time game constraints. To bridge this pedagogical gap, this study investigates the effectiveness of ZEP QUIZ-based assessment in measuring synonym vocabulary mastery among eleventh-grade Social Studies students. Consequently, this research aims to address the following specific research questions. The study addresses two research questions: 1) How is the ZEP QUIZ-based assessment implemented in evaluating synonym vocabulay mastery among the eleventh-grade Social Studies students at MAS Simbang Kulon?; and 2) Is there any significant difference in synonym vocabulary mastery between students assessed using the ZEP QUIZ-based platform and those assessed through conventional paper-based methods?

## LITERATURE REVIEW

### The Dynamics of Synonym Vocabulary Mastery

Vocabulary serves as the essential foundation for acquiring the English language; as noted by Wilkins (as references in Obata, 2022), limited grammatical understanding may slightly restrict communication, whereas a lack of vocabulary completely hinders the ability to convey meaning. Conversely, research by (Siregar & Henni, 2023) indicates that the vocabulary knowledge of students in Indonesia frequently shows inconsistency and often falls below the expeced level. The importance of vocabulary is further emphasized by (Sari et al., 2022), who argue that it is vital for students' ability to engage with and understand intricate materials. Particularly, expertise in synonyms allows students to prevent langugae redundancy and articulate their thoughts with enhanced accuracy (Stasya Prima Sari & Ria Pandiangan, 2021). In the context of language assessment, synonym vocabulary

requires high cognitive flexibility because students must quickly distinguish subtle semantic nuances between near-synonoyms based on specific contexts. This unique linguistic characteristic makes synonym mastery highly suitable for game-based assessments. Unlike general vocabulary size, which relies on rote memorization, evaluating synonyms under real-time game constraints tests students' fluid lexical retrieval and rapid problem-solving skills, making the assessment process more dynamic and authentic.

**Digital Game-Based Assessment (DGBA)**

In the current digital environment, Digital Game-Based Assessment (DGBA) has emerged as a new method to reduce test anxiety while accurately reflecting the cognitive processes of students (Zhu et al., 2023). The DGBA framework consists of combined components such as key goals, learning results, gameplay tasks, evaluation standards, and marking systems (Udeozor et al., 2023). From an educational viewpoint, the use of such digital tools aligns with the modern trend towards digital technology and AI-based teaching resources, requiring a significant change in how educational materials and assessments are presented (Syahrizal et al., 2024). This study utilizes the DGBA framework to transition from passive, conventional paper methods to an organized, software-based testing system. By transforming evaluative tasks into interactive gameplay, DGBA provides an operational mechanism where formative evaluation functions not as a daunting threat, but as a continuous, low-stakes feedback loop that sustains high cognitive engagement.

**ZEP QUIZ in Language Education**

ZEP QUIZ, an innovative creation from the South Korean company ZEP (a collaboration between NAVER Z and Supercat), transforms traditional assessments into an interactive gaming experience (Adipat et al., 2021; Ratnawati & Wahyuni, 2026). this platform integrates game-like features such as personalized avatars, challenges with time limits, and instat feedback to improve students' mental involvement (Santoso et al., 2025). The launch of ZEP QUIZ is also consistent with the Discovery Learning approach, which encourages analyical thinking and rapid decision-making in understanding both story and informational texts (Ramadhani & Sya, 2025). In this study, ZEP QUIZ serves as the main independent treatment variable. Its real-time multiplayer format actively contributes to the research by creating a vibrant, competitive atmosphere in which eleventh-grade students can participate in interactive contests with their peers. This setup allows for the immediate display of final scores, which acts as an instant measure of their actual vocabulary retention.

**Synthesis of Previous Studies and Present Research**

Numerous investigations have examined the acquisition of synonym vocabulary and gamified platforms, yet they exhibit distinc methodological and conceptual boundaries when compared to this study. To establish a clear critical evaluation, the similaraties and differences between earlier literature and the present research are synthesized across five core parameters: participants, learning context, platform, vocabulary focus, and research design. First, regarding vocabulary focus and instruction, Baltayeva (2024) discovered that involving learners in interactive exercises designed to teach synonyms significantly enhances both their communication skills and confidence. Although Baltayeva (2024) shares a similar focus on synonym instruction, that study operated within a descriptive qualitative framework, whereas the present research utilizes a rigorous quasi-experimental design to empirically quantify assessment outcomes. Similarly, in an experimental investigation, Kamalesh & C (2024) demonstrated that the integration of digital media within the Content and Language Integrated Learning (CLIL) framework greatly outperforms conventional approaches in helping students recognize the nuanced differences between synonyms. However, while Kamalesh & C (2024) share an experimental nature, their learning context operates within a broad curriculum framework and does not isolate a specific, software-based assessment tool. In a different setting, Hameed & Challob (2022) emphasized the importance of synonyms as a unifying lexical component in essay writing, but their research focused on writing production among university-level participants rather than targeted formative assessment among secondary school learners.

Second, regarding the specific application of the ZEP QUIZ platform, (Amajida et al. (2025) noted a significant increase in student engagement, rising from 63% to 93%. Nevertheless, the critical difference lies in the research focus; while Amajida et al. (2025) focused purely on a classroom action research (CAR) design to measure academic mastery. The finding is reinforced by Ratnawati & Wahyuni (2026), who observed that ZEP QUIZ not only improves students' comprehension but also lowers performance anxiety when compared to traditional paper-based assessments. Although Ratnawati & Wahyuni (2026) lay an important groundwork regarding cognitive benefits, their learning context targeted general reading comprehension rather than specific synonym vocabulary. In this study, ZEP QUIZ serves as the main independent treatment variable. Its real-time multipayer format actively contribures to the research by creating a vibrant, competitive atmosphere in which eleventh-grade student can participate in interactive contests with their peers. This setup allows for the immediate display of final scores, which acts as an instant measure of their actual vocabulary retention.

## METHOD

### Design and Samples

This study utilized a quasi-experimental framework, particularly a design featuring a non-equivalent control group. This method was chosen to assess the effectiveness of the ZEP QUIZ-based assessment in comparison to traditional assessment techniques, all while preserving the authentic classroom environment. Two pre-existing groups were involved: one experimental group and one control group. Initially, both groups participated in a pre-test ($O_1$) to set a baseline measurement. The experimental group then underwent treatment ($X$) using the ZEP QUIZ-based assessment, whereas the control group adhered to onventional assessment methods. Subseqeuntly, a post-test ($O_2$) was conducted with both groups to evaluate their progress.

The subjects of this study were eleventh-grade Social Studies students at MAS Simbang Kulon for the 2025/2026 school year. An intact group sampling technique was employed to select two existing classes for the study: class 11P8, which included 31 students, served as the experimental group, while class 11P9, consisting of 31 students, acted as the control group, resulting in a total of 62 participants. This non-probability sampling approach was strictly selected based on the school's administrative policy, which prohibited the random assignment of individual students to avoid disrupting the established classroom organization and academic schedule. To address the baseline equivalence and control potential pre-existing differences resulting from this non-random assignment (where the experimental group's pre-test mean appeared slightly higher), the researchers intgerated Normalized Gain Score (N-gain) analysis within the statistical framework. By analyzing the actual improvement ratio than raw post-test outcomes, the initial variance between the two intact groups was statistically neutralized.

### Instrument and Procedure

Data were gathered using a 20-item multiple-choice synonym vocabulay test fully developed by the author to align with the Phase F (Grade XI) Learning Outcomes of the Indonesian national curriculum. The test evaluates students' lexical depth across three essential word classes with distributed difficulty levels (Easy, Moderate, and Difficult), as detailed in Table 1.

*Table 1. Test Specifications and Difficulty Levels of the Synonym Vocabulary Instrument*

| Lexical Category | Target Synonm Items | Difficulty Level | Item Numbers | Total Items |
|---|---|---|---|---|
| Adjectives | Ambiguous, Durable | Easy | 2, 5 | 6 items |
| | Prominent, Rigid | Moderate | 1, 3 | |
| | Supercial, Hazardous | Difficult | 4, 6 | |

| | | | | |
|---|---|---|---|---|
| Verbs | Collaborate, Vanish | Easy | 11, 13 | 7 items |
| | Acquire, Commence | Moderate | 9, 10 | |
| | Maintain, Inhibit, Diminish | Difficult | 7, 12, 8 | |
| Nouns | Initiative, Outcome | Easy | 16, 17 | 7 items |
| | Component, Conflict | Moderate | 18, 19 | |
| | Obstacle, Prosperity, Objective | Difficult | 14, 15, 20 | |
| **Total** | | | | **20 items** |

Before conducting statistical analysis, the assessmet tool was reviewed qualitatively by an English language education lecturer to determine its face validity and the relevance of its content. The data collection process followed three operational stages: 1) Pre-test: A simultaneous 45-minute testing slot to evaluate foundational synonym mastery prior to intervention. 2) Treatment: Conducted over two separate sessions each lasting 90-minute, both sessions shared the same instructional content and time distribution. The only variation was confined to the last 30 minutes of each session: Experimental Group (11P8 – ZEP QUIZ): Students utilized a QR code or web link to access the platform on their smartphones, identifying themselves with their real names. Within the live virtual environment, they maneuvered 2D avatarrs to engage in synchronized live multiple-choice synonym tasks, with a time constraint of 45 seconds per question. Prompt cognitive feedback and correct answers were shown on the screen, accompanied by real-time updates to the leaderboard. This deliberate arrangement guaranteed that any effects observed were a dirrect result of the gamified framework and the synchronous nature of the activities. Control Group (11P9 – Conventional): Students responded to the same 20 multiple-choice questions during the identical 30-minute timeframe, using conventional paper-based worksheets. Formative feedback was provided through a traditional method, involving a teacher-led review session of answers at the end of the class period. 3) Post-test: A final 45-minute testing slot using the same protocol to assess ultimate vocabulary mastery. Regarding ethics, formal administrative permission was granted by the principal of MAS Simbang Kulon. Informed student assent was secured voluntarily, and participant confidentiality was strictly maintained by anonymizing all data.

**Data Analysis**

The quantitative data from the pre-tests and post-tests were statistically analyzed using SPSS 20 through four sequential steps. First, student scores were calculated on a 100-point scale and classified into five mastery levels: Poor (0 - 49), Less (50 - 59), Fair (60 - 74), Good (75 - 89), and Excellent (90 - 100) based on Arikunto's framework as cited in Maula (2024) this established classification provides a standardized educational metric to interpret the depth of students' lexical competence. Second, descriptive statistics, including means and standard deviations, were computed to summarize the students' baseline and final

vocabulary achievement. Third, a normality distribution check was conducted as a prerequisite test. Since the check revealed that the dataset violated normal distribution assumptions ($p \leq 0.05$), non-parametric inferential statistics were applied. Finally, a Wilcoxon Signed-Rank Test was executed to evaluate the internal progress within each group, and a Mann-Whitney U Test was performed on the gain scores to contrast the assessment effectiveness between the experimental and control classes.

## RESULT AND DISCUSSION

### Scoring Method and Mastery Classification

In accordance with the first step of data analysis, the raw scores of individual students were calculated and categorized into mastery levels. Instead of displaying extensive raw data, Table 5 outlines the distribution and frequency of students across the mastery thresholds specified by Arikunto as cied in Maula (2024).

*Table 5. Distribution of Students' Vocabulary Mastery Levels*

| Range | Qualification | Experimental (11P8) Pre-Test (f) | Experimental (11P8) Post-Test (f) | Control (11P9) Pre-Test (f) | Control (11P9) Post-Test (f) |
|---|---|---|---|---|---|
| 90 - 100 | Excellent | 7 | 27 | 1 | 2 |
| 75 – 89 | Good | 4 | 4 | 8 | 11 |
| 60 – 74 | Fair | 5 | 0 | 4 | 12 |
| 50 -59 | Less | 6 | 0 | 4 | 2 |
| 0 - 49 | Poor | 9 | 0 | 14 | 4 |
| Total | | 31 | 31 | 31 | 31 |

Table 5 reveals a drastic shift in the experimental group's mastery levels after the treatment. On the pre-test, the majority of experimental students clustered in the "Poor" (9 students) and "Less" (6 students) categories. However, on the post-test, 27 out of 31 students successfully advanced into the "Excellent" classification. Conversely, the control group exhibited a slower progression, with most students remaining heavily concentrated within the "Fair" (12 students) and "Good (11 students) threshold during the post-tes stage.

### Descriptive Analysis Summary

To offer a comprehensive summary of the changes in these scores, descriptive statistics were calculated. Table 6 displays the overall profile that includes the lowest and highest values, mean, and standard deviation of the datasets.

*Table 6. Distribution of Students' Vocabulary Mastery Levels*

| | N | Range | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|---|
| Pre-Test Esperiment | 31 | 100 | 0 | 100 | 62.94 | 26.835 |
| Post-Test Esperiment | 31 | 25 | 75 | 100 | 94.84 | 7.581 |
| Post-Test Control | 31 | 75 | 15 | 90 | 54.52 | 22.074 |
| Post-Test Control | 31 | 60 | 40 | 100 | 71.94 | 14.758 |
| Valid N (listwise) | 31 | | | | | |

The descriptive statistics show that the experimental group's mean score increased significantly from 62.94 to an outstanding final average of 94.84. Meanwhile, the control group showed a more modest increase, rising from a baseline of 54.52 to a post-test mean of 71.93.

**Prerequisite Analysis Tests**

Before evaluating the research hypotheses, a normality assessment was conducted on the datasets. The Shapiro-Wilk test was utilized to analyze the statistical distribution of the scores from both the pre-test and post-test for the experimental and control groups. The results indicated that the pre-test scores from both groups, as well as the post-test scores of the control group, conformed to a normal distribution ($p > 0.05$). However, the post-test scores of the experimental group yielded an empirical significance value of 0.000 ($p < 0.05$).

*Table 7. Tests of Normality*

| Class | | Kolmogrov-Smirnov[a] | | | Shapiro-Wilk | | |
|---|---|---|---|---|---|---|---|
| | | Statistic | df | Sig. | Statistic | df | Sig. |
| Result | Pre-test Experiment | 7 | 31 | .200 | 27 | 1 | .282 |
| | Post-test Experiment | 4 | 31 | .000 | 4 | 8 | .000 |
| | Pre-test Control | 5 | 31 | .091 | 0 | 4 | .077 |
| | Post-test Control | 6 | 31 | .048 | 0 | 4 | .281 |

This specific finding signifies that the post-test data for the experimental group significantly strayed from the expected normal distribution, thus contravening the

essential distributional assumption required for parametric statistics. As a result, in order to uphold methodological integrity, the homogeneity assessment was omitted, leading to a decisive shift in the analytical approach towards non-parametric inferential statistical techniques, specifically implementing the Wilcoxon Signed-Rank test and the Mann-Whitney U test.

**Inferential Analysis and Hypothesis Testing**

In order to assess the research hypotheses, non-parametric inferential methods were utilized given that the datasets did not exhibit a normal distribution. Initially, a Wilcoxon Signed-Rank Test was performed to examine the internal development and the statistical significance of the changes in vocabulary secores from the pre-test to the post-test within each group. The test statistics that were calculated can be found in Table 8.

*Table 8. Wilcoxon Signed-Rank Test Statistical Output*

| Comparison Dataset | *Z*-value | Asymp. Sig. (2-tailed) |
|---|---|---|
| Post-Test – Pre-Test (Experimental Class) | -4.518 | .000 |
| Post-Test – Pre-Test (Control Class) | -3937 | .000 |

As shown in Table 8, the Wilcoxon Signed-Rank Test performed on the experimental group resulted in a notably significant internal change ($Z$ = - 4.518, $p$ = .000). Given that the asymptotic significance level is considerably lower than the critical alpha threshold ($p < 0.05$), it is conclusively demonstrated that the implementation of ZEP QUIZ-based formative assessment led to a significant enhancement in students' mastery of synonym vocabulary over time.

At the same time, the control group also showed a meaningful baseline improvement ($Z$ = -3.937, $p$ = .000), suggesting that traditional worksheet-driven instruction also fostered positive vocabulary growth. Since both settings resulted in internal learning advancements, an additional analysis of relative progression employing Normalized Gain (N-gain) scores along with a Mann-Whitney U test is necessary to identify which method provided a statistically superior educational effect.

*Table 9. Descriptive Summary of Normalized Gain (N-Gain) Scores*

| Group | Mean N-Gain (%) | Std. Deviation | Minimum | Maximum | Effectivess Category |
|---|---|---|---|---|---|
| Experimental (11P8) | 85.40% | 26.15 | -25% | 100% | Effective |
| Control (11P9) | 28.71% | 46.59 | -100% | 100% | Ineffetive |

According to the simplified statistics presented in Table 9, the experimental group achieved a robust average N-gain score of 85.40%, positioning the ZEP QUIZ-based evaluation comfortably within the "Effective" category for enhancing students' vocabulary mastery of synonyms. In contrast, the control group recorded a significantly lower average N-gain score of merely 28.71, which classifies the conventional paper-based assessment approach as "Ineffective." While there were differences in scores among individuals in both groups, as shown by the lowest and highest scores, the overall results indicate that the gamified digital intervention fostered a significantly higher rate of advancement compared to conventional classroom techniques.

To empirically confirm if the observed differences between the two groups were statistically meaningful, a non-parametric Mann-Whitney U test was performed on the individual N-gain scores of the students. The definitive statistics derived from this computational analysis are presented in Table 9.

*Table 10. Mann-Whitney U Test Statistical Output*

| **Test Statistic** | **Operational Value** |
|---|---|
| Mann-Whitney U | 72.500 |
| Wilcoxon W | 568.500 |
| *Z* - value | -5.823 |
| Asymp. Sig. (2-tailed) | .000 |

As presented in Table 10, the Mann-Whitney U test generated an empirical value of 72.500 along with a *Z*-score of -5.823. Importantly, the asymptotic significance value resulted in .000 (Asymp. Sig. < .05). Since the empirical significance level is significantly lower than the accepted alpha level, the null hypothesis ($H_0$) is effectively rejected, leading to the acceptance of the alternative hypothesis ($H_a$). this statistical finding offers strong empirical support that the use of the ZEP-QUIZ-based assessment has a significantly greater impact on the mastery of synonym vocabulary among eleventh-grade students compared to traditional paper-based assessment methods.

The empirical results of this research indicate that the use of ZEP QUIZ for formative assessment significantly enhances the synonym vocabulary skills of elventh-grade students when compared to conventional paper-based assessment methods. This is evidanced by the notable disparity between the experimental group's mean N-gain score of 85.40% ("Effective") and the control group's score of 28.71% ("Ineffective"). These results highlight the unique educational and structural benefits found within interactive, gamified learning settings that promote rapid language development, rather than simply reitating numerical scores.

A key factor contributing to the vocabulary mastery of the experimental group is the structured format of ZEP QUIZ as an interactive, real-time educational platform. Instead of working on standard, individual worksheets, students

participate in synonym exercies within a lively, game-oriented environment along with their classmates. This multiplayer approach transforms conventional formative assessments into collaborative and competitive gameplay, fostering sustained cognitive engagement throughout the activity. From a cognitive standpoint, these interactive features directly accelerate vocabulary acquisition rather than merely raising student engagement. Specifically, the instant feedback allows students to recognize and correct lexical errors immediately, preventing the internal stabilization of incorrect synonym associations. Furthermore, the synchronous game challenges demand heightened attention and active participation, which naturally drives repeated mental exposure to the target words. This repetitive cognitive processing strengthens the neural pathways responsible for word association, allowing students to map, store, and recall complex synonyms far more efficiently than through passive paper-based drilling. Additionally, the instant calculation and presentation of final scores at the end of the game provide immediate academic feedback, encouraging a shared sense of accomplishment that reinforces vocabulary retention.

Critically, the outcomes of this investigation expand upon previous literature while offering nuanced insight into gamified pedagogy. The robust vocabulary gains align with prior findings by Zhu et al. (2023) and Udeozor et al. (2023) regarding the efficacy of digital learning tools, as well as the contemporary framework by Ratnawati & Wahyuni (2026) on interactive formative assessments in EFL settings. However, a critical comparison reveals that while earlier studies primarily attribute vocabulary improvement ot general student motivation, the data in this study suggests that the strict structuring of formative exercises within a live game interface is what drives the stabilization of learning outcomes. While traditional methodologies often suffer from declining student attention toward the end of a session, the interactive nature of ZEP QUIZ maintains operational focus, resulting in a more uniform upward shift in post-test scores

To achieve a well-rounded and academically valid viewpoint, it is essential to recogniz specific empirical and practical limitations. Primarily, the use of ZEP QUIZ is significantly dependent on reliable internet access and personal smartphone availability, which can lead to unexpected technical issues within traditional classroom environments. Secondly, the effective operation of the gamified sessions relies heavily on students' prior experience with digitial learning tools; those who are not accustomed may encounter some initial cognitive strain while navigating the platform. Thirdly, from an instructional perspective, creating and organizing synchronized synonym exercises on the platform demands considerable preparation time from instructors compared to using conventional printed worksheets. Finally, the notable vocabulary advancement observed in the experimental group might be somewhat influenced by the "novelty effect," where students display increased interest and higher performance levels due to the introduction of a new and unfamiliar digital tool, rather than genuine long-term understanding. Acknowledging these systematic limitations allows for a realistic

approach to the implementiation of gamified formative assessment in terms of expectations for lasting sustainability.

Lastly, regarding the emotional aspects of the classroom experience, it is important to recognize specific limitations. Although the gamified aspect of ZEP QUIZ may imply a decrease in test anxiety or a boost in motivation, such claims should be made with scholarly caution. Since the research did not employ dedicated psychometric surveys, structured interviews, or systematic observational methods to assess student anxiety and motivational levels, firm conclusions regarding anxiety alleviation cannot be confirmed. Nonetheless, it can be reasonably suggested that the platforms's formative, low-pressure approach promoted heightened student involvement. By presenting synonym tasks within a familiar and accessible smartphone interface, the design likely created a less daunting evaluative setting compared to traditional paper-based worksheets, enabling students to concentrate their mental efforts on lexical challenges rather than on the stress of evaluation.

## CONCLUSION

This research concludes that utilizing ZEP QUIZ-based formative assessment is very affective in improving students' understanding of synonyms, as demonstrated by a significant average N-gain score of 85.40% ("Effective") in the experimental group, compared to only 28.71% ("Ineffective") in the control group. The non-parametric inferential analysis conducted using the Mann-Whitney U test provides empirical evidance for this difference ($U$ = 72.500, $P$ = .000), indicating that the engaging, interactive nature of the game and the immediate feedback from the scoring system fosters better vocabulary retention than traditional paper-based approaches. However, certain limitations should be recognized to prevent overgeneralization. This study was specifically limited to a quasi-experimental design involving just two intact classes ($N$ = 62) at a single secondary school, MAS Simbang Kulon, during a relatively short teaching period. Additionally, the assessment focused solely on synonym mastery without examining broader vocabulary aspects or other essential language skills, and student emotional factors such as anxiety or motivation were not thoroughly assessed using specialized psychological tools.

Based on these findings, practical suggestion are made for English language teaching and futre research endeavors. English instructors are strongly encouraged to incorporate real-time, game-like digital assessment tools such as ZEP QUIZ into their reguler formative evaluation processes, as the interactive multiplayer format effectively captures student engagement and facilitates quick performance assessments through instant grading. For future researchers, it is advisable to broaden this investigational framework by including larger and more varied student populations, extending the intervention duration, and using independent methods to collect psychological data. Furthermore, upcoming studies should examine the effectiveness of gamified platforms in other areas of vocabulary development, such

as collocations and antonyms, as well as different macro-language skills like listening and reading comprehension.

## REFERENCES

Adipat, S., Laksana, K., Busayanon, K., Asawasowan, A., & Adipat, B. (2021). Engaging students in the learning process with game-based learning: The fundamental concepts. *The International Journal of Technology in Education, 4*(3), 542–552. https://doi.org/10.46328/ijte.169

Amajida, A., Khoiri, N., & Rozzaqi, A. R. (2025). Efektivitas media ZEP QUIZ dengan model pembelajaran Teams Games Tournament terhadap motivasi dan hasil belajar matematika di PKBM Putra Bangsa Demak. *Papanda Journal of Mathematics and Sciences Research (PJMSR), 4*(2), 425–434. https://doi.org/10.56916/pjmsr.v4i2.2944

Baltayeva, N. T. qizi. (2024). Methodology for developing students' ability to learn from synonyms. *Mental Enlightenment Scientific-Methodological Journal, 5*(3), 37–42. https://doi.org/10.37547/mesmj-V5-I3-06

Boontam, P. (2022). The effect of teaching English synonyms through data-driven learning (DDL) on Thai EFL students' vocabulary learning. *Shanlax International Journal of Education, 10*(2), 80–91. https://doi.org/10.34293/education.v10i2.4458

Hameed, Y. S., & Challob, I. (2022). Synonymy as a cohesive device in students' English essay writing. *Al-Adab Journal, 2*(140), 91–110. https://doi.org/10.31973/aj.v2i140.3632

Kamalesh, F. J. D., & C, S. (2024). Intertextualizing interactive texts for teaching synonyms using digital platforms: An experimental study engaging the first-year undergraduate students. *Journal of Language Teaching and Research, 15*(2), 674–683. https://doi.org/10.17507/jltr.1502.36

Korkmaz, S. (2021). Using Kahoot to improve reading comprehension of English as a foreign language learners. *International Online Journal of Education and Teaching (IOJET), 8*(2), 1138–1150.

Lee, T., Song, G., Ahn, H., Oh, H., Byun, M., Han, E. C., Kim, S., Kim, C. W., Kim, H. J., Hong, S., & Song, K. (2024). Public effect of the 2022 colorectal cancer awareness campaign delivered through a metaverse platform. *Annals of Coloproctology, 40*(2), 145–153. https://doi.org/10.3393/ac.2023.00122.0017

Maula, I. (2024). The effectiveness of word search puzzle to teach vocabulary (an experimental research of 11th grade of TKR major at SMK Muhammadiyah Kajen). *8*(6), 787–792.

Muhtarom, & Maghfiroh, A. (2022). Student difficulties in English learning during limited PTM period. *International Journal of Research on English Teaching and Applied Linguistics, 2*(2), 10–14. https://doi.org/10.30863/ijretal.v2i2.2450

Muryani, R. N., & Yunus, M. (2024). Students' digital vocabulary learning: Experiences, barriers, and adaptive strategies exploration. *English Review: Journal of English Education, 12*(1), 31–42.

https://doi.org/10.25134/erjee.v12i1.9328

Obata, A. (2022). Effects of synonyms in English as second language in Kenya: A critical literature review. *European Journal of Linguistics, 1*(1), 26–36. https://carijournals.org/journals/ejl/article/view/892/3526

Ramadhani, A. D., & Sya, M. F. (2025). Analisis isi modul ajar berbasis game-based learning: Pemanfaatan media AI Zep Quiz untuk meningkatkan pemahaman teks naratif. *Karimah Tauhid, 4*(9), 6980–6990.

Ratnawati, D., & Wahyuni, S. (2026). The effect of 'Zep Quiz' in assessing students' reading comprehension skill. *Journal of Educational Sciences, 10*(1), 468–479. https://doi.org/10.31258/jes.10.1.p.468-479

Santoso, D. R., Tamara, P. R., Nabhila, C. A., Ningsih, A. W., Sujarwo, G. D., & Fadlilah, N. L. Q. (2025). ZepQuiz game-based learning for improving students' narrative reading skills: ZepQuiz: Pembelajaran berbasis permainan untuk meningkatkan keterampilan membaca naratif siswa. *Academia Open by Universitas Muhammadiyah Sidoarjo, 10*(1), 1–16.

Sari, A. S. P., & Pandiangan, S. R. (2021). The application of guessing strategy to improve students' vocabulary mastery on English phrasal verbs. *Jurnal Sulu Pendidikan, 9*(1), 24–38.

Sari, A. S. P., Sembiring, N., & Wau, K. (2022). Improving students' vocabulary mastery on synonyms through the application of Think Pair Share (TPS) technique. *Budapest International Research and Critics Institute-Journal (BIRCI-Journal), 5*(2), 16466–16475. https://doi.org/10.33258/birci.v5i2.5561

Setyaningrum, T. D., Saputro, H., & Apriani, D. R. (2025). Implementation of the Team Games Tournament (TGT) learning model assisted by Zep Quiz media in science learning to improve students' learning motivation in Grade VIII at SMP Muhammadiyah. *Proceedings of International Conference on Teacher Profession Education Yogyakarta, 3*(1), 820–831.

Siregar, F. L., & Henni. (2023). An investigation of EFL students' vocabulary size and level at an Indonesian private tertiary institution. *Eralingua: Jurnal Pendidikan Bahasa Asing dan Sastra, 7*(2), 305–315. https://doi.org/10.26858/eralingua.v7i2.41379

Sun, D., Chen, Z., & Zhu, S. (2023). What affects second language vocabulary learning? Evidence from multivariate analysis. *Frontiers in Educational Research, 11*, 1–12. https://doi.org/10.3389/feduc.2023.1210640

Syahrizal, S., Yasmi, F., & Mary, T. (2024). AI-enhanced teaching materials for education: A shift towards digitalization. *International Journal of Religion, 3538*(1), 203–217. https://doi.org/10.61707/j6sa1w36

Udeozor, C., Abegao, F. R., & Glassey, J. (2023). Measuring learning in digital games: Applying a game-based assessment framework. *British Journal of Educational Technology, 55*(3), 957–991. https://doi.org/10.1111/bjet.13407

Zhu, S., Guo, Q., & Yang, H. H. (2023). Beyond the traditional: A systematic review of digital game-based assessment for students' knowledge, skills, and affections. *Sustainability (Switzerland), 3*, 1–19. https://doi.org/10.3390/su15054693.