JADHUR

Vol.3, Issue.4, (2024)

Journal of Advances in Humanities Research

https://doi.org/10.56868/jadhur.v3i4.234

Mobile Gamification for Early Literacy: An Analysis of Learning Outcomes and Engagement

Li Liu[1], Jing Zhang[2*], Ziyan Guo[3] and Yanyu Wang[3]

1. Guizhou Lifelong Education Research Institute, Guiyang Preschool Education College, P.R, China
2. INTI International University: Kampung Baharu Nilai, Malaysia
3. School of Art, Guiyang Preschool Education College, P.R, China

Article Information

Article Type: Research Article

Dates:
Received: 12 July 2024
Revised: 10 December 2024
Accepted: 18 December 2024
Available online: 21 December 2024

Corresponding Author: Jing Zhang

1809030129@stu.hrbust.edu.cn

ORCID: https://orcid.org/0009-0009-8323-5457

ABSTRACT

This study aimed to evaluate the impact of gamified apps on phonemic awareness, vocabulary, and reading comprehension. The study aims to explore the relationship between engagement levels and literacy outcomes. A quasi-experimental design involving pretest and post-test assessments of 200 children aged 5 to 7 divided equally into experimental and control groups. The experimental group used gamified learning apps, while the control group engaged in traditional learning activities. Data were analyzed using paired t-tests, independent sample t-tests, ANCOVA, and multiple regression analyses. The findings revealed that the experimental group showed significant improvements in all literacy measures compared to the control group. Higher engagement with the gamified apps was positively associated with more significant literacy gains. The study concludes that gamified mobile learning apps enhance early childhood literacy by increasing engagement and motivation. These results have important implications for educators and policymakers, suggesting that integrating gamified elements into early childhood education can significantly improve literacy outcomes. Educational technology developers are encouraged to design engaging and motivational learning tools to maximize educational benefits.

Keywords: Gamification, Early Childhood Literacy, Vocabulary Development, Reading Comprehension, Engagement

1. INTRODUCTION

The integration of gamification into mobile learning applications has become a radical revolution in initial literacy learning. The term "Gamification" refers to using game mechanics in areas unrelated to games, aiming to harness some of the motivational aspects promoted in game playing to learning environments (Zainuddin et al., 2024). This applies primarily to education at the initial stages of a child's learning when it is imperative, for instance, to sustain the child's attention and interest (Mandujano et al., 2023). Smartphones and tablet-based applications allow students to have a mobile form of learning whenever they articulate the courses preferred by more and more educational institutions (Huang & Zhou, 2021). Studies have shown that educational gamification solutions can enhance learners' motivation, engagement, and persistence (Koivisto & Hamari, 2019). Specific to gamification, apps in the domain of

early literacy involve points, badges, quest maps, and leaderboards as crucial components that form a good learning process (De-Marcos et al., 2014). The problems associated with early learners' classes, including short attention spans and inequality of interests, can be dealt with using gamification to enhance learning. Correspondingly, the use of mobile apps and their interactivity and integration of multimedia in games make it a versatile tool in early literacy learning that suits the different needs and requisites of the learners (Kapp, 2012). For instance, interactive storytelling programs and phonics games facilitate the learning of phonemic awareness and vocabulary (AlAwadhi and Al-Daihani, 2019). Given this flexibility, the following are key advantages: One can effectively meet the needs of the learners, especially those with special needs. It is possible to introduce various activities within the framework of the gamified apps, allowing working with the children's different learning rates and thus avoiding leaving some of them behind in terms of literacy development (Oliveira et al., 2023; Xu & Hamari, 2023). Among the primary benefits of using gamified learning is that it is the surest way of bringing about a sense of achievement and progress among the learners. If those children get the points or badges for the tasks they perform or the new skills attained, they stand to gain something tangible, which will be instrumental in bolstering their self-esteem and increasing their desire to learn (Nicholson, 2015).

In addition, introducing technologies in early childhood education has merits and demerits. Mobile learning apps may offer learners efficient and practical learning tools; however, there are issues related to learners' limited time in front of the screens at different ages, including early childhood (Oliveira et al., 2023; Cheong and Cheong, 2014). It is through screen time that various adverse effects, such as decreased physical activity, poor sleep quality, and behavioral problems, are associated (Hamari et al., 2014). Thus, parents and teachers must control and restrain the usage of such applications by focusing on encompassing these apps in learning processes (Rosli & Omar Zaki, 2023; Hanus & Fox, 2015).

The primary objective of this study is to evaluate the effectiveness of gamified mobile applications in enhancing early literacy skills, specifically focusing on phonemic awareness, vocabulary acquisition, and reading comprehension among children aged 5 to 7. By employing a quasi-experimental design, the research seeks to compare the learning outcomes of children using gamified apps with those engaged in traditional literacy activities, providing empirical evidence on the efficacy of gamification in early education. Additionally, the study aims to explore the relationship between engagement levels and literacy gains, addressing a critical gap in the understanding of how interactive and motivational elements of gamification contribute to improved educational outcomes. Moreover, investigating these objectives, the study seeks to contribute valuable insights into the pedagogical potential of gamified learning tools, offering practical implications for educators, policymakers, and technology developers in early childhood education.

2. LITERATURE REVIEW

Gamification, defined as applying game design elements in non-game contexts, has emerged as a transformative pedagogical strategy in education (Seaborn & Fels, 2015). Gamification leverages points, badges, leaderboards, and immediate feedback to foster engagement and sustain learner motivation (Hanus & Fox, 2015). In early childhood education, gamification also relates to a play-based paradigm in which the tasks and activities are learner-centered, fun, and purposeful, contributing to cognitive and socio-emotional development. Mobile applications are one of the tools that refer to digital gamified solutions as they can be easily accessed, involving and flexible (Hammedi et al., 2024). These apps are composed of smart learning functions and intelligent and adaptive feedback, which are personalized for the learner's

needs, making them particularly helpful for young and novice learners in developing foundational literacy functions (Hammedi et al., 2024).

2.1 Positive Impacts of Gamification on Students

Incorporating gamification into the classroom has shown many positive results for learners at different graded levels and fields of study. The first issue solved is that it increases engagement and motivation. The study of Jaskari and Syrjala (2023), which employed Self-Determination Theory (SDT), revealed that gamification fulfills the psychological need for autonomy, competence, and relatedness, enhancing engagement in learning activities. These are important in early childhood literacy because the motivational affordances keep children engaged and motivated to tackle complex elements of emergent literacy, such as word recognition and reading comprehension (Gee, 2003). The strengths of gamifying tools in increasing academic performance are reinforced in empirical studies. To illustrate, (Hamari et al., 2014) played a gamified learning environment that achieved more knowledge retention and task completion than in regular classrooms. Real-time feedback and rewards create a growth mindset, as students have features like that for real-time feedback and rewards to track their progress and know where they are lacking. It is excellent for young learners who can learn just by being rewarded immediately. Furthermore, gamifying tools allow learners to be included at different paces and learning styles (Hanus & Fox, 2015).

2.2 Negative Impacts of Gamification on Students

Although Gamification has numerous advantages, it faced criticism among researchers, and significant concern is that the reliance on extrinsic rewards may undermine intrinsic motivation, a phenomenon referred to as the "overjustification effect" (Zvereva et al., 2023). If learners begin to rely excessively on external rewards like points or badges, the long-term engagement of learners within the space can grow when they are present (Kashive & Mohite, 2023). In particular, the staying power one needs to acquire lifelong reading habits is closer to the bone in early literacy education, where permitting interest to lag once outside gamified contexts is pressing (Nacke & Deterding, 2017). In addition, there is a potential for cognitive overload if students are exposed to too many gamified applications with a complicated or defective interface. According to Cooper's (1990) Cognitive Load Theory, this overload could hinder learning by consuming cognitive resources that should be used during meaningful engagement with the material. Despite studies reporting that poorly implemented gamification can lead to distractions that impede the overall success of instructional activities (Mazarakis & Bräuer, 2023), gamification is still prevalent.

Gamification has also raised concerns about the possibility that competition between students may be heightened by gamification, making many students feel inferior or nervous if they cannot get high scores or accomplish game objectives. Negative emotions can reduce participation and erode self-esteem, erecting risks to the young learners' social-emotional development (Goldman, 2008). Additionally, gamification may be overdone, but measurable outcomes should be emphasized over more profound learning objectives. This reductionist approach has been criticized by critics, who argue that this will trivialize education and leave out how education and the holistic development of students are neglected (Zvereva et al., 2023).

2.3 Access to Gamification and Mobile Learning

Gamified mobile learning tools are widely uneven in access, and the barriers are socioeconomically- and culturally based (Nugroho, 2024). However, promising though gamification may

be for democratizing education, its success depends on access to digital resources, a crucial gap in very low-income and disenfranchised communities (Lopes et al., 2024). Gamified learning apps function better where parents have smartphones and tablets, allowing their children to utilize them throughout learning.

Where parents come from disadvantaged backgrounds, it might be difficult for the child to access these types of apps while learning, thus widening the digital divide (Hamari et al., 2014). Collectivist cultures may require adaptations to gamified designs emphasizing collaboration and shared achievements. Gamification design with cultural alignment achieves better results. Furthermore, gamified tools can be shaped by gender norms and stereotypes related to access and engagement. Based on research, boys are usually more attracted to competitive and action-oriented games, girls may prefer collaborative and story-related experiences, and there are diverse design strategies to make experiences inclusive (Liatukaitė, 2024). Additionally, learners in non-native language environments face hurdles due to language barriers (Lopes et al., 2024). Because many gamified apps are developed in English, they are unavailable to children in English-speaking regions unless a localized version is developed.

3. METHODOLOGY

3.1 Research Design

This study employed a quantitative research design to investigate the role of gamification in mobile learning apps on early childhood literacy development in China. A quasi-experimental design was utilized, incorporating both pretest and post-test measures to assess the impact of gamified learning applications. The independent variable in this study was the use of gamified mobile learning applications, while the dependent variables included various literacy outcomes such as phonemic awareness, vocabulary, and reading comprehension. The study involved two groups: an experimental group that used gamified learning apps and a control group that used traditional, non-gamified learning materials. This design facilitated a direct comparison between the two groups, thereby allowing for a robust evaluation of the effectiveness of gamification in enhancing early literacy skills. The experimental group interacted with mobile applications to engage children through game-like elements such as points, badges, and interactive narratives. In contrast, the control group continued conventional literacy activities, including story reading and phonics exercises, without including gamified elements.

3.2 Sampling and Sample Size

The study employed a stratified random sampling technique to ensure that the sample was representative of the diverse population of young learners in Nanjing, China. Schools from various socio-economic backgrounds across urban and rural areas were included to capture a broad spectrum of participants. This approach was crucial for obtaining a sample accurately reflecting the country's demographic diversity. The target sample size for the study was set at 200 children, with 100 allocated to the experimental group and 100 to the control group. This sample size was determined based on a power analysis conducted before the study, which indicated that 200 participants would provide sufficient statistical power to detect significant differences between the groups. The power analysis considered expected effect sizes, the standard deviation of outcomes, and the desired significance level to ensure the study was adequately powered to test the research hypotheses. To make both groups as homogenous as possible, the criteria used for stratification were age, gender, and socio-economic status. A more refined target population was children aged 5-7 years, boys and girls, from different socio-economic backgrounds.

This stratification was done in a way that separated the schools by different socio-economic statuses; they used household income, parents' education level, and the availability of educational resources. The schools were selected randomly in each bracket, and the children were randomly posted in either the experimental or control groups in each school.

3.3 Data Collection Procedure

Data collection in this study occurred in three phases: pretest, intervention, and post-test. Each phase was meticulously planned and executed to ensure the accuracy and reliability of the collected data, as well as the overall integrity of the study. In the pretest phase, baseline literacy skills of all participants were assessed using standardized literacy tests. These tests included the Phonemic Awareness Literacy Screening (PALS) and the Peabody Picture Vocabulary Test (PPVT) to evaluate phonemic awareness and vocabulary. The Early Grade Reading Assessment (EGRA) was also used to measure reading comprehension.

As for the intervention phase of the study, the experimental group used gamified mobile learning apps within a fixed amount of time, usually 30 minutes in real time. These were conducted parallel to the regular school program, meaning the child had a fixed and uninterrupted day to use the apps. The control group, however, had printed worksheets and storybooks for the same period instead of commonly used IT aids during learning. Cross-group equivalence was also ensured, and the teachers and parents of the learners were trained vigorously on how to implement and oversee the intervention. This training focused on using apps in the form of games, procedures for handling conventional physical materials, management of daily records, and other perceived behaviors. Logs were kept and documented in detail to document all the participants' daily activities during the intervention's implementation. In the case of the experimental group, usage data of the mobile apps were gathered using logs that contained information such as time spent on the apps, levels, and achievements.

In the post-test, the same standardized literacy tests were given at the pretest to determine any changes in literacy. This ensured that whichever differences were observed in performance could result from the intervention, not differences in the assessment instruments. Like in the pretests, post-tests were taken under similar circumstances, and the findings were documented in the slightest detail. Also, data collection in the form of self-administered questionnaires with questions created by the researcher was provided to the teachers and parents to establish their perception of the benefit accrued from the intervention. More importantly, the questionnaire was in Chinese, and after collecting the data, it was converted into English.

3.4 Reliability Analysis

The reliability of the engagement and motivation scales was assessed using Cronbach's alpha, a measure of internal consistency. Cronbach's alpha values above 0.70 were considered acceptable, indicating that the items within each scale consistently measured the same construct. The results from the pilot study indicated that the engagement and motivation scales had Cronbach's alpha values ranging from 0.78 to 0.85, demonstrating good reliability.

Table 1: Reliability Analysis of Engagement and Motivation Scales

Scale	Number of Items	Cronbach's Alpha
Engagement Scale	10	0.82
Motivation Scale	12	0.78
Overall	22	0.85

Table 1 indicates that Cronbach's alpha was performed in the reliability analysis of the engagement and motivation scale items to establish internal consistency. Regarding the validity part of the assessment, the engagement scale, which included ten items, had a Cronbach's alpha coefficient of 0. 82, which again portrays a high degree of internal consistency, and this means that all the items used in the study measure a common construct known as engagement. The motivation scale, which has 12 items, has a Cron Bach Alpha of 0.78, again showing good reliability and establishing that all the items consistently gauge the construct motivation. If combined, the overall reliability for the 22 items was 0 for the pilot and primary samples, which is 0.85, confirming high internal consistency across the engagement and motivation indicators.

Table 2: Feedback from Cognitive Interviews

Questionnaire Item	Feedback from Children	Revision Made
Enjoyment Item 1	"It's fun" vs "I have fun using the app"	Simplified wording to "I have fun using it."
Perceived Difficulty Item 3	"This part is hard to understand."	Added examples to clarify the question
Frequency of Use Item 2	"I don't know what 'frequency' means"	Reworded to "How often do you use the app?"

Table 2 shows the problems that could be encountered in the main study concerning the research instruments and procedures. The modifications based on the pilot study's findings made it possible to include the standardized tests and the developed questionnaires as feasible, valid and credible instruments for the target population. Thus, this preparatory phase helped considerably enhance the methodological reliability of the main study. The pilot study results supported how the standardized tests and the engagement and motivation questionnaires would map onto early literacy and the related concept in children of a young age. The reliability coefficients in the current study were relatively high, suggesting that the scales captured the intended constructs used in the main study.

3.5 Scales or Measurements

The study used standardized and custom-designed instruments to comprehensively measure literacy outcomes and other relevant variables. The selection of these instruments was guided by their established validity and reliability in assessing early childhood literacy skills and their appropriateness for the cultural and linguistic context of China. Phonemic awareness literacy screening (PALS) was utilized to assess phonemic awareness. PALS is a widely recognized tool that evaluates young children's understanding of the sound structure of language, which is a crucial component of early literacy development (Invernizzi et al., 2004). This assessment included tasks such as rhyme awareness, initial sound identification, and phoneme segmentation, providing a detailed picture of each child's phonemic awareness abilities. Vocabulary development was measured using the Peabody Picture Vocabulary Test

(PPVT), which assesses receptive vocabulary knowledge by requiring children to select pictures representing spoken.

The PPVT was chosen for its ease of administration and ability to provide a standardized measure of vocabulary that is comparable across different populations. Reading comprehension was evaluated using the Early Grade Reading Assessment (EGRA). EGRA is designed to measure a range of foundational reading skills, including letter recognition, familiar word reading, and reading comprehension through simple passages followed by comprehension questions. This instrument is suitable for early readers and has been adapted in various international contexts, ensuring its relevance to the study population. In addition to these standardized tests, custom-designed questionnaires were developed to assess engagement and motivation. These questionnaires included enjoyment, perceived difficulty, and Frequency of use of the gamified mobile learning apps. The engagement scale was based on established user engagement and flow theories, incorporating items that measured children's immersion, interest, and sustained attention while using the apps (Hamari et al., 2014).

Apart from those usual tests, questionnaires constructed and designed for the purpose were used to measure engagement and motivation. These questionnaires focused on enjoyment, perceived difficulty, and the Frequency of use of the mobile learning apps that incorporated gamification. The engagement scale was developed using previously proposed theories about user engagement and flow. It included items that assessed the level of children's immersion interest and time spent focused on the apps (Hamari et al., 2014).

The socio-economic status (SES) classification into low, middle, and high categories was based on a composite index constructed using household income, parental education level, and parental occupation, as these are widely recognized indicators of SES. Household income was categorized into tertiles based on the income distribution of the study population, with the lowest third classified as low SES, the middle third as middle SES, and the highest third as high SES. Parental education level was assessed by the highest level of education completed by either parent, with categories aligned to primary education or below (low), secondary education (middle), and tertiary education (high). Parental occupation was classified based on occupational prestige scores, distinguishing unskilled or manual labour (low), skilled or administrative work (middle), and professional or managerial roles (high).

3.6 Data Analysis Techniques

The analysis was done in several steps to determine the factors that impacted students' learning experiences in different higher-learning institutions. First, to obtain the mean, count, and percentage of the participant's demographic characteristics and initial literacy levels. This involved the computation of means, standard deviations, and frequencies on age, gender, socio-economic status and PALS, PPVT and EGRA score baselines. It is noted that several percentages and means are presented in the current tables, which allows the painting of a picture of the current sample and helps to compare the experimental and control groups right from the beginning.

The Analysis of Covariance (ANCOVA) procedure was used to partially out any possible covariates, including socioeconomic status, parental education, and home literacy environment. This way, ANCOVA enabled the researchers to make the most accurate estimation of the impact of the gamified learning intervention on literacy scores while excluding the effect of other variables. To conduct the engagement and motivation data analysis, questionnaire responses were analyzed through Exploratory Factor Analysis (EFA) to confirm the constructs of engagement and motivation derived from the developed

questionnaires. Cronbach's alpha was calculated to assess the internal consistency of the scales, ensuring the reliability of the measures used. In addition, this analysis offered a preliminary picture of whether increased levels of activation correlated with improved literacy learning – although the sample and design of this study were not well-suited to this question – and, therefore, potentially how and why gamification may boost learning. All the statistical analysis was attained using SPSS software; the significance criterion was $p < 0.05$.

3.7 Ethical Considerations

This study paid special attention to ethical issues because children under 7 years old were examined. Parental consent was sought for children and all participants; guardian consent was obtained if parents were not around. According to the research ethics, the work was coordinated under the recommendation of an institutional ethics review board. The aim was to protect the identity of participants with anonymity and confidentiality of all data being preserved while undertaking the analysis. Moreover, beneficence was observed in this research since the participants were not harmed. These stages involved subjecting the content to peers and other specialists like educational technology and child psychology authorities. Safeguards to the study were explained to the parents, including the possible risks involved, and they were given a choice to remove their children from the study at any one time at no charge. This was done to ensure that children's participation was voluntary and that the parents were comfortable with the safety and suitability of the study for their children.

4. RESULTS AND DISCUSSION

4.1 Descriptive Statistics

Table 3 shows the demographic characteristics of the participants as mean. The experimental group determines standard deviations (N=100, mean =6.20 and S. D=0.82), followed by the control group total number of participants is N=100 (M=6.05 and S. D=0.79). About 51% of male and 49.0% of female participants are in the experimental group, and 53.0 vs 47.0% of males and females are in the control group, respectively. The socio-economic status of the participants indicates that about 34.0% are low-income, 44.0 % are middle, and 22.0% are high-income in the experiment group. The control group shows 32.0% low-income followed by 46.0% middle and 22.0% high-income students' respondents in this study.

Table 3: Demographic Characteristics of Participants

Characteristic	Experimental group (n=100)	Control Group (n=100)	Total sample (n=200)
Mean (S.D)	6.02 (0.82)	6.05 (0.79)	6.04 (0.80)
Gender			
Male (%)	51 (51%)	53 (53%)	104 (52%)
Female (%)	49 (49%)	47 (47%)	96 (48%)
Socio-economic Status			
Low (%)	34 (34%)	32 (32%)	66 (33%)
Middle (%)	44 (44%)	46 (46%)	90 (45%)
High (%)	22 (22%)	22 (22%)	44 (22%)

4.2 Baseline Literacy Scores

Table 4 shows the baseline literacy skills were assessed using three standardized tests: Phonemic Awareness Literacy Screening (PALS), Peabody Picture Vocabulary Test (PPVT), and Early Grade Reading Assessment (EGRA). The mean reported scores were closely appealed for PALS, PPVT, and EGRA, and the standard deviations presented a normal distribution with each group.

The minor baseline score disparity means that any variation in literacy performance after using the apps as intervention can, therefore, be attributed to the effects of the learning apps and not due to pre-intervention variation between the two sets of students. The baseline scores' means, standard deviations, and ranges are presented below.

Table 4: Baseline Literacy Scores

Test	Experimental group (n=100)	Control Group (n=100)	Total Sample (n=200)
PALS			
Mean (SD)	47.2 (6.8)	46.9 (7.1)	47.0 (6.9)
Range	30-60	29-60	29-60
PPVT			
Mean (SD)	98.5 (10.2)	99.1 (9.8)	98.8 (10.0)
Range	75-120	76-121	75-121
EGRA			
Mean (SD)	85.3 (11.5)	84.7 (11.9)	85.0 (11.7)
Range	60-100	58-102	58-102

4.3 Pretest and Post-test Comparisons

This section analyses the results of the paired sample t-tests to ascertain the difference between the experimental and control groups' pretest and post-test scores. They intended to evaluate the extent of the transition of phonemic awareness, vocabulary, and reading comprehension after the intervention times.

4.3.1 Changes in Literacy Outcomes Within the Experimental Group

To assess the effectiveness of the proposed integrated gamified mobile learning applications on the experimental group's literacy performance, paired sample t-tests were executed. The results indicated significant improvements in all three areas: phonemes, number of words, and reading skills. Table 5 presents the results of the paired sample t-test for the experimental group, highlighting significant improvements across all literacy measures. For the Phonemic Awareness Literacy Screening (PALS), the mean score increased from 47.2 (pretest) to 53.4 (post-test), with a mean difference of 6.2 ($SD = 5.2$, $t = 9.52$, $p < 0.001$). Similarly, vocabulary development, as measured by the Peabody Picture Vocabulary Test (PPVT), showed a mean increase from 98.5 to 104.3, yielding a mean difference of 5.8 ($SD = 6.1$, $t = 7.55$, $p < 0.001$). Reading comprehension, assessed using the Early Grade Reading Assessment (EGRA), demonstrated the most considerable mean difference of 6.8, with pretest and post-test means of 85.3 and 92.1, respectively ($SD = 7.3$, $t = 8.79$, $p < 0.001$). These results indicate statistically significant improvements in all literacy skills assessed following the intervention with gamified learning apps.

Table 5: Paired Sample T-test Results for Experimental Group

Test	Mean Pretest	Mean Post-test	Mean Difference	Standard Deviation	t-value	p-value
PALS	47.2	53.4	6.2	5.2	9.52	<0.001
PPVT	98.5	104.3	5.8	6.1	7.55	<0.001
EGRA	85.3	92.1	6.8	7.3	8.79	<0.001

4.3.2 Changes in Literacy Outcomes Within the Control Group

In the control group, paired sample t-tests were also carried out to establish if there were changes in literacy regarding traditional approaches. The results' analysis revealed slight enhancements in all three domains; however, the growth rates painted a less optimistic picture, indicating that the increase was less significant than the one observed in the experimental group. Table 6 presents the paired sample t-test results for the control group, indicating modest but statistically significant improvements across all literacy measures. For the Phonemic Awareness Literacy Screening (PALS), the mean increased from 46.9 (pretest) to 48.3 (post-test), with a mean difference of 1.4 (SD = 3.5, t = 2.83, p = 0.006). The Peabody Picture Vocabulary Test (PPVT) showed a mean increase from 99.1 to 100.4, with a mean difference of 1.3 (SD = 4.2, t = 2.15, p = 0.034). Similarly, the Early Grade Reading Assessment (EGRA) demonstrated a mean improvement from 84.7 to 86.1, yielding a mean difference of 1.4 (SD = 4.8, t = 2.29, p = 0.024).

Table 6: Paired Sample T-test Results for Control Group

Test	Mean Pretest	Mean Post-test	Mean Difference	Standard Deviation	t-value	p-value
PALS	46.9	48.3	1.4	3.5	2.83	0.006
PPVT	99.1	100.4	1.3	4.2	2.15	0.034
EGRA	84.7	86.1	1.4	4.8	2.29	0.024

4.4 Group Comparisons

Table 7 analysis sought to isolate the changes in literacy, as measured by the intention and post-test, that were statistically significant in terms of the effect of the gamified apps. Also, the analysis of effect sizes was performed to determine the size of these differences. Table 7 displays the results of the independent sample t-test comparing post-test scores between the experimental and control groups. For the Phonemic Awareness Literacy Screening (PALS), the experimental group achieved a higher mean score (M = 53.4, SD = 5.2) compared to the control group (M = 48.3, SD = 3.5), with a significant difference (t = 7.67, p < 0.001, Cohen's d = 1.09), indicating a large effect size. Similarly, for the Peabody Picture Vocabulary Test (PPVT), the experimental group had a mean score of 104.3 (SD = 6.1) versus 100.4 (SD = 4.2) in the control group, with a significant difference (t = 6.09, p < 0.001, Cohen's d = 0.86), representing a large effect size. The Early Grade Reading Assessment (EGRA) showed a mean score of 92.1 (SD = 7.3) for the experimental group compared to 86.1 (SD = 4.8) for the control group, with a significant difference (t = 5.71, p < 0.001, Cohen's d = 0.80), also reflecting a large effect size.

Table 7: Independent Sample T-test Results for Post-test Scores

Test	Group	Mean	Standard Deviation	t-value	p-value	Effect Size (Cohen's d)
PALS	Experimental	53.4	5.2	7.67	<0.001	1.09
	Control	48.3	3.5			
PPVT	Experimental	104.3	6.1	6.09	<0.001	0.86
	Control	100.4	4.2			
EGRA	Experimental	92.1	7.3	5.71	<0.001	0.8
	Control	86.1	4.8			

4.5 Analysis of Covariance (ANCOVA)

As a more statistically stringent method to establish causality and assess the influence of the game-based mobile learning apps on literacy while ignoring potentially confounding variables, an Analysis of Covariance (ANCOVA) was performed. It is a statistical method founded on variance and regression analysis, aiming to assess the impact of covariates, which impact the dependent variable. In this study, potential confounding variables were socio-economic status, parental education, and home literacy environment, which could, in themselves, moderate the literacy outcomes of the program. By adding these variabilities as covariates, ANCOVA brings the post-test scores back to a common baseline and adjusts for the confounding variables, which might have contributed to the difference between the scores of the experimental and the control groups during the pretest.

4.5.1 Controlling for Confounding Variables

Table 8 summarizes the ANCOVA results for literacy outcomes, comparing the adjusted post-test means between the experimental and control groups while controlling for pretest scores. For the Phonemic Awareness Literacy Screening (PALS), the adjusted mean for the experimental group was 53.1, significantly higher than the control group's 48.6 ($F = 65.23$, $p < 0.001$, partial $\eta^2 = 0.251$), indicating a large effect size. Similarly, for the Peabody Picture Vocabulary Test (PPVT), the experimental group's adjusted mean of 104 exceeded the control group's 100.7 ($F = 43.67$, $p < 0.001$, partial $\eta^2 = 0.182$), also reflecting a significant effect. The Early Grade Reading Assessment (EGRA) showed an adjusted mean of 91.8 for the experimental group compared to 86.4 for the control group ($F = 36.54$, $p < 0.001$, partial $\eta^2 = 0.158$), indicating a moderately large effect size.

Table 8: ANCOVA Results for Literacy Outcomes

Test	Group	Adjusted Mean	F-value	p-value	Partial η^2
PALS	Experimental	53.1	65.23	<0.001	0.251
	Control	48.6			
PPVT	Experimental	104	43.67	<0.001	0.182
	Control	100.7			
EGRA	Experimental	91.8	36.54	<0.001	0.158
	Control	86.4			

4.6 Exploratory Factor Analysis (EFA) for Engagement and Motivation

Table 9 used in data analysis was Exploratory Factor Analysis (EFA), which was used to establish the validity of the constructs of engagement and motivation; reliability analysis through the use of Cronbach alpha and t-tests was used to compare engagement and motivation of the experimental group with that of the control group. The EFA showed that two factors require extraction: engagement and motivation. For Engagement (Factor 1), the items included enjoyment and perceived difficulty. The Frequency of use ranged from Factor Loading 0. 72 to 0. 78. Thus, Factor 2, named "Motivation", concerned items related to intrinsic and extrinsic motivation and had factor loadings ranging from 0. 74 to 0. 81. These results validate the author's assertion that the questionnaire items suitably measured the envisaged constructs.

Table 9: Factor Loadings for Engagement and Motivation Constructs

Item	Factor 1 (Engagement)	Factor 2 (Motivation)
Enjoyment of the app	0.78	
Perceived difficulty	0.72	
Frequency of use	0.75	
Intrinsic motivation (enjoyment)		0.81
Intrinsic motivation (interest)		0.77
Extrinsic motivation (usefulness)		0.74
Extrinsic motivation (goal setting)		0.76

4.6.1 Comparative Analysis

To establish whether gamified apps impacted engagement and motivation scores, the independent samples t-test was employed to compare the experimental group's results with those of the control group, which learnt from conventional learning materials. Table 10 shows that the results of engagement observed in the experimental group were relatively higher than that of the control group, with a mean of 4. 20 (SD = 0. 65) for the experimental and 3. 75 (SD = 0. 70) for the control group t = 5. 34($p<0.001$). In the same regard, the mean motivation score of the experimental group was higher, as acknowledged through the mean score (M = 4. 35, SD = 0. 60), than that of the control group (M = 3. 90, SD = 0. 65) with t = 4. 99 ($p < 0.001$).

Table 10: Independent Sample T-test Results for Engagement and Motivation Scores

Scale	Group	Mean	Standard Deviation	t-value	p-value
Engagement Scale	Experimental	4.2	0.65	5.34	<0.001
	Control	3.75	0.7		
Motivation Scale	Experimental	4.35	0.6	4.98	<0.001

4.7 Regression Analysis

To establish the impact of engagement levels on literacy, we undertook a series of multiple regression analyses. In these analyses, we sought to determine how the level of engagement with the given gamified mobile learning applications explained the variations in improvements in phonemic awareness, word recognition, and comprehension among the students.

4.7.1 Relationship Between Engagement and Literacy Outcomes

The regression analysis revealed that mobile learning apps incorporating games positively affected all three literacy domains. Table 11 shows phonemic awareness, the engagement coefficient was (B = 2. 35, SE = 0. 45, β = 0. 42) and was statically significant at (t (198) = 5. 22, p < 0. 001), which means that the higher engagement level is directly related to the higher improvements of Phonemic Awareness. The proposed model was capable of predicting 52% of the total variance of post-test PALS scores (R^2 = 0. 52, Adjusted R^2 = 0. 51). Table 12 shows vocabulary, the engagement coefficient (B = 1. 98, SE = 0. 50, β = 0. 38) was also significant, t (198) = 3. Students in the high treatment group had a higher vocabulary reception score than those in the low treatment group F = 96, p < 0. 001, indicating that a higher level of engagement would enhance vocabulary improvements.

The proposed model explained 46% of the post-test PPVT score (F = 29. 737; R^2 = 0. 46 adjusted R^2 = 0. 45). Table 13 shows Concerning the reading comprehension, data showed that engagement coefficient (B = 2. 45, p < 0. 001 denoting that the level of engagement has a positive correlation with the improvements of students' reading comprehension. The model accounted for 49 per cent of the change in the scores in the post-test EGRA, R^2 = 0. 49, Adjusted R^2 = 0. 48).

Table 11: Regression Analysis for Phonemic Awareness (PALS)

Variable	B	SE	B	t-value	p-value
Engagement	2.35	0.45	0.42	5.22	<0.001
Baseline PALS Score	0.78	0.1	0.6	7.8	<0.001
R^2 = 0.52					
Adjusted R^2 = 0.51					

Table 12: Regression Analysis for Vocabulary (PPVT)

Variable	B	SE	B	t-value	p-value
Engagement	1.98	0.5	0.38	3.96	<0.001
Baseline PPVT Score	0.84	0.12	0.65	7	<0.001
R^2 = 0.46					
Adjusted R^2 = 0.45					

Table 13: Regression Analysis for Reading Comprehension (EGRA)

Variable	B	SE	B	t-value	p-value
Engagement	2.45	0.55	0.4	4.45	<0.001
Baseline EGRA Score	0.7	0.14	0.55	5	<0.001
R^2 = 0.49					
Adjusted R^2 = 0.48					

4.8 Discussion

This study aimed to evaluate the effectiveness of gamified mobile learning apps in enhancing early childhood literacy development. The results offer thus strong support for the statement that the introduction

of the learning intervention based on gamification positively influenced phonemic awareness, with the effect size reflecting moderate to high improvements in vocabulary and reading comprehension among young first-graders. The experimental group using the game-based apps improved significantly on all the literacy indices compared to the control group, which carried out non-game-related activities. In addition, other significant and positive correlation findings included engagement levels and literacy, meaning that the more students' participation in the apps was enhanced, the better the literacy improvement. The improvements identified in the experimental group are noteworthy and result from the features of gamified applications, which incorporate active and motivational approaches to learning. The elements of points, badges, interactive narratives, and feedback might have helped keep and motivate the students' interest. This corresponds with the Self-Determination Theory (Deci & Ryan, 2013).

High engagement emerged as a significant predictor of literacy performance according to the regression analyses conducted in the study. A more substantial positive coefficient for the dependent variable, engagement across phonemic awareness, word, and reading comprehension, implies that students who played with the gamified apps more had a better improvement across the three. This is because students may need more motivation to participate in educational intercessions; therefore, experts must develop intercessions that will engage students and encourage results-oriented learning. It also accords with earlier findings stating that involvement is another crucial element in determining academic performance (Fredricks et al., 2004). This study also aligns with the previous work on gamification in education and its positive outcomes. Research in this area has revealed that learning enhanced through games positively impacts motivation, participation, and academic achievement (Denden et al., 2024). In extending the literature, this study hopes to contribute to knowledge in the field by presenting the positive effect of gamified mobile learning apps on early childhood literacy learning.

5. CONCLUSION AND RECOMMENDATIONS

This research aimed to assess the effects of mobile learning apps incorporating game design elements in young children's literacy learning in China. The studies prove that such game-based learning instruments improve core reading skills such as phonemic awareness, vocabulary, and reading comprehension. The experimental group that utilized the gamified apps significantly increased the mean scores of literacy skills compared to the control group, which went through the conventional activities. On this score, the presented outcomes confirm the efficiency of the gamified approach toward enhancing learning outcomes among students of a younger age. According to the study and results, engagement was also a significant predictor of literacy gains. It was also found that significantly higher levels of engagement in the gamified apps were highly correlated with the degree of improvement in literacy, proving that the use of related forms of interactive education, as in the case of gamification, can indeed capture the interest and sustain the learning interest of the children. This corroborates the concepts developed in motivational and engagement theories, thus promoting gamified aspects in educational interventions.

There are crucial practical implications of the facts; therefore, teachers are advised to adopt elements of games in learning activities to foster system interaction and thus raise the fun quotient of learning. Authorities should encourage the use of game-based learning applications, especially in less-provided districts, to advance learners' literacy comprehensively. Technology developers are encouraged to create stimulating educational applications appropriate to the student's cultural background and appropriate the age of students in schools. Nevertheless, there are significant limitations, such as the study duration and the cultural background. Further studies should look at the post-intervention utility of the

usage of gamification and, at the same time, control factors such as cultural differences and educational calendars to ensure the results obtained are consistent.

6. LIMITATIONS AND FUTURE STUDIES

Despite the significant findings demonstrating the effectiveness of gamified mobile learning apps in enhancing early childhood literacy development in China, several limitations must be acknowledged. Firstly, the duration of the study was relatively brief, spanning only three months. This limited study period might not capture the overall impact of the changes on the observed literacy levels and their sustainability. As a result of the brief nature of the studies considered, future research should aim to continue the investigations for a more extended period to estimate the long-term outcomes of the game-based learning approach to literacy acquisition. Second, it is evidenced that other factors could impact the research results besides the variables corrected by ANCOVA, including socio-economic status, parental education, and home literacy environment. Other limitations that may have influenced the findings include differences in the learners' learning styles, past technology literacy, and variances in teachers' application of the gamified apps. Subsequently, the same variables should be included in subsequent studies to improve the understanding of the factors that impact the efficiency of gamified learning.

Author contributions: L.L. developed the main idea and, with the help of J.Z, conducted the fieldwork, Z.G. contributed to methods, and Y.W. worked on literature and analysis. All authors agreed and reviewed the published version of the manuscript.

Ethical Statement: Parental consent was sought for children and all participants; guardian consent was obtained when parents were not around. According to the research ethics, the work was coordinated under the recommendation of an institutional ethics review board. The aim was to protect the identity of participants with anonymity and confidentiality of all data being preserved while undertaking the analysis. Moreover, beneficence was observed in this research since the participants were not harmed. These stages involved subjecting the content to peers and other specialists like educational technology and child psychology authorities. Safeguards to the study were explained to the parents, including the possible risks involved, and they were given a choice to remove their children from the study at any one time at no charge. This was done to ensure that children's participation was voluntary and that the parents were comfortable with the safety and suitability of the study for their children.

Competing Interests: not declared.

Grant/Funding information: The author declared that no grants supported this work.

Data Availability Statement: The associated data is available upon request from the corresponding author.

Declaration Statement of Generative AI: The author has not used any AI tool to prepare the manuscript.

REFERENCES

AlAwadhi, S., & Al-Daihani, S. M. (2019). Marketing academic library information services using social media. *Library Management*, *40*(3/4), 228-239.https://doi.org/10.1108/LM-12-2017-0132.

Cheong, C., Filippou, J., & Cheong, F. (2014). Towards the gamification of learning: Investigating student perceptions of game elements. *Journal of Information Systems Education*, *25*(3), 233. https://jise.org/volume25/n3/JISEv25n3p233.html.

Cooper, G. (1990). Cognitive load theory as an aid for instructional design. *Australasian Journal of Educational Technology*, *6*(2). https://doi.org/10.14742/ajet.2322.

Deci, E. L., & Ryan, R. M. (2013). *Intrinsic motivation and self-determination in human behavio*ur. Springer Science & Business Media. https://doi.org/10.1007/978-1-4899-2271-7.

De-Marcos, L., Domínguez, A., Saenz-de-Navarrete, J., & Pagés, C. (2014). An empirical study comparing gamification and social networking on e-learning. *Computers & Education*, *75*, 82-91. https://doi.org/10.1016/j.compedu.2014.01.012.

Denden, M., Tlili, A., Chen, N. S., Abed, M., Jemni, M., & Essalmi, F. (2024). The role of learners' characteristics in educational gamification systems: A systematic meta-review of the literature. *Interactive Learning Environments*, *32*(3), 790-812. https://doi.org/10.1080/10494820.2022.2098777.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, *74*(1), 59-109. https://doi.org/10.3102/00346543074001059.

Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in Entertainment (CIE)*, *1*(1), 20-20. https://doi.org/10.1145/950566.950595.

Goldman, J. D. (2008). Responding to parental objections to school sexuality education: A selection of 12 objections. *Sex Education*, *8*(4), 415-438. https://doi.org/10.1080/14681810802433952.

Hamari, J., Koivisto, J., & Sarsa, H. (2014, January 6-9). Does gamification work? A literature review of empirical studies on gamification. In *2014 47th Hawaii International Conference on System Sciences* (pp. 3025-3034). IEE. https://doi.org/10.1109/HICSS.2014.377.

Hammedi, W., Leclercq, T., & Steils, N. (2024). Gamification Myopia: satiation effects in gamified activities. *Journal of Service Research*, *27*(2), 213-230. doi.org/10.1177/10946705231190873.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, *80*, 152-161. doi.org/10.1016/j.compedu.2014.08.019.

Huang, J., & Zhou, L. (2021). Social gamification affordances in the green IT services: perspectives from recognition and social overload. *Internet Research*, *31*(2), 737-761. https://doi.org/10.1108/INTR-03-2020-0121.

Jaskari, M. M., & Syrjälä, H. (2023). A mixed-methods study of marketing students' game-playing motivations and gamification elements. *Journal of Marketing Education*, *45*(1), 38-54. https://doi.org/10.1177/02734753221083220.

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.

Kashive, N., & Mohite, S. (2023). Use of gamification to enhance the e-learning experience. *Interactive Technology and Smart Education*, *20*(4), 554-575. https://doi.org/10.1108/ITSE-05-2022-0058

Liatukaitė, M. (2024). *Gamification in mobile applications: influence of immersion, achievement and social-related features on brand loyalty* [Doctoral dissertation, Vilniaus universitetas]. https://epublications.vu.lt/object/elaba:191573735/.

Lopes, J. M., Gomes, S., & Rodrigues, I. (2024). Playing the gamification and co-creation game: a bibliometric literature review. *Management Review Quarterly*, *74*(1), 353-376. https://doi.org/10.1007/s11301-022-00302-8.

Mandujano, G. G., Riar, M., Morschheuser, B., & Hamari, J. (2023). Gamification as a Catalyst to the Circular Economy. H. Lehtimäki, L. A. Stenroos, A. Jokinen, P. Jokinen (Eds). In *The Routledge Handbook of Catalysts for a Sustainable Circular Economy* (pp. 312-336). Routledge. https://doi.org/10.4324/9781003267492-18.

Mazarakis, A., & Bräuer, P. (2023). Gamification is working, but which one exactly? Results from an experiment with four-game design elements. *International Journal of Human-Computer Interaction*, *39*(3), 612-627. https://doi.org/10.1080/10447318.2022.2041909.

Nacke, L. E., & Deterding, S. (2017). The maturing of gamification research. *Computers in Human Behavior*, *71*, 450-454. https://doi.org/10.1016/j.chb.2016.11.062.

Nicholson, S. (2014). A Recipe for Meaningful Gamification. T. Reiners, C. W. Lincoln (Eds). In *Springer eBooks* (pp. 1–20). https://doi.org/10.1007/978-3-319-10208-5_1.

Nugroho, S. S. (2024). Gamification aspects affecting mobile app continued use, attitude, and satisfaction. *Jurnal Siasat Bisnis*, *28*(1), 19–36. https://doi.org/10.20885/jsb.vol28.iss1.art2.

Oliveira, W., Hamari, J., Shi, L., Toda, A. M., Rodrigues, L., Palomino, P. T., & Isotani, S. (2023). Tailored gamification in education: A literature review and future agenda. *Education and Information Technologies*, *28*(1), 373-406. https://doi.org/10.1007/s10639-022-11122-4.

Rosli, N., & Omar Zaki, H. (2023). A Bibliometric Review of Research on Gamification in Marketing: Reflections for Moving Forward. *International Journal of Management Studies (IJMS)*, *30*(2), 271-300. http://dx.doi.org/10.32890/ijms2023.30.2.4.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of human-computer Studies*, *74*, 14-31. https://doi.org/10.1016/j.ijhcs.2014.09.006.

Xu, H., & Hamari, J. (2023). How to improve creativity: a study of gamification, money, and punishment. *Behaviour & Information Technology*, *42*(15), 2545-2559. doi.org/10.1080/0144929X.2022.2133634.

Zainuddin, Z., Chu, S. K. W., & Perera, C. J. (2024). Gamification in Education. In *Gamification in A Flipped Classroom: Pedagogical Methods and Best Practices* (pp. 67-113). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-97-2219-8_3.

Zvereva, G., Xi, N., Hamari, J., & Pirkkalainen, H. (2023, December 5-8th). Gamification in Ethics Education: A Literature Review. In *Australasian Conference on Information Systems (ACIS) 2023* (Australasian Conference on Information Systems, Wellington, New Zealand). https://aisel.aisnet.org/acis2023/70.