Exploring the Impact of Gamified Artificial Intelligence–Driven English Vocabulary Learning Systems on Learner Retention and Motivation

Mrs. Gogineni Aswini[1], Dr. Madhu Munagala[2], V.Saranya[3], Keerthana R[4],
Aseel Smerat[5], Dr. Vinisha Sumra[6], Ahmed I. Taloba[7]
Research scholar, Department of English, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India [1]
Assistant Professor of English, Dhanekula Institute of Engineering and Technology, Vijayawada, AP, India [1]
Research Supervisor & Associate Professor, Department of English, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India[2]
Assistant Professor, Department of English, Panimalar Engineering College, Chennai, India[3]
Assistant Professor, Department of Information Technology,
Vel Tech High Tech Dr Rangarajan Dr Sakunthala Engineering College, Chennai, India[4]
Faculty of Educational Sciences, Al-Ahliyya Amman University, Amman, 19328, Jordan[5]
Department of Biosciences, Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences, Chennai, 602105, India[5]
Assistant Professor, Department of Computer Science and Engineering, MMEC,
Maharishi Markandeshwar (Deemed To Be University), Haryana, India[6]
Department of Computer Science, College of Computer and Information Sciences, Jouf University, Saudi Arabia[7]
Information System Department-Faculty of Computers and Information, Assiut University, Assiut, Egypt[7]

Abstract*—The growing development of digital learning platforms has posed an increased demand on gamified and artificial intelligence-based methods of enhancing English vocabulary learning. However, existing studies often treat gamification and AI as loosely pair components, relying on static game mechanics or post-hoc analytics that limit personalization, adaptability, and long-term learning impact. To address these limitations, this study proposes the Gamified AI-Driven Vocabulary Retention and Motivation Enhancer (GAI-VRME), an adaptive learning framework that integrates machine-learning–based learner modeling, real-time difficulty calibration, and adaptive gamification strategies. In contrast to the previous systems, GAI-VRME can dynamically regulate the complexity of the task, the frequency of feedback and the sequence of rewards according to the performance and the motivational state of a specific learner, and can thus be constantly customized to the individual level as the process of learning progresses. The implementation and empirical assessment of the framework were conducted with the help of Python, TensorFlow, and Jupyter Notebook and Teaching-Learning Gamification Dataset of Mendeley Data. Mixed method analysis of vocabulary retention with paired t-tests and sentiment-analysis-based motivation modelling was used. The experimental outcomes show that GAI-VRME has much higher predictive accuracy, vocabulary retention, and learner motivation than the traditional gamified systems. These findings provide empirical evidence that deeply integrated AI-driven adaptive gamification, jointly optimizing cognitive retention and affective engagement, offers a scalable and pedagogically robust solution for modern digital vocabulary learning environments.

Keywords*—*AI-driven gamified learning; adaptive educational systems; English vocabulary acquisition; learner motivation and engagement; vocabulary retention

I. Introduction

The high rate of growth of digital learning classrooms has fueled the need to explore instructional methods that would effectively promote learning success as well as learner interest [1]. Gamified learning is one of these methods that have shown potential in boosting motivation and maintaining the engagement of learners since they involve the use of the game theory of motivation including points, levels, badges, and challenges in the educational process [2] [3]. Simultaneously, artificial intelligence (AI) has emerged as a primary facilitator of personalization to the learning systems, promoting adaptive content delivery, tracking learner performance, as well as personalized feedback [4], [5]. The meeting of gamification and AI has thus gained an increasing interest in the field of educational research, especially in the field of English vocabulary learning, which makes a structural part of language mastery and academic achievement [6], [7]. Regardless of this accumulating body of work, published research has a number of weaknesses that limit the effectiveness and generalizability of gamified AI-based systems of vocabulary learning in the long term [8] [9]. To begin with, most gamified vocabulary systems use static or slightly adaptive games, providing homogenous difficulty levels and game reward systems that fail to capture the personality of a specific learner [10]. Second, although AI methods are often included, they are often simple analytics or recommendation systems without being integrated in a systematic way with pedagogically significant outcomes like long-term vocabulary retention [11]. Third, and most importantly, previous studies tend to measure system effectiveness through one-dimensional measures, which are learning performance or engagement measures, but rarely both in a combined fashion [12] [13].

The process of vocabulary learning is, nevertheless, intrinsically conditioned by cognitive factors, i.e., the ability to remember and master a vocabulary, as well as by the affective ones, i.e., motivation, interest, and emotion involvement. The gains in test scores do not always correspond to significant and sustainable learning when learners are not motivated to proceed with the system interaction or to use learned vocabulary in real-life situations. On the contrary, high engagement in the form of learning gains which are not proven increases the question of pedagogical effectiveness. In turn, the evaluation frameworks that would allow studying retention and motivation together are in high demand, so that the impacts of AI-based gamified systems on the language learning outcomes can be better understood holistically.

A. Research Question

The research examines the impacts of gamified and AI-based learning of English vocabulary on learning results. In particular, it deals with the following research question:

What is the effect of adaptive AI-based gamification on vocabulary retention and learner motivation in English as a Foreign Language (EFL) learners?

B. Research Gap

Even though previous studies have investigated the field of gamification, artificial intelligence and adaptive learning in teaching English vocabulary, the majority of systems provide the integration of these elements in a loosely paired or fixed fashion [14]. Gamification tends to be rule-based, whereas AI is utilized mainly to do post-hoc analytics or-recommendation, and provides limited pedagogical adaptation-in-the-moment [15]. Furthermore, motivation and vocabulary retention are usually measured as independent outcomes. The proposed GAI-VRME framework can be used to fill these gaps by integrating learning modeling as a machine-learned learner model and dynamically adaptive gamification. Through continually adjusting the task difficulty, feedback and reward systems depending on the performance of the learner and motivation extent, and collaborative assessment of the cognitive retention and affective involvement, GAI-VRME contributes to the development of intelligent and learner-oriented systems of vocabulary learning.

C. Research Motivation

The increase in the use of digital learning spaces has led to an increase in the benefits of vocabulary learning strategies that can facilitate the overcoming of personal engagement and long-term learning [16]. Traditional vocabulary teaching processes are usually based on the fixed material and homogenous level of challenge so that the motivation of learners will be minimized and the knowledge will not be retained [17]. Despite the adaptive and interactive learning experiences, adaptive and interactive learning systems based on AI are often considered separately in the literature, even though they should be evaluated collectively. The research is inspired by the fact that there is a lack of analytic methodology that considers both cognitive retention and learner motivation in adaptive gamified learning systems of vocabulary.

D. Research Significance

The importance of the present study is that it presents a comprehensive assessment model of AI gamified vocabulary learning in which the measured variables are the retention and motivation of the learner. The proposed solution incorporates statistical retention analysis and sentiment-driven motivational evaluation, which leaves the realm of classical single-metric-based assessment. The results provide empirical research on the effect of adaptive gamification on cognitive and affective learning outcomes. The work also adds a reproducible evaluation model to help educators, system designers, and researchers to create successful and learner-centered digital vocabulary learning settings.

E. Problem Statement

Even though gamified and AI-driven language learning systems have gained an interest now, empirical evidence regarding their effectiveness is inconsistent [18]. A number of intervention studies indicate positive changes in learner engagement, but insignificant increases in vocabulary acquisition, and that gamification is not an assurance of meaningful learning without pedagogical underpinning and adaptive design. Meta-analyses also show that there are only significant overall effects of gamification on intrinsic motivation, significant levels of heterogeneity, and a lack of support of gamification on learner autonomy and competence [19] [20]. These gaps warrant the need to conduct research that clearly relates adaptive AI-based gamification processes to validated retention scales and motivation predictors that are theoretically based. This paper fills this gap by using a mixed retention-sentiment assessment test.

F. Key Contribution

- Introduces GAI-VRME a single framework combining real-time learner modeling and adaptive gamification processes.
- Introduces dynamic adaptation logic recalibrating difficulty, feedback, and rewards using continuous learner performance signals.
- Establishes a hybrid cognitive - affective assessment regime involving the combination of retention rates with sentiment-driven motivation.
- Empirically proves that the trajectories of learner motivation are correlated significantly with the improvement in vocabulary retention.

G. Rest of the Section

- Section II includes a thorough overview of the current studies in the area of gamified learning, AI-based vocabulary system, motivational psychology and mixed-method evaluation methods.
- Section III elaborates all the methodological framework, such as set up of data, pre-processing, quantitative retention analysis, qualitative motivation assessment, and synthesis of both results.

- Section IV presents the empirical results of statistical tests and sentiment analysis with the aid of tables, graphics, and comparison knowledge.
- Section V brings the study to the conclusion with a conclusion on key contributions, practical implications of AI-based vocabulary learning and the future directions of the research improvement and expansion.

II. Literature Review

Sadeghi et al. [21] suggested a research to examine the impact of the application of the gamified instruction to language learners in terms of vocabulary acquisition and their motivation. The research design was quasi-experimental design with participants being schools where pre-test and post-test were taken as the dataset. This was conducted to investigate whether the gamification methods would be able to increase the vocabulary retention and the motivation of the learners. The technique incorporated gamified learning intervention, which consisted of points, badges, and leaderboards, and traditional teaching as a control group. Findings showed a great deal of improvement in vocabulary acquisition during the experimental group as opposed to the control group, and learners also reported a greater level of engagement. The limitations of the study are a small sample size and school-specific setting (which limits the generalizability of the research). The originality is in the fact that gamification is applied directly to the teaching of EFL vocabulary in a classroom setting.

Putri Fachri Aulia Fatah [22] the investigation of the efficacy of gamification in improving the results of learning the English language. The study was performed on 60 participants of a school and it was conducted as the quasi-experimental one with a structured pre-test/post-test evaluation. This was done to determine whether gamified instructional strategies had the potential to enhance vocabulary retention and general learning outcomes. In the research, the methodology included both game-based activity and scoring because it would involve the engagement of the learners and the results would be compared to the traditional forms of teaching. The results indicated that there was a substantial increase in the vocabulary test scores and that the learner's findings were more satisfied with the gamified learning, potentially because gamified learning promotes more performance and engagement. Weaknesses are that the sample used is relatively small and homogeneous, and the intervention is quite short, which limits the extent to which the results can be generalized. The originality of the study lies in the combination of gamatherics in an English curriculum in school.

Kumar and Vairavan [23] presented an experimental research on the effect of gamification on motivation and retention in language learning with the help of a gamified language learning application. The data were collected through a quasi-experimental design involving 100 participants, and entailed the survey results of the respondents and performance outcomes. The aim of the study was to establish what the effectiveness of a mobile gamified app is in contrast to the traditional methods. Techniques used included interactive challenge, points, and progress monitoring in the app, pre- and post-intervention assessment in terms of retention and motivation. The findings showed that positive changes were recorded, with the experimental group having better scores at the post-test level and having a greater engagement level. Weaknesses are the use of self-reported survey data and limited application to a specific context. These are novel approaches where digital gamification is used in conjunction with empirical retention and motivation measurement in a controlled experimental design.

Chen and Zhao [24] suggested conducting research to know how Chinese EFL learners accept gamified vocabulary learning apps. This was to combine Self-Determination Theory (SDT) and the Technology Acceptance Model (TAM) in order to investigate the perception and engagement of the learners. The data used consisted of the answers to the surveys that were carried out among Chinese learners applying gamified apps in both formal and informal contexts. The techniques used structural equation modeling (SEM) to analyze the interrelation between perceived usefulness, autonomy, competence, and motivation. The findings showed that the intrinsic motivation and the perceived simplicity of the app had a significant impact on the adoption and use of the app, which is why the key aspects of gamified learning adoption are psychological. Limitations are based on self-reported perceptions, use of survey data where there is no performance measurement. The innovativeness is the integration of SDT with TAM to study the gamification acceptance in an EFL situation.

Yu [25] suggested research on the topic of learning outcomes, motivation, and satisfaction in gamified learning of English vocabulary. The research question was to investigate the role gamified instruction plays in cognitive and affective aspects of learners. The data was composed of the pre-test and post-test vocabulary scores and survey data that included motivation and satisfaction. Gamified elements like points, levels and feedback were incorporated into the quasi-experimental design. The findings revealed that the vocabulary retention and high satisfaction and motivation among learners became statistically significant, which points to a positive correlation between gamification and engagement. The weaknesses are a small group of participants and a brief intervention period. The originality of the research is the joint analysis of learning outcomes, motivation, and satisfaction of a gamified EFL learning space.

Liu [26] explored the implications of AI-based gamified learning on English as a Foreign Language (EFL) vocabulary retention and dynamic motivational patterns. The aim of the research was to make comparisons between adaptive learning paths, conversational agents and storytelling strategies in AI-enhanced gamified settings. It was conducted using an experimental research design that involved performance evaluation and longitudinal motivation study. The dataset was in the form of learner interaction logs and vocabulary performance scores obtained through an AI-based gamified learning platform. The adaptive learning pathways made a significant impact on vocabulary retention in the study, whereas storytelling methods had greater intrinsic motivation. Nevertheless, the results were confined to platform dependence, inconsistency of AI interactions, as well as a lack of applicability to a wider range of educational settings.

Caiza et al. [27] suggested a research that assessed an interactive app combining the use of Virtual Reality (VR) and the Artificial Intelligence (AI) to enhance the pronunciation of the English language. The aim was to determine the potential of the immersive gamified experiences to improve the language acquisition outcomes that is, pronunciation accuracy. The data consisted of performance indicators, user feedback, and self-reported feedback. Techniques were experimental testing, VR-based pronunciation training under AI tutoring, and the pronunciation accuracy level of participants at the start and the end of the intervention. Findings indicated a significant change in the pronunciation scores and increased interest of learners in immersive game-based aspects. Limitations are the constraints imposed by the technology availability and low sample sizes. The uniqueness is that VR and AI are introduced in gamified language learning to help students in pronunciation training.

Liu et al. [28] performed a mixed-method research to evaluate the effect of motivation and enjoyment on AI-mediated informal digital learning of English (AI-IDLE) in Chinese learners. The study was intended to comprehend the impact of gamified AI tools on informal learning of vocabulary, learner engagement, and fun. The quantitative analysis of vocabulary learning combines with the qualitative sentiment analysis and thematic analysis of the feedback of the learners. The data represented performance scores, interaction records, and open-ended responses by learners with the help of AI-based gamified applications. The research proved that the increase in the level of enjoyment and motivation had a strong positive correlation with better vocabulary performance. However, its weaknesses were the possible self-selection bias and varying degrees of informal learning settings.

TABLE I. LITERATURE REVIEW OF GAMIFIED AI-BASED ENGLISH LEARNING

Author & Year	Purpose	Findings	Limitations
Sadeghi et al. [21]	Examine whether gamified instruction improves vocabulary retention and motivation in school learners	Significant improvement in vocabulary scores; increased learner engagement	Small sample, school-specific context limits generalizability
Putri Fachri Aulia Fatah [22]	Assess the effect of gamified activities on English learning outcomes	Improved vocabulary test scores and learner satisfaction	Small, homogeneous sample; short intervention duration
Kumar and Vairavan [23]	Explore motivation and retention using a gamified app	Higher post-test scores and motivation in experimental group	Reliance on self-reported surveys; context-specific app use
Chen and Zhao [24]	Investigate learners' acceptance of gamified apps using SDT and TAM	Intrinsic motivation and ease of use positively influenced adoption	Survey-based data; lacks performance measures
Yu [25]	Examine learning, motivation, and satisfaction in gamified settings	Significant improvement in vocabulary retention; high satisfaction	Narrow participant demographic; short duration
Liu [26]	Compare adaptive learning, conversational agents, and storytelling effects	Adaptive learning and storytelling improved retention and engagement	Limited generalizability; variability in AI interactions
Caiza et al. [27]	Evaluate VR and AI-based gamified app for pronunciation	Improved pronunciation scores and engagement	Small sample; technology accessibility constraints
Liu et al. [28]	Examine motivation and enjoyment in informal AI-based learning	Higher motivation and enjoyment correlated with better learning outcomes	Self-selection bias; informal learning variability

Table I reviewed literature highlights that gamified and AI-based learning systems significantly increase vocabulary retention, motivation, and learner satisfaction in a variety of contexts. Studies consistently show that gamification elements like points, badges, storytelling, adaptive learning paths, and VR/AI applications improve engagement and cognitive outcomes. However, limitations such as small, homogeneous samples, short intervention periods, reliance on self-reported surveys, and context-specific platforms restrict generalization. Despite these obstacles, the innovation lies in integrating gamification with AI, adaptive learning, and immersive technologies. Collectively, these studies provide strong empirical support for implementing gamified AI-powered approaches in English vocabulary instruction, justifying their adoption in the current study. Review summarizing limitations of prior studies, such as small samples, short interventions, or separate evaluation of retention and motivation, and explain how the current study addresses these gaps with integrated AI-based gamified learning.

III. METHODOLOGICAL FRAMEWORK FOR GAMIFIED AI VOCABULARY LEARNING

As opposed to traditional gamified learning assessment tools which view adaptation and assessment as non-adaptive or non-evolving processes, the proposed methodology modulates the learner behavior as a dynamic, evolving process. Learner modeling based on AI is used to implement modifications to gamification mechanics in real-time, and retention and motivation are measured as interdependent variables and not independent ones. Such an approach to the methodology allows exploring the effects of adaptive decisions in more depth and understanding the impact of such decisions on both the cognitive learning gains and the affective engagement in the long-term. The study is based on a systematic mixed-method technique to investigate the effectiveness of an artificial intelligence-based system of gamified English vocabulary learning in terms of retention and motivation of learners. A quantitative-qualitative type of research method was utilized to allow a thorough study that would cover not only the objective results of learning but also the subjective experiences of the learners. To assess all empirical findings, the Teaching-Learning Gamification Dataset at Mendeley Data was used and offers comprehensive records of learner, interactions, vocabulary performance, and motivational feedback in a gamified learning setting to establish statistical hypothesis testing. Quantitative analysis was aimed at studying the vocabulary retention gains in terms of paired performance comparisons, whereas qualitative analysis was to be made concerning the motivational reactions in terms of interpretation of the learner feedback using sentiment as the measuring instrument. Moreover, machine-learning-assisted modelling was used to find patterns of behavior and plot performance trends among learners. The combination of the quantitative results with the qualitative data allows conducting a multidimensional analysis of the learning effectiveness to gain both empirical and contextual insights. All the methodological steps were implemented in accordance with the study objectives in order to make them accurate, transparent and reproducible. In general, this systematic approach to methodology is not only able to measure the educational efficiency of AI-based gamification, but also has a replication trail to direct such research in the future technology-enhanced language learning. To make it robust, mean imputation and z-score normalization were used to preprocess the learner inputs to eliminate missing values, outliers, and inconsistencies. These measures reduce the influence of noisy or adversarial data to increase stability of the model, however, extreme or adversarial inputs can still have an impact on predictions.

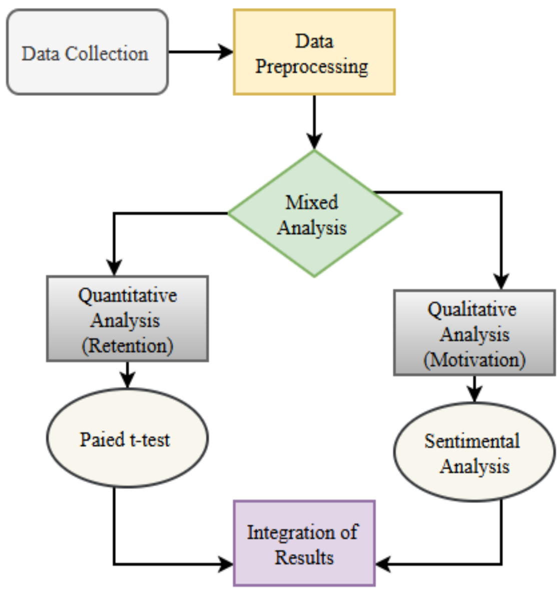

Fig. 1. Workflow of the gamified AI vocabulary learning.

The Fig. 1 depicts a mixed-methods research process in which gathered data is preprocessed, then subjected to quantitative analysis using paired t-tests and qualitative analysis using sentiment analysis. The results are then integrated to produce thorough, verified insights.

A. Learner Modeling and AI Architecture

In GAI-VRME, learner modeling is based on the well-defined input features, such as pre-test and post-test vocabulary scores, the participation rates, and completion rates, interaction logs and sentiment-based motivation indicators. The neural network used in the model architecture has fully connected layers and ReLU activation functions, which are optimized to learn patterns of cognitive performance and affective signals. The training of the system is in such a way that it predicts various targets such as post-test vocabulary performance, likelihood of completing task and real-time motivational state. The specifications of features, architecture, and prediction targets allow transparency, reproducibility, and creates a solid base of adaptive task calibration and individualized gamified learning paths.

B. Research Design

The research design is based on a mixed-method quasi-experimental research design, aiming to examine the impact of a gamified system of learning English vocabulary enhanced with AI on retention and motivation of learners. The method combines both the quantitative and qualitative analysis in the form of the same group of participants, which allows a thorough evaluation of the cognitive and affective learning results. The quantitative strand measures vocabulary retention by comparing pre-test and post-test vocabulary by administering Paired Sample t-tests to statistically effectively screen the evidence of improvements in learning between the matched participants. The qualitative strand will measure the motivation of the learners through sentiment analysis of open-ended textual responses, which will record the engagement, enjoyment, and emotion towards the gamified learning procedure. Its non-random assignment, which is similar to the quasi-experimental nature, is a way of reflecting the real-life classroom situation, where learners are free to engage with AI-driven gamification tools. The research design is effective because it combines numerical performance indicators and sentiment-based evaluation, so the results of retention are understood in a comprehensive and practical way and provide an objective picture of AI-based gamifier interventions in helping students to improve vocabulary acquisition in technology-enhanced learning settings.

C. Dataset Description

The study uses the Teaching-Learning Gamification Dataset from Mendeley Data [29] https://data.mendeley.com/datasets/yc4np572zs/1 , a publicly available dataset that records detailed learner interactions in gamified English vocabulary activities. It includes both quantitative data, such as pre-test and post-test vocabulary scores, participation rates, task completion, and performance metrics, and qualitative data, such as learner feedback on engagement, enjoyment, and perceived difficulty. This combination allows a mixed-method analysis, examining cognitive retention through statistical measures and affective motivation through sentiment evaluation. The dataset is fully anonymized, ethically sound, and supports robust analysis, making it ideal for evaluating the effectiveness of AI-mediated gamified vocabulary learning.

D. Dataset Accessibility and Reproducibility

The publicly available Teaching-Learning Gamification Dataset of Mendeley Data is used in the study, which allows transparency and reproducibility. The hybrid methodology will be based on the synthesis of the quantitative analysis of retention and qualitative analysis of motivation, which will enable the effective examination of the results of the AI-based gamified learning of English vocabulary.

E. Data Preprocessing

Preprocessing of data involved in this research is a very important phase as it will establish that the data obtained in Mendeley Data is accurate, consistent and is fit to undergo quantitative and qualitative analysis. As the research will be examining retention and motivation based on numerical scores and textual feedback, preprocessing aims at preparing each type of data to be interpreted reliably. This involves detecting and fixing missing/inconsistent values, outliers, normalizing score-based data and cleaning textual responses to use in sentiment analysis. Preprocessing increases the statistical accuracy of the analysis by converting the raw data into a standardized and noise-free data format, which increases the credibility of the retention and motivation results in the gamified AI learning environment.

1) Data cleaning and handling missing values: Handling missing values and cleaning the data make sure that the dataset is complete, consistent and retention and motivation analysis can be done accurately. Numerical variables like pre-test and post-test scores are verified on the missing values and imputed with the help of mean substitution to maintain the distribution of the scores. The formula of mean imputation is in Eq. (1):

$$x_{\text{new}} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1}$$

where, x_{new} is the inferred value of a new score that is used in place of a missing score. x_i is used to denote each of the possible existing numerical scores in the dataset, and (n) is used to denote the total amount of valid observations. The combination of the two gives a balanced mean value that ensures that the dataset is distributed correctly to be analyzed.

Sentiment analysis is performed by cleansing the text of symbols, digits, and other meaningless characters to remove noise before the analysis. The resulting data quality, decreased bias, and valid integration of quantitative and qualitative insights are a benefit of this combined cleaning process.

2) Outlier detection and normalization: Outlier extraction and normalization is what makes sure that the numerical information in this research (pre-test and post-test vocabulary scores) are similar and consistent across students. The z-score technique is applied to detect outliers, so that all the values above some threshold (which is usually 3) are treated as extreme. The z-score is computed in Eq. (2):

$$z = \frac{x-\mu}{\sigma} \tag{2}$$

where, x is an individual score of a learner, μ is the mean of the scores, and σ is the standard deviation, the purpose of which is to eliminate or normalize all the numerical characteristics to ensure the same consistency in the analysis of the Paired t–test. The normalization equation is given in Eq. (3):

$$x_{\text{scaled}} = \frac{x-x_{min}}{x_{max}-x_{min}} \tag{3}$$

where, x_{scaled} is the normalized score on a range of 0 to 1 guaranteeing that the scores are comparable across learners. All these variables, combined, contribute to the clean and objective data preparation.

This action will retain the relative differences in scores, minimizes the effect of extreme values, and puts data in the right format to compare the data statistically. The original score is denoted by x once more whereas xmin and xmax are the lowest and highest values in the dataset in the Min-Max normalization formula.

3) Text cleaning and lemmatization: The qualitative feedback in this study is cleansed and lemmatized using text cleaning, and it is sent to sentiment analysis. Text cleaning is used to clean the noise that is found in the text by removing symbols, numbers, extra spaces, and any unwanted characters and keeps only the meaningful words. The sentences are tokenized into individual words and the stopwords are removed to leave the non-informative words like the or and, and then lemmatization is performed to change each and every word to its base form; running becomes run. The following process can be stated in Eq. (4):

$$L(w_i) = w_{\text{root}} \tag{4}$$

where, $L(w_i)$ is the mapping of each word w_i to the root form of the word. The steps make motivation analysis clear, consistent, and reliable.

4) Sentiment lexicon preparation: Sentiment lexicon preparation will make sure that every processed learner feedback record can be assessed with motivational polarity in this case. VADER sentiment model is loaded so as to assign a polarity value of positive, negative, and neutral emotions. A compound sentiment score is calculated by analyzing each text of feedback, and this is in the form of Eq. (5):

$$S_{\text{compound}} = \frac{S_{\text{positive}} - S_{\text{negative}}}{S_{\text{total}}} \tag{5}$$

where, S_{positive} is the cumulative strength of positive words in the feedback of a learner and s_{negative} is the cumulative strength of negative words. S = The total magnitude of all the sentiment components. The resulting S_{compound} is a measure of the sum of emotional polarity of the learner feedback.

This score defines the general feeling category of every response. By creating such a lexicon, one can relax in the fact that the patterns of motivational behavior can be labeled in a uniform manner and that their emotional engagement of learners can be meaningfully interpreted in the context of the gamified AI-based vocabulary learning environment.

F. Quantitative Analysis: Retention Measurement

The vocabulary retention in this research is assessed through the qualitative analysis of the learners who are exposed to the gamified AI based English learning system, based on the Teaching-Learning Gamification Dataset given by Mendeley Data. The pre-test and post-test learners were compared to assess cognitive improvement, and the Paired Sample t -test was used to establish the statistical significance of the difference in scores among the two related measures. All data were preprocessed before analysis i.e. it was cleaned, outlier detected and normalized so that all data was accurate and comparable across the participants. The difference in the calculated scores actually measures individual learning improvements, the t-test measures the overall learning improvements which can be attributed to the intervention. Such quantitative findings are further supported by the qualitative sentiment analysis of the learner feedback, which offers a well-rounded, multidimensional analysis of the system in terms of its efficacy in improving the learner vocabulary retention and engagement.

1) Data extraction: Teaching- Learning Gamification dataset was used to obtain the corresponding data which were then analyzed quantitatively to determine vocabulary retention. Pre-test and post-test scores were taken as the main variables, whereas other parameters including completing rates of activity gave contextual performance. The information on the participants was tabularized, allowing matching comparisons to be done correctly. This automatic extraction makes the data integrity intact, promotes the calculation of differences in the scores, and provides support to the further Paired Sample t-test between the engagement and retention outcomes.

2) Data cleaning and preparation: Data cleaning and preparation provided the Teaching-Learning Gamification dataset with the analytical reliability. Missing pre- and post-test scores were filled by mean imputation, z-scores identified outliers, and Min -Max scaled the numerical attributes. This standardized data allowed calculating the difference in scores accurately and validly using paired sample t-test based assessment of vocabulary retention.

3) Calculation of score differences: Score differences are an important component of the quantitative analysis, which would be a direct indication of individual learning gain with the help of gamified AI-based English vocabulary system. With the Teaching-Learning Gamification dataset of Mendeley Data, pre-test scores are the baseline knowledge of the learners, and post-test scores are post-intervention performance. The difference between post and pre-test score is used to compute individual learning gain. This measure is a true measure of the efficiency of the gamified system to increase the vocabulary retention among every one of the participants, which would be a valid basis to later conduct a statistical analysis. The difference between the scores of all learners may be developed mathematically in Eq. (6):

$$d_i = \text{Post}_i - \text{Pre}_i \tag{6}$$

where, d_i is the difference of the i-th learner, Post_i indicates post-test score and Pre_i indicates pre-test score. Appendage of positive values of di means an improvement, whereas zero means no change and negative values mean a decline in performance, but these cases are not very common in a well-designed intervention of gamified learning.

Once d_i is determined in all the learners, the differences will be summed up to obtain the mean difference $\bar{d}$ and standard deviation of differences s_d which will be required in the Paired Sample t-test. Mean difference is determined by Eq. (7):

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} \tag{7}$$

where, n= the total number of learners. Equally, the standard deviation of differences is given as Eq. (8):

$$s_d = \sqrt{\frac{\sum_{i=1}^{n}(d_i - \bar{d})^2}{n-1}} \tag{8}$$

where, s_d is the standard deviation of the differentiation in the scores $\bar{d}$ is the difference between the post-test and the pre-test scores of the i th learner, d is the average of all the differentiation and n is the number of learners. This measure indicates consistency of retention results among the participants. These computations determine those gains on an individual level and system efficiency, match the data to test them in a valid statistic, and also offers an opportunity to correlate cognitive changes with motivational results of the qualitative analysis.

4) Paired sample t-test computation: The paired sample t-test is used to find out that the gamified AI-based vocabulary learning system can provide a statistically significant improvement in the retention of learners. Paired observations are used when pre-test and post-test scores are concerned with the same participants. The differences in individual scores are calculated, and the mean difference and standard deviation are calculated, on which the t-statistic will be determined and the effectiveness of the intervention will be assessed. It is measured in Eq. (9).

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \tag{9}$$

where, n denotes total number of learners. The t-value measures the distance between the observed difference in the means and the distance between the variability in the differences. The higher the t-value the more the gamified intervention affects vocabulary retention. The calculated t-statistic is compared with the critical value of t-distribution with n- 1 degrees of freedom or a p-value is obtained. The post-test scores improvement is said to be statistically significant in case the p-value is below the significance threshold of 0.05.

This process will make sure that the analysis considers the paired nature of the data and that it controls individual differences and isolates the impact of the gamified AI system. The Paired Sample t-test is used to determine the score changes that are associated with statistical significance and thus constitute solid empirical data that the study relies to make its conclusion about retention as the quantitative foundation of the research.

5) Significance testing and interpretation: The significance testing used in this paper will test whether the improvements in vocabulary retention in learners after using the gamified AI-based English learning system are statistically relevant. After calculating the t- statistic of the Paired Sample t-test, the p-value is obtained. The p-value shows how likely it is to get the calculated difference in the means or a more severe statistical outcome under the null hypothesis no improvement exists (H0 1 =0). Mathematically, the null hypothesis is tested as in Eq. (10).

$$H_0{:}\bar{d} = 0, \quad H_1{:}\bar{d} \neq 0 \tag{10}$$

The hypotheses are defined as $H_0{:}\bar{d} = 0$ and $H_1{:}\bar{d} \neq 0$, where $\bar{d}$ represents the mean score difference. If the computed p-value is less than $\alpha = 0.05$, H_0 is rejected, confirming that the gamified AI-based intervention produces a statistically significant improvement in vocabulary retention.

The statistical interpretation presented above proves that the vocabulary improvement achieved has been observed to be credible, meaningful, and associated to the gamified AI-based learning method.

6) Reporting results: The results in this study will be reported by providing the findings of the quantitative analysis to indicate the efficacy of the gamified AI-based English vocabulary learning system. The descriptive statistics of the mean pre-test score, mean post-test score, the mean difference, and standard deviation of differences (di =Posti) are summarized after calculating the score differences (di =Posti) and carrying out the Paired Sample t-test. The measures of effect size, like Cohens d may be used to determine the extent of improvement. It is illustrated in Eq. (11).

$$d = \frac{\bar{d}}{s_d} \tag{11}$$

All these statistics can be easily interpreted to understand the gains in retention and both the statistical and practical significance are clear and clear. Reporting individual differences, mean improvements and significant values, the study is well-evidenced to conclude that the gamified AI system has a positive effect on vocabulary learning, both in numerical terms and analytical rigor.

G. Qualitative Analysis: Motivation Measurement

The qualitative analysis measures the motivation of learners through emotional tone in the feedback about the gamified AI-based vocabulary system. Systematic sentiment analysis classifies responses to be positive, neutral or negative. This method is a complement to quantitative findings that will provide better enlightenment on the interaction and its influence on the general effectiveness of learning.

1) Data selection: Appropriate textual feedbacks of the Teaching-Learning Gamification dataset are mined to evaluate motivation. Open ended questions that represent participation, fun, difficulty, and impressions of learning are chosen. This means that only meaningful entries of emotional and behavioral reactions will be analyzed, which will give a strong basis of sentiment assessment and relate motivational patterns to vocabulary retention results.

2) Text cleaning and preprocessing: Symbols and digits, punctuation and unwanted characters are cleared off the feed back and the same is converted into lower case. The words are made to be standardised at their root forms by tokenization, removal of stopwords and lemmatization. This guarantees consistency, less noise, and textual data readiness to correctly score sentiments in order to have effective quantification of learner motivation.

3) Sentiment lexicon assignment: Cleaned learner feedback is analyzed using the VADER sentiment model to assign polarity scores ranging from -1 to +1. These numerical scores quantify emotional engagement and classify responses as positive, neutral, or negative, providing a systematic measurement of motivation levels during interaction with the gamified AI-based vocabulary system.

4) Sentiment categorization: Calculated sentiment scores by the learner feedback are classified into meaningful motivation levels in this study to assess interaction with the gamified AI-based vocabulary learning system. The VADER model scores compounds on a scale of -1 to +1, where scores above 0.05 are positive motivation, -0.05 to 0.05 is neutral and the scores below 0.05 show negative motivation. This categorization enables a generalization of motivation patterns on all learners indicating proportions of high, moderate, or low engagement. The analysis transforms the numerical sentiment data into qualitative categories, which prove the definite correlation between emotions reactions and the engagement of the learners, and supplement the quantitative results on retention to complete the evaluation of the learning process.

5) Aggregation and interpretation: The process of aggregation and interpretation in this study will be based on the synthesis of single scores in sentiment measures that can be used to measure the general level of motivation among learners during English vocabulary learning through gamification and AI. Once feedback has been divided into positive, neutral, and negative, the scores are summed to give the average level of motivational on all participants. The percentage of every type is calculated in Eq. (12).

$$P_{\text{category}} = \frac{N_{\text{category}}}{N_{\text{total}}} \times 100 \tag{12}$$

Where, N_{category} is the count of a specific sentiment, and N_{total} is the total feedback entries. This measures motivational distribution, identifies engagement trends, and links emotional responses to learning outcomes, complementing quantitative retention analysis.

H. Integration of Retention and Motivation Outcomes

It offers an in-depth insight into the effect of the gamified AI-based vocabulary system on the learning performance and on the engagement of learners. The analysis makes use of statistically confirmed retention differences in the paired t-test and the motivation scores (based on a sentiment) to determine whether a higher level of motivation relates to a better level of vocabulary retention. This summary reinforces the pedagogical meaning of findings, and emphasizes the cognitive and affective synergistic effects of AI-based gamified learning conditions. The investigation of this association is achieved by the aid of a simple correlation measure in Eq. (13).

$$r = \frac{\sum(M_i-\bar{M})(R_i-\bar{R})}{\sqrt{\sum(M_i-\bar{M})^2 \sum(R_i-\bar{R})^2}} \tag{13}$$

where, the correlation coefficient r measures the strength and direction of the relationship between learner motivation scores (M_i) and vocabulary retention improvements (R_i). It reveals whether higher emotional engagement is associated with greater learning gains in the gamified AI-based vocabulary system.

I. Ethical Considerations

In this study, ethical issues are considered so that all research processes observe privacy, transparency, and responsible data usage. The data used is completely anonymized, which does not allow identifying a participant and protects personal data during the analysis process. Every step is taken in terms of the ethical research practice, such as the safe processing of quantitative and qualitative data. The research is ethical because it has not manipulated data, and it is also honest in reporting the results and making the method of analysis reproducible. Also, the analysis of sentiments and the statistical tests are performed objectively without the introduction of bias and misunderstanding. All these measures will guarantee the ethical rigor and the credibility and trustworthiness of the research findings.

IV. Results and Discussion

The findings section demonstrates the most important facts of the results of the quantitative and qualitative analyses that were performed to measure the efficiency of the gamified AI-based vocabulary learning system. The quantitative results of statistical testing indicate the measurable increase in retention among learners, and the qualitative analysis of feedback can be reduced to the sentiment analysis, which indicates motivational patterns and the level of engagement. The results give a holistic account of the system effects by incorporating cognitive performance and emotional reactions. Tabular summaries and graphical visualizations also explain trends, which serve to interpret the data. On the whole, this part emphasizes the improvement of the vocabulary acquisition and the motivation of the learners through AI-driven gamification in a logical and evidence-based way

A. Descriptive Statistics

Table II displays the summary of core motivation dimensions with the mean of 2.672 to 2.9 meaning moderate engagement levels of Group Motivation, Role Performance, Task Completion, and Learning Interaction constructs. The consistency of the median of 3 indicates that the learner answers are homogenous and it can be used in a reliable reference to understand motivation patterns in the dataset under analysis.

TABLE II. DESCRIPTIVE STATISTICS OF LEARNER PERFORMANCE AND MOTIVATION

Dimension	Mean	SD	Median	Median
Group Motivation	2.9	1.03	3	1–4
Role Performance	2.74	0.94	3	1–4
Task Completion	2.77	0.98	3	1–4
Learning and Interaction	2.67	0.96	3	1–4
Group Integration	2.77	1	3	1–4

TABLE III. FREQUENCY DISTRIBUTION OF GROUP MOTIVATION CATEGORIES

Motivation Category	Score Range	Frequency
Low	1.0 – 1.9	7
Moderate	2.0 – 2.9	18
High	3.0 – 3.9	25
Very High	4.0 – 5.0	10

Table III shows how the learners were distributed in the various categories of motivation, 7 in Low, 18 in Moderate, 25 in High, and 10 in Very High. The differences demonstrate how a minimal fluctuation in the motivation scores can dislocate the learners to the other categories and influence the general analysis and give the insights on the engagement tendencies in the gamified learning environment.

B. Preprocessing Outcomes

Fig. 2 shows the performance of learners prior to and after normalization and it is quite evident how raw scores of 32 to 95 will be changed into scaled scores between 0-1. The Before Scaling line goes up gradually among the learners and this is a reflection of the initial score variations. The pattern is similar to the one before Min-max scaling, except that each value is proportionally scaled e.g. 32 would be represented by 0.00, 45 by 0.18, 70 by 0.61 and 95 by 1.00. This correspondence shows that normalization maintains the relationship between scores and makes the range standardized.

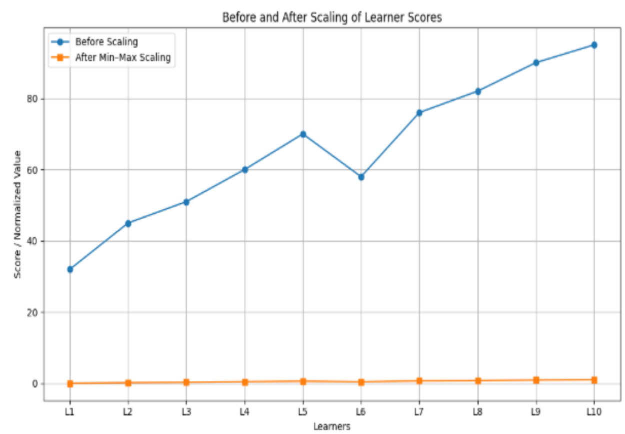

Fig. 2. Before and after scaling of learner scores.

C. Quantitative Analysis: Retention Measurement

Fig. 3 shows bar chart that is used to compare the performance of vocabulary prior to the learning intervention and the performance post the learning intervention. The score in pre-tests is between 40 and 78 including L1 with a score of 40 and L10 with a score of 78. Test scores are also evidently better as L1 shows the scores to be 55 and L10 95 on the post-test. All of the learners have positive gains with gains as high as L3 of 50 to 68 and L7 of 65 to 85. The difference in the visual representation between the paired bars shows uniform learning across all the participants.

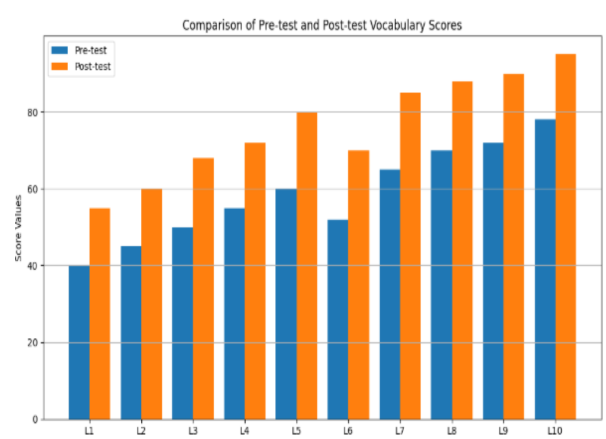

Fig. 3. Comparison of pre-test and post-test vocabulary scores.

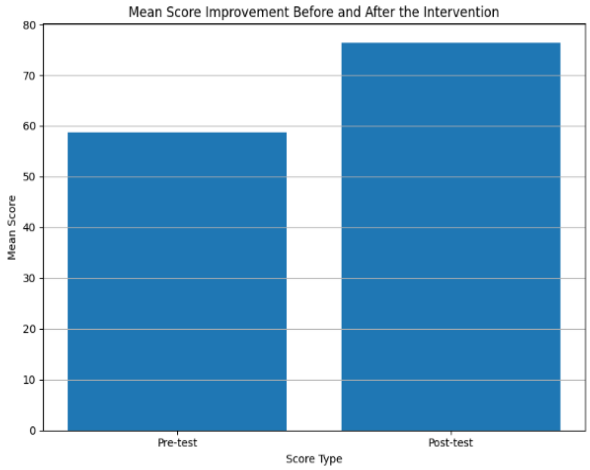

Fig. 4. Mean score improvement before and after the intervention.

Fig. 4 shows bar graph presents the average scores in vocabulary of the two days before and after the learning intervention. The mean of the pre-test is 58.7, which represents the level of performance at the beginning of the learners. The post-test mean also increases significantly after the intervention to 76.3 meaning that the improvement is evident by 17.6 points. This is an apparent growth in vocabulary learning as the post-test bar appears taller than the pre-test bar. The comparison brings out the general performance of the instructional strategy in improving the performance of the learners.

TABLE IV. PAIRED SAMPLE T-TEST RESULTS FOR VOCABULARY RETENTION

Metric	Value
Mean Difference	4.12
Standard Deviation (SD)	2.35
t-value	7.28
p-value	0.00001
Metric	Value

Table IV presents a significant increase in learner performance with a mean of 4.12 and moderate variability (SD = 2.35). Vocabulary gains in retention are statistically significant due to the high t-value (7.28) and very low p-value (0.00001).

D. Qualitative Analysis: Motivation Measurement

Fig. 5 shows the polarity of sentiment scores distribution, where most of the scores are concentrated at the range (0.2 0.6) of strong satisfaction with learners. Eighteen of the answers indicate a very positive motivation, only four negative responses and two neutral answers are received, which proves the generally positive emotional reaction towards the gamified AI learning process.

Fig. 6 represents the distribution of the level of learner sentiment, with the majority of moderate and low responses happening, which implies the presence of neutral and slightly-decreased motivation. The high and very high sentiments are less and this implies a few strong motivational reactions in all.

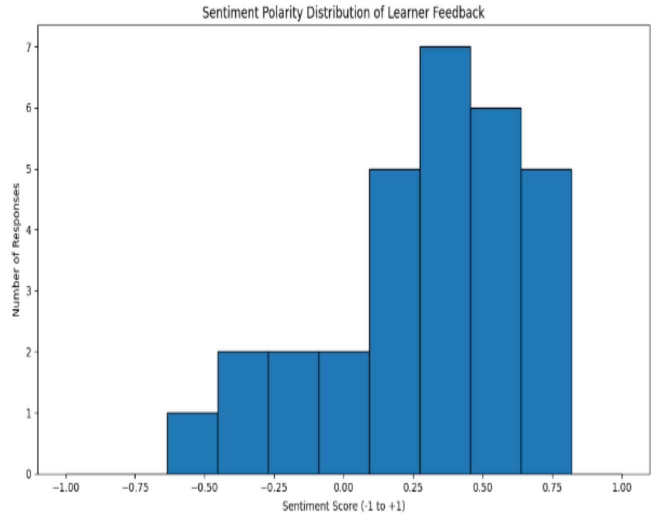

Fig. 5. Sentiment polarity distribution of learner feedback.

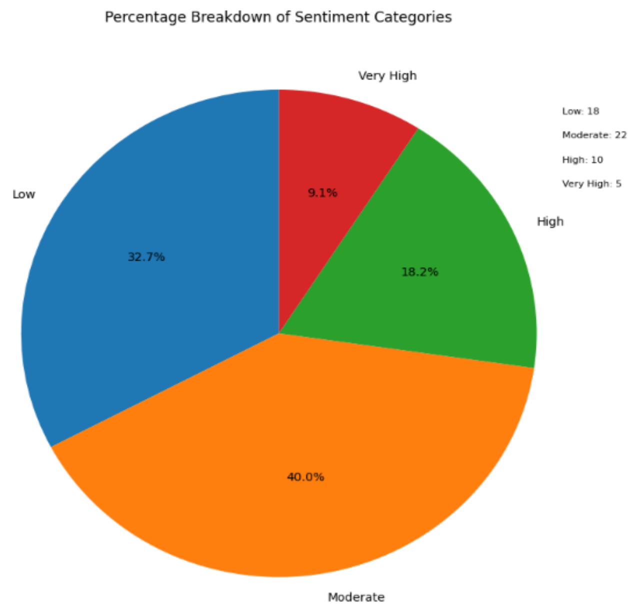

Fig. 6. Percentage breakdown of sentiment categories.

TABLE V. FREQUENCY DISTRIBUTION OF GROUP MOTIVATION CATEGORIES

Sentiment Category	Count
Low	18
Moderate	22
High	10
Very High	5

Table V shows the most common categories of sentiments based on the feedback of learners. High and very high levels of motivation are predominant, and 18 and 22 responses, respectively, show that the level of engagement is high. The moderate responses constitute 10 entries, and the low motivation can only be attributed to 5 entries indicating the lowest level of dissatisfaction and general positive perception of the learner.

E. Integrated Retention: Motivation Analysis

Fig. 7 shows the direct proportions between the motivation scores and retention gains at the paired value of motivation 2.1 retention gain 4, 3.5 retention gain 11 and 4.8 retention gain 18. The trend of the plotted points is to the upward direction meaning that the more the motivation the more the retention increase. All points are marked in order to easily represent the specific numeric pair unmistakably, without overlap, to facilitate easy interpretation of the pattern. This visual data confirms that motivated learners normally gain vocabularies better after the intervention.

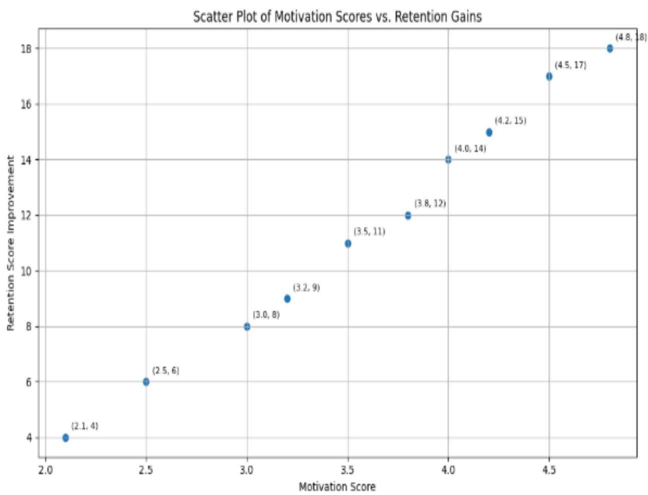

Fig. 7. Scatter plot of motivation scores vs. Retention gains.

TABLE VI. CORRELATION BETWEEN RETENTION IMPROVEMENT AND MOTIVATION SCORE

Metric	Value
Correlation Coefficient (r)	0.62
Significance (p)	0.003

Table VI indicates that the increase in vocabulary retention has a statistically significant positive relationship with the motivation of learners ($r = 0.62$, $p = 0.003$). The presence of this moderately high correlation proves that the greater the degree of motivational engagement is, the better the learning outcomes are, which legitimizes the educational power of combining gamification with AI-enhanced vocabulary training.

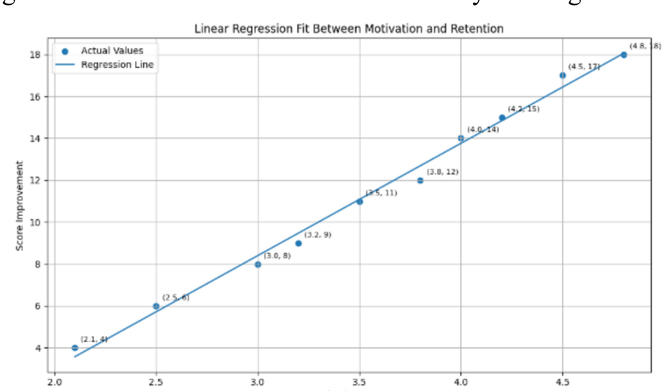

Fig. 8. Linear regression fit between motivation and retention.

Fig. 8 presents the linear association between the motivation scores and retention improvement thus indicating an apparent upward trend. The regression line fitted fits very well with the observed data points implying that there is a strong predictive association. This trend indicates that increased motivation will always result in maximum vocabulary retention in the gamified AI learning environment.

F. Additional Performance and Interaction Metrics

Fig. 9 demonstrates the enhancement of task completion with the improvement of group interaction. As an example, a rating of 1.8, 2.5, 3.2, and 4.0 translates to a 2.0, 2.8, 3.5, and 4.2 level of task completion respectively. All the points are well marked to prevent overlaps to make it easy to interpret the values. The positive trend means that the more successful interaction of learners in their groups, the more successful they would be in completing their tasks, which proves the positive impact of the collaborative engagement on the performance results.

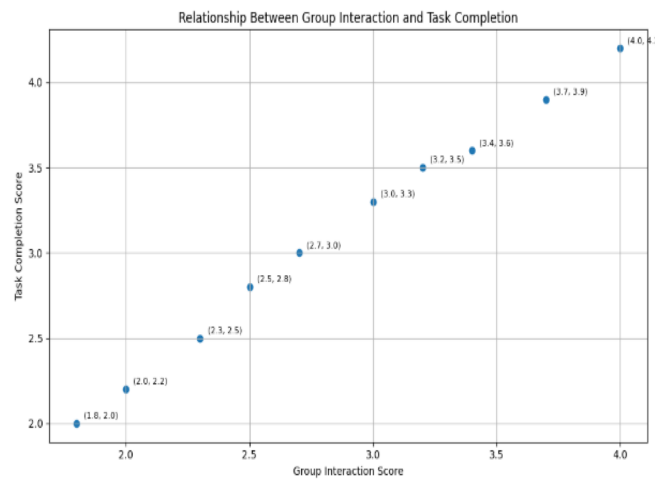

Fig. 9. Relationship between group interaction and task completion.

G. Discussion

The results show that the presented gamified AI-based vocabulary learning tool has a significant positive impact on retention and motivation of learners. The outcomes of the paired sample t-test show that there is a statistically significant positive change in the post-test vocabulary scores which proves that AI-assisted adaptive tasks are effective in ensuring a lasting engagement in cognitive processes. Motivation analysis also indicated that the majority of learners had moderate to high positive feelings, showing their satisfaction with the main aspects of gamification like instant feedback, reward systems, and progress monitoring. The positive relationship existing between motivation and retention has been observed, reiterating the fact that learning environments that are emotionally involving encourage better academic performance. The correlation between these two variables is supported by regression and scatter plot analyses that reveal that there are regular vocabulary gains in highly motivated learners. Moreover, the higher the level of task completion and the team-working interaction patterns, the higher the indication is that AI-based gamification facilitates active engagement and quality learning activities. The research is based on one publicly available data, and it is hard to extrapolate it to other settings. The behaviour of a learner is platform-dependent, resistance to noisy or adversarial examples is not fully assessed, motivation metrics are partially based on self-reports, and the architecture of the AI model is not completely optimized or compared with various baselines.

V. Conclusion and Future Work

The research investigated the efficiency of a gamified learning system of artificial intelligence-based learning of English vocabulary, by analyzing vocabulary retention and learner motivation together. The quantitative findings indicated the statistically significant enhancement of vocabulary performance post-intervention, which proves the usefulness of applying adaptive AI guidance along with structured gamification to facilitate cognitive growth. The qualitative results were complementary, and most of the motivation responses were positive, suggesting that interactive activities, timely feedback, and game-based mechanics increased the engagement. The combined analysis also proved a effective correlation between the motivation and retention because highly motivated learners gained higher vocabulary. Altogether, the results can serve as solid proof that AI-mediated gamification positively influences both the cognitive and the affective learning of language.

A. Theoretical and Practical Implications

The results of the proposed research contribute to the theoretical knowledge base of AI-driven education by showing that the highest learning results are achieved when cognitive retention and affective motivation are considered as mutually reliant constructs as opposed to individual indicators. Findings do confirm the learner-focused and self-determination-focused viewpoints, with emphasis placed on the significance of adaptive feedback, autonomy, and emotional involvement in vocabulary acquisition in the long-term. Practically speaking, the GAI-VRME framework proposed can provide a blueprint of scalable application of adaptive gamification, which is based on real-time model learner. The sentiment-based motivation analysis can be integrated to have practical implications to instructional design and personalization, and these implications are not only in vocabulary learning, but also in more general intelligent tutoring and adaptive learning systems.

Even though the findings affirm the efficacy of AI-facilitated gamified vocabulary learning, the findings also point to the future research directions. The validity of the external validity can be essentially enhanced by conducting future research in terms of larger and more heterogeneous populations of learners in different educational settings. Other behavioral measurements, including time-on-task, the frequency of interaction, and adaptive challenge responses can provide more information on the engagement patterns. The use of sentiment analysis with the use of transformers would provide more emotional insights. The longitudinal studies should be encouraged to evaluate the sustainability of retention, and personalization by use of real time adaptive feedback and reinforcement learning can also be applied to improve the results of learning.

References

[1] A. Christopoulos and S. Mystakidis, "Gamification in education," Encyclopedia, vol. 3, no. 4, pp. 1223–1243, 2023.

[2] S. Suresh Babu and A. Dhakshina Moorthy, "Application of artificial intelligence in adaptation of gamification in education: A literature review," Comput. Appl. Eng. Educ., vol. 32, no. 1, p. e22683, 2024.

[3] N. Chandrakant, "Gamified learning and NLP: Enhancing student engagement through AI-driven interactive education models," Int. J. Sci. Res. Arch., vol. 9, no. 1, pp. 813–824, 2023.

[4] B. Panmei and B. Waluyo, "The pedagogical use of gamification in English vocabulary training and learning in higher education," Educ. Sci., vol. 13, no. 1, p. 24, 2022.

[5] N. P. Wulantari, A. Rachman, M. N. Sari, L. J. Uktolseja, and A. Rofi'i, "The role of gamification in English language teaching: A literature," J. Educ., vol. 6, no. 01, pp. 2847–2856, 2023.

[6] W. Chen, "Gamification as a Strategy for Enhancing Long-Term Memory of Low-Frequency Vocabulary in Primary English Education," Res. Adv. Educ., vol. 4, no. 4, pp. 51–58, 2025.

[7] H. Gadra, A. Guebli, and A. Boumediene, "Exploring the Effect of Gamified Learning on the Acquisition of Vocabulary among EFL Learners: Case of Fourth Year Pupils at El Mujahid Kadoun Mohamed Primary School-Ain Temouchent.," PhD Thesis, UNIVERSITY OF AIN TEMOUCHENT, 2025.

[8] X. Ren, "Gamified systems: exploring state-of-the-art design beyond points and badges," 2025.

[9] A.-S. Hellberg and J. Moll, "A point with pointsification? Clarifying and separating pointsification from gamification in education," in Frontiers in education, Frontiers Media SA, 2023, p. 1212994.

[10] S. Bennani, A. Maalel, and H. Ben Ghezala, "Adaptive gamification in E-learning: A literature review and future challenges," Comput. Appl. Eng. Educ., vol. 30, no. 2, pp. 628–642, 2022.

[11] A. H. Al-Hoorie and O. Albijadi, "The motivation of uncertainty: Gamifying vocabulary learning," RELC J., vol. 56, no. 2, pp. 332–345, 2025.

[12] X. Chen, W. Sun, and R. Zhang, "An Interval-Valued Neutrosophic Framework: Improved VIKOR with a Preference-Aware AHP–Entropy Weight Method for Evaluating Scalp-Detection Algorithms," Appl. Sci., vol. 15, no. 22, p. 11937, 2025.

[13] C. Lu and M. Cutumisu, "Online engagement and performance on formative assessments mediate the relationship between attendance and course performance," Int. J. Educ. Technol. High. Educ., vol. 19, no. 1, p. 2, 2022.

[14] L. Liu, "Impact of AI gamification on EFL learning outcomes and nonlinear dynamic motivation: Comparing adaptive learning paths, conversational agents, and storytelling," Educ. Inf. Technol., vol. 30, no. 8, pp. 11299–11338, 2025.

[15] F. Naseer, M. N. Khan, A. Addas, Q. Awais, and N. Ayub, "Game Mechanics and Artificial Intelligence Personalization: A Framework for Adaptive Learning Systems.," Educ. Sci., vol. 15, no. 3, 2025.

[16] I. Patra, N. Shanmugam, S. M. Ismail, and G. Mandal, "An Investigation of EFL Learners' Vocabulary Retention and Recall in a Technology-Based Instructional Environment: Focusing on Digital Games," Educ. Res. Int., vol. 2022, no. 1, p. 7435477, 2022.

[17] D. Liu, "The effects of segmentation on cognitive load, vocabulary learning and retention, and reading comprehension in a multimedia learning environment," BMC Psychol., vol. 12, no. 1, p. 4, 2024.

[18] Z. Zhang and J. Crawford, "EFL learners' motivation in a gamified formative assessment: The case of Quizizz," Educ. Inf. Technol., vol. 29, no. 5, pp. 6217–6239, 2024.

[19] Z. Zhang and X. Huang, "Exploring the impact of the adaptive gamified assessment on learners in blended learning," Educ. Inf. Technol., vol. 29, no. 16, pp. 21869–21889, 2024.

[20] L. Li, K. F. Hew, and J. Du, "Gamification enhances student intrinsic motivation, perceptions of autonomy and relatedness, but minimal impact on competency: a meta-analysis and systematic review," Educ. Technol. Res. Dev., vol. 72, no. 2, pp. 765–796, 2024.

[21] K. Sadeghi, E. Sağlık, E. Mede, Y. Samur, and Z. Comert, "The effects of implementing gamified instruction on vocabulary gain and motivation among language learners," Heliyon, vol. 8, no. 11, 2022.

[22] Putri Fachri Aulia Fatah, "The Effectiveness of Gamification in Enhancing English Language Learning Outcomes," J. Pendidik. Dan Sastra Ingg., vol. 5, no. 2, pp. 304–320, Jun. 2025, doi: 10.55606/jupensi.v5i2.5238.

[23] P. K. Kumar and C. Vairavan, "The Impact of Gamification on Motivation and Retention in Language Learning: An Experimental Study Using a Gamified Language Learning Application," INTI J., vol. 2024, 2024.

[24] Y. Chen and S. Zhao, "Understanding Chinese EFL learners' acceptance of gamified vocabulary learning apps: an integration of self-determination theory and technology acceptance model," Sustainability, vol. 14, no. 18, p. 11288, 2022.

[25] Z. Yu, "Learning outcomes, motivation, and satisfaction in gamified English vocabulary learning," Sage Open, vol. 13, no. 2, p. 21582440231158332, 2023.

[26] L. Liu, "Impact of AI gamification on EFL learning outcomes and nonlinear dynamic motivation: Comparing adaptive learning paths, conversational agents, and storytelling," Educ. Inf. Technol., vol. 30, no. 8, pp. 11299–11338, 2025.

[27] G. Caiza, C. Villafuerte, and A. Guanuche, "Interactive Application with Virtual Reality and Artificial Intelligence for Improving Pronunciation in English Learning," Appl. Sci., vol. 15, no. 17, p. 9270, 2025.

[28] G. L. Liu, R. Darvin, and C. Ma, "Unpacking the role of motivation and enjoyment in AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation in the Chinese context," Comput. Hum. Behav., vol. 160, p. 108362, 2024.

[29] P. Acosta-Vargas, C. Uchima-Marin, and L. Salvador-Acosta, "Dataset on Teaching-Learning Gamification." Mendeley Data, 2025. doi: 10.17632/YC4NP572ZS.