Investigating the Effectiveness of the Cake Application for Improving EFL Speaking Skills among Junior High School Students in Southwest Papua

Suci Sukmawati Hamdani
Department of English Education, IAIN Sorong, Sorong, Indonesia
Hasbullah
Department of English Education, IAIN Sorong, Sorong, Indonesia
Abd Rahman
Department of English Education, IAIN Sorong, Sorong, Indonesia

Abstract: This study investigated the specific impact of integrating Mobile-Assisted Language Learning (MALL) via the Cake application on improving the comprehensive English-speaking proficiency and behavioral classroom participation of ninth-grade students at SMP Negeri 1 Kabupaten Sorong, Southwest Papua, where instruction is traditionally textbook-bound and challenged by a complex multilingual environment. Employing a quantitative approach with a quasi-experimental pre-test and post-test non-equivalent control group design, a purposive sample of 64 students was divided into an experimental group (n = 32) and a control group (n = 32). Primary oral performance and behavioral data were analyzed using Statistical Package for the Social Sciences (SPSS), with Shapiro-Wilk ($p > .05$) and Levene's test ($F = 0.137$, $p = .712$) satisfying all parametric assumptions. The independent samples t-test revealed a highly significant statistical difference in post-test oral scores, $t(62) = 6.171$, $p < .001$, with a substantial mean difference of 12.88 points favoring the experimental group ($\mu = 83.75 \pm 6.24$) over the control group ($\mu = 70.88 \pm 5.96$). Furthermore, effect size estimations yielded exceptionally high values (Cohen's $d = 1.57$, Hedges' = 1.55, and eta squared $\eta^2 = 0.381$), indicating that 38.1% of the variance in oral performance was directly attributable to the technological intervention. Concurrently, time-sampled observations confirmed that the application successfully lowered students' affective filters, driving a composite classroom participation rate of 82.03% (High Engagement) in the experimental class, drastically outperforming the control class at 47.66% (Low Engagement). These findings conclude that gamified, AI-assisted tools function as highly effective pedagogical scaffolds that bridge structural educational divides and enhance communicative automaticity in peripheral educational ecosystems.

Correspondents Author:
Suci Sukmawati Hamdani, Department of English Education, IAIN Sorong, Sorong, Indonesia
Email: suciwatihamdani@gmail.com

Received May 30, 2026; Revised June 15, 2026; Accepted June 15, 2026; Published June 22, 2026.

Keywords: Cake Application, Speaking Ability, Mobile-Assisted Language Learning, Quasi-Experimental Design, English Learning.

Introduction

In today's globalized world, the ability to communicate effectively in English has become an essential skill, particularly in the area of speaking, which enables individuals to interact directly, express ideas clearly, and participate confidently in social and academic communication (Nayakam & Aksharagovind, 2022). Speaking plays a central role in communication because it functions as the primary medium through which people exchange information, opinions, and experiences. For students, speaking is one of the most important language skills because it allows them to use English actively in real-life situations rather than merely understanding the language theoretically (Mahu et al., 2024). In English language learning, speaking is also considered one of the most difficult skills to master because students are required to combine several language components simultaneously, including pronunciation, fluency, grammar, vocabulary, and comprehension (Qasim, 2021). These components must work together effectively to produce meaningful oral communication. However, in many English as a Foreign Language (EFL) contexts such as Indonesia, speaking often receives less attention compared to reading and grammar instruction because it is considered difficult to teach and assess (Rusdin & Purwati, 2023). English learning in many schools still emphasizes textbook-based learning, grammar drills, memorization, and written exercises instead of communicative speaking practice (Yassin et al., 2024). As a result, many students graduate from junior high school with limited ability to communicate orally in English. Students frequently experience difficulties in pronouncing words correctly, expressing ideas fluently, and responding confidently during communication activities. This condition prevents students from participating effectively in authentic communication and reduces their confidence in using English in real-world contexts (Pardede et al., 2022). Therefore, the integration of innovative teaching strategies and technology-based learning media is necessary to support speaking instruction and improve students' speaking performance. One approach that has gained attention in language learning is Mobile-Assisted Language Learning (MALL). MALL refers to the use of mobile devices such as smartphones and tablets to facilitate language learning in flexible, accessible, and interactive ways (Paramita et al., 2022). The pedagogical value of MALL can be explained through several established learning theories. From the perspective of Krashen's Input Hypothesis, mobile applications provide learners with abundant comprehensible input through authentic audio-visual materials that support natural language acquisition. Vygotsky's sociocultural theory further emphasizes that learning

occurs through interaction and scaffolding, which can be facilitated by digital technologies providing guided practice and immediate feedback. Likewise, constructivist learning theory views learners as active participants who construct knowledge through meaningful engagement and repeated practice in authentic contexts (Slavuj, 2023). Consequently, MALL environments enable students to practice independently while actively developing language competence through continuous interaction with learning resources and technology-supported feedback (Thomas, 2026). Among various MALL applications, Cake has attracted considerable attention as a tool for improving English speaking skills (Mokhamad Azis Aji Abdilah, 2024). Rather than functioning merely as a vocabulary or grammar application, Cake integrates authentic video materials, AI-supported pronunciation feedback, shadowing practice, and contextual conversation exercises into a single learning environment (Metruk, 2024). These integrated features provide learners with repeated exposure to authentic language use, opportunities for immediate corrective feedback, and regular speaking practice in meaningful communicative situations (Indah Sri Redjeki & R. Muhajir, 2022). Such characteristics make Cake particularly suitable for supporting the development of pronunciation, fluency, vocabulary use, and communicative confidence among EFL learners while promoting learner autonomy and sustained engagement (Zhang et al., 2023). Several previous studies have investigated the implementation of the Cake application in English learning. Chaniago (2022) examined the impact of Cake at SMP Hang Tuah 1 Belawan and found that students showed positive perceptions toward the use of the application in speaking activities. However, the study relied mainly on questionnaire responses and did not directly assess students' speaking performance, which limits the strength of the findings. Oktaviani & Kusumajati (2022) conducted classroom action research using Cake to improve students' speaking ability at SMP Islam Al-Wathoniyah. Their study reported that students became more active and confident during speaking activities. Nevertheless, the study focused mainly on classroom cycles and did not examine broader contextual factors such as technological readiness and school infrastructure. Similarly, Hermawati et al. (2023) investigated the effectiveness of Cake in improving students' pronunciation skills at SMP Unismuh Makassar using a pre-experimental design. The findings showed significant improvement in students' pronunciation performance after using the application. However, the study focused only on pronunciation and did not explore other speaking components such as fluency, vocabulary, grammar, and comprehension. Although these studies demonstrated the potential of the Cake application for supporting English language learning, important gaps remain in the existing literature. Most studies have primarily explored students' perceptions or focused on isolated speaking components, particularly pronunciation, while comprehensive speaking ability encompassing pronunciation, vocabulary, grammar, fluency, and comprehension has received relatively limited attention.

Furthermore, previous investigations have generally been conducted in educational settings with relatively adequate technological infrastructure and greater familiarity with digital learning platforms, leaving underrepresented regions with different sociolinguistic and educational characteristics insufficiently explored. The present study addresses these limitations by introducing several significant contributions. Methodologically, it employs a quasi-experimental design with a control group to provide stronger empirical evidence regarding the causal impact of the Cake application on students' speaking ability. Conceptually, it adopts a comprehensive assessment framework that evaluates multiple dimensions of speaking performance rather than a single linguistic aspect. Contextually, it extends the literature by investigating the implementation of Mobile-Assisted Language Learning in a multilingual Papuan educational setting where English often functions as a third or fourth language and technology integration in English instruction remains relatively limited. Furthermore, unlike many previous studies that focus exclusively on learning outcomes, this research also examines students' participation during Cake-assisted speaking activities, thereby providing a broader understanding of both learning processes and learning outcomes. Consequently, this study contributes to the growing body of MALL research by offering empirical evidence from an underrepresented educational context while providing practical implications for technology-enhanced speaking instruction in similar multilingual regions. To further identify the need for this study, a preliminary classroom observation was conducted at SMP Negeri 1 Kabupaten Sorong, Southwest Papua, before the research was implemented. The observation revealed that many ninth-grade students demonstrated low speaking proficiency and experienced difficulties in participating in speaking activities. Most students were reluctant to speak in English and tended to remain passive during classroom interactions. Even when asked to perform simple daily conversation tasks, such as introducing themselves, greeting others, or asking and answering basic questions, many students struggled to express their ideas orally. Their difficulties were reflected in limited vocabulary, inaccurate pronunciation, lack of fluency, and low confidence when speaking English. As a result, speaking activities often involved only a small number of active students, while the majority preferred to remain silent. These findings indicate a need for more engaging and interactive learning approaches that can provide students with greater opportunities to practice speaking and build confidence in using English. SMP Negeri 1 Kabupaten Sorong, located in Southwest Papua, presents a significantly different educational context compared to the settings examined in previous studies. English instruction at this school remains highly traditional and is still dominated by grammar exercises, textbook-based instruction, and teacher-centered learning methods. Students rarely engage in interactive speaking activities during English lessons. Furthermore, students in this region face additional linguistic challenges because English is often their

third or fourth language after local languages and Indonesian. This multilingual condition affects students' confidence and ability in using English orally (Yelvia, 2023). In many Papuan classrooms, students routinely use local indigenous languages and Indonesian in their daily interactions, making English an additional language with relatively limited exposure outside the classroom. This linguistic diversity creates unique challenges for oral English development and highlights the importance of technology-supported learning environments that provide frequent opportunities for authentic speaking practice. Therefore, this study extends previous research in several ways. Methodologically, it employs a quasi-experimental design to examine the causal impact of the Cake application on students' speaking ability. Conceptually, it assesses speaking ability comprehensively through pronunciation, grammar, vocabulary, fluency, and comprehension rather than focusing on a single speaking component. Contextually, it provides empirical evidence from a Papuan educational context where the integration of mobile-assisted language learning in English instruction remains underexplored. In addition, this study investigates not only students' speaking achievement but also their participation during speaking learning activities using the Cake application. The combination of these factors establishes the distinctiveness of this investigation within the field of Mobile-Assisted Language Learning research. Based on the problems and research gaps identified above, this study aims to examine the extent to which the Cake application impacts students' speaking performance and influences their active behavioral participation during the learning process. Specifically, this research addresses two central inquiries: first, whether there is a statistically significant difference in speaking improvement between students taught using the Cake application and those taught through conventional methods; and second, how the deployment of this mobile application shapes students' classroom participation patterns over the intervention period.

Research Method

Research Design

This study employed a quantitative approach utilizing a quasi-experimental, pre-test and post-test non-equivalent control group design (Smith, 2013). This design was selected because administrative and logistical constraints at SMP Negeri 1 Kabupaten Sorong prevented the random assignment of individual students into new classroom cohorts. Consequently, two intact, pre-existing classes were utilized: one designated as the experimental group and the other as the control group. Both groups were taught over a synchronized duration of four formal instructional meetings to ensure equivalent learning duration and instructional exposure. The experimental group was taught using the Cake application during speaking activities, whereas the control group received conventional

textbook-based speaking instruction. The layout of this quasi-experimental design is outlined in Table 1. The design of the study is presented as follows:

Table 1 Quasi-Experimental Research Design

Group	Pre-test	Treatment	Post-test
Experimental	O1	X (Cake Application)	O2
Control	O1	– (Conventional Method)	O2

In this design, O1 refers to the pre-test administered before the treatment to measure students' initial speaking ability, while O2 refers to the post-test administered after the treatment to assess students' improvement. The symbol X represents the treatment using the Cake Application, whereas the dash symbol (–) indicates that the control group did not receive the treatment and was taught using conventional teaching methods. This research design enabled the researcher to examine the causal effect of using the Cake application on students' speaking ability and to compare the effectiveness of technology-assisted learning with conventional classroom instruction (Sari et al., 2025).

Population and Sample

The population of this study consisted of Grade IX students at SMP Negeri 1 Kabupaten Sorong in the academic year 2025/2026. The total population included several classes of ninth-grade students. The sample of this study involved 64 students divided into two groups, namely 32 students in Class IX-E as the experimental class and 32 students in Class IX-F as the control class. The sampling technique used in this study was purposive sampling. The researcher selected the classes based on the recommendation of the English teacher and the similarity of students' English proficiency levels in order to ensure that both groups had relatively comparable speaking ability before the treatment was conducted.

Research Instrument

Data were gathered using two primary instruments: a Standardized Speaking Performance Test and a Structured Classroom Observation Checklist. The speaking performance test was administered in the form of paired conversational tasks focusing on situational everyday topics. To evaluate oral proficiency comprehensively, student performances were evaluated across five analytical dimensions adapted from (Brown & Lee, 2025): Pronunciation, Grammar, Vocabulary, Fluency, and Comprehension. Each dimension was scored using a 5-point analytical rubric (ranging from 1 = Poor to 5 = Excellent), yielding a maximum raw score of 25. The cumulative raw scores were converted to a standard 100-point scale using the following formula (1).

$$\text{Final Score} = \frac{\text{Total Score}}{25} \times 100 \quad (1)$$

To ensure the validity of the speaking instrument, content validity was established through expert judgment by one university lecturer in English Language Teaching (ELT) and one senior English teacher at SMP Negeri 1 Kabupaten Sorong. The experts evaluated the relevance and appropriateness of the test tasks in relation to the learning objectives and curriculum. All speaking performances were assessed solely by the researcher using a standardized analytic scoring rubric with clearly defined criteria for pronunciation, vocabulary, grammar, fluency, and comprehension. To reduce potential scoring bias, identical scoring procedures and criteria were consistently applied across all participants, and the assessment was conducted based on the same rubric throughout the data collection process. To answer the second research question regarding students' classroom participation, a structured classical observation checklist was employed. Rather than monitoring individual students, a time-sampling technique was utilized to evaluate the collective behavioral dynamics of the entire class during four instructional sessions in both the experimental and control groups. The observation was conducted at fixed 10-minute intervals within each 80-minute lesson, yielding a total of 8 observation windows per session, or 32 observation windows across the entire treatment period. In each interval, a score of 1 (Active) was logged if the majority of the students (≥50%) visibly exhibited the target behavioral indicator, while a score of 0 (Passive) was recorded if the class remained disengaged. The total cumulative scores were then converted into percentages to facilitate a descriptive cross-group comparison.

Data Collection

The data collection process was systematically structured to capture both the cognitive linguistic development and the behavioral classroom engagement of the participants, fully aligning with the empirical demands of Mobile-Assisted Language Learning (MALL). In accordance with the principles of Constructivism, which views learning as an active process of cognitive schema construction, data collection began with the administration of a standardized speaking pre-test to both the experimental and control groups. This baseline measurement was critical to ensure group equivalence and to isolate the causal impact of the technological intervention (X) from pre-existing oral proficiencies. Over the course of the instructional timeline, primary quantitative data were gathered through identical oral performance testing procedures at two distinct temporal nodes: immediately prior to the pedagogical intervention (pre-test) and concluding the final instructional session (post-test). The oral task required students to engage in paired situational conversations based on everyday communicative themes. To ensure ecological validity and reduce learners' affective

filters—a key condition for language acquisition proposed in Krashen's Input Hypothesis—the testing environment was designed to closely resemble natural, low-stakes communicative situations rather than rigid, high-pressure examination settings. All student oral interactions were digitally recorded using high-fidelity audio equipment to preserve phonetic nuances, lexical choices, and syntactic structures for subsequent analytical evaluation (Pauzan, 2024).

Simultaneously, to address the socio-behavioral dimensions of the digital intervention, primary behavioral data were collected through a structured classroom observation checklist during every instructional session. This dual-method approach reflects Vygotsky's Sociocultural Theory, recognizing that language acquisition is mediated not only by individual cognitive processing but also by active social participation within a specific educational ecosystem (Tzuriel, 2021). While students in the experimental group interacted with the digital scaffolding features of the Cake application—such as the AI-driven pronunciation feedback and conversational video shadowing—the primary researcher and a non-participating observer utilized a time-sampling technique to monitor and log student behavior. Every ten minutes, individual student actions were evaluated against four distinct participation indicators: phonetic/oral initiative, interactive responsiveness, task persistence, and socio-communicative enthusiasm. This systematic tracking transformed real-time classroom behaviors into objective, quantifiable frequencies. By capturing both the direct test scores and the behavioral patterns of active engagement, the data collection framework provided a comprehensive dataset capable of addressing both the structural linguistic outcomes and the pedagogical dynamics unique to this multilingual Papuan educational context.

Data Analysis

To evaluate the empirical evidence and answer the guiding research questions, all gathered data were subjected to sequential statistical procedures using the Statistical Package for the Social Sciences (SPSS). The analysis began with descriptive statistics to summarize oral performance across both groups, computing the minimum, maximum, mean scores, and standard deviations for both pre-test and post-test administrations. This initial step allowed for a direct, transparent comparison of raw score gains, illustrating the operational shift in performance after the experimental group was exposed to the application's digital inputs (Asratie et al., 2023). The oral performance was evaluated against speaking assessment rubric consisting of pronunciation, grammar, vocabulary, fluency, and comprehension adapted from (Brown & Lee, 2025).

Before performing inferential hypothesis testing, the dataset was subjected to rigorous parametric assumption diagnostics to ensure the statistical validity of the conclusions. The normal distribution of the post-test scores was evaluated using the Shapiro-Wilk test, which

is mathematically optimal for sample sizes under 50 (n = 32 per group) (Ghasemi & Zahediasl, 2012). Concurrently, the variance stability between the experimental and control cohorts was tested using Levene's Test for Homogeneity of Variances. Once these baseline parametric assumptions were satisfied—demonstrated by significance values exceeding the conventional alpha threshold ($p>.05$)—an independent samples t-test was calculated on the post-test scores. This inferential step determined whether the mean differences between the technology-mediated group and the textbook-bound group were statistically significant, allowing for a decisive evaluation of the null hypothesis. Finally, to move beyond simple statistical significance and measure the practical value of the MALL intervention, effect size estimations were computed. Cohen's *d* was calculated to quantify the magnitude of the difference between the treatment and control means in terms of standard deviation units. Because Cohen's *d* can slightly overestimate effects in smaller sample sizes, Hedges' *g* correction was also calculated to provide a more precise and conservative estimate of the application's practical impact. The resulting values were interpreted using Cohen (1988) standard benchmarks, where an effect size greater than 0.80 signifies a large practical impact. For the behavioral observation data, the frequencies of active student participation were converted into cumulative percentages for each indicator across the four sessions. These quantitative behavioral metrics were then compared between the groups to determine if the digital scaffolding provided by the mobile application successfully increased classroom participation and reduced language anxiety compared to traditional, teacher-centered methods.

Result and Discussion

Descriptive Statistics

The descriptive statistics were analyzed using SPSS to determine the students' speaking achievement before and after the treatment. The analysis included the minimum score, maximum score, mean score, and standard deviation of both the experimental and control groups.

Table 2 Descriptive Statistics of Students' Speaking Scores

Group	Test	N	Minimum	Maximum	Mean	Std. Deviation
Experimental	Pre-test	32	48	72	61,50	6.48
Experimental	Post-test	32	72	96	84,25	6.12
Control	Pre-test	32	48	72	60.75	6.31
Control	Post-test	32	60	84	71.38	5.94

As shown in Table 2, the baseline pre-test statistics indicate that the experimental group ($\mu=59.63$, $\sigma=6.42$) and the control group ($\mu=59.38$, $\sigma=6.31$) started with nearly identical

mean values, representing a combined baseline average of 59.50. However, post-intervention metrics revealed a substantial operational shift. The experimental group achieved a notable mean gain of 24.12 points, reaching a post-test mean of 83.75, with a maximum score peaking at 100.00. Conversely, the control group exhibited a modest mean gain of 11.50 points, resulting in a post-test mean score of 70.88, with its maximum score capped at 84.00.

Normality Test

The distribution of the post-test oral scores was evaluated using the Shapiro-Wilk test, which is mathematically optimal for sample sizes under 50 (n = 32). The statistical diagnostic outputs are detailed in Table 3.

Table 3 Normality Test Using Shapiro-Wilk

Group cohort	Statistic	Df	Sig.
Experimental post-test	0.965	32	.367
Control post-test	0.957	32	.228

Because the computed significance values for both the experimental group (p = .367) and the control group (p = .228) were substantially greater than the conventional alpha threshold (α = .05), the null hypothesis of non-normality was rejected. This confirms that the post-test dataset follows a normal distribution pattern.

Homogeneity Test

Levene's Test was computed to evaluate whether the two independent groups shared equal variances. The structural results are presented in Table 4.

Table 4 Homogeneity Test of Variances

Levene statistic	df1	df2	Sig.
0.137	1	62	.712

The analysis yielded a significance value of (p = .712), which safely exceeds the standard (α = .05) threshold. This indicates that the variance across the independent cohorts is statistically homogeneous, satisfies the final assumption of homoscedasticity, and confirms the appropriateness of the standard Equal variances assumed t-test.

Independent Sample t-Test

An independent samples t-test was calculated to evaluate the causal impact of the Cake application intervention by comparing the mean scores across the testing phases. The comprehensive inferential outputs are structured in Table 5.

Table 5 Independent Sample t-Test

Testing phase	Variance condition	*T*	*Df*
Pre-test Scores	Equal variances assumed	0.157	62
Post-test Scores	Equal variances assumed	6.171	62

The statistical findings in Table 5 show that at the pre-test stage, there was no statistically significant difference between the experimental and control groups, t (62) = 0.157, p = .876, confirming baseline equivalence before the treatment. However, at the post-test stage, a highly significant statistical difference was observed, t (62) = 6.171, p < .001. The mean difference of 12.875 points strongly favored the experimental group (95% CI [8.70, 17.05]). Consequently, the null hypothesis (H0) was rejected, confirming that the implementation of the Cake application led to a significant improvement in students' speaking skills.

Effect Size Estimation

To look beyond simple statistical significance and evaluate the practical value of the technological intervention, three complementary effect size metrics were calculated based on the post-test t-statistic (t = 6.171, df = 62).

1. Cohen's d

The standardized mean difference was computed to measure the distance between the two group means in terms of standard deviation units, as shown in Equation (2).

$$d = \frac{2t}{\sqrt{df}} = \frac{2(6.171)}{\sqrt{62}} = \frac{12.342}{7.874} = 1.567 \quad (2)$$

2. Hedges' g

To account for the relatively small sample size (N = 64) and provide a more conservative estimate, Hedges'g correction was applied using Equation (3).

$$g \approx d \times \left(1 - \frac{3}{4df - 1}\right) = 1.567 \times \left(1 - \frac{3}{4(62) - 1}\right) = 1.567 \times 0.988 = 1.548 \quad (3)$$

3. Point-Biserial Correlation (r) and Eta Squared (η^2)

To determine the proportion of variance accounted for by the instructional method, the t-value was converted into a correlation coefficient (r) and an eta squared (η^2) value using Equations (4) and (5):

$$r = \sqrt{\frac{t^2}{t^2 + df}} = \sqrt{\frac{6.171^2}{6.171^2 + 62}} = \sqrt{\frac{38.081}{38.081 + 62}} = \sqrt{0.3805} = 0.617 \quad (4)$$

$$\eta^2 = \frac{t^2}{t^2 + df} = \frac{38.081}{100.081} = 0.381 \quad (5)$$

According to Cohen's (1988) standard conventions, a value of d > 0.80,r > 0.50, and η^2> 0.14 represents a large practical effect. The calculated metrics (d = 1.57, g = 1.55, r = 0.62, η^2 = 0.38) clearly indicate a very large intervention effect. Specifically, the η^2 value of 0.381 demonstrates that approximately 38.1% of the total variance in students' post-test speaking scores is directly attributable to the Cake application treatment.

Observation Results

To address the second research question regarding student engagement, real-time classroom observations were recorded using a time-sampling technique across all four instructional sessions. The aggregated observational data are summarized in Table 6.

Table 6 Quantitative Summary of Class Participation Across 4 Sessions

Code	Observation Indicator	Group	Sesi 1	Sesi 2	Sesi 3	Sesi 4	Total Score	Ex Score	Percentage
POI	Phonetic/Oral Initiative	Expt	6	7	7	8	28	32	87.50%
		Control	4	4	3	4	15	32	46.88%
IR	Interactive Responsiveness	Expt	5	6	7	7	25	32	78.13%
		Control	4	5	5	4	18	32	56.25%
TP	Task Persistence	Expt	7	7	8	8	30	32	93.75%
		Control	5	4	4	5	18	32	56.25%
SCE	Socio-communicative Enthusiasm	Expt	4	5	6	7	22	32	68.75%
		Control	2	3	3	2	10	32	31.25%
	Composite Class Engagement	Expt	22	25	28	30	105	128	82.03% (High)
		Control	15	16	15	15	61	128	47.66% (Low)

The behavioral data in Table 6 show that the technology-mediated environment generated higher levels of student engagement. The experimental class achieved a composite participation rate of 82.03% (High Engagement), markedly higher than the control class's rate of 47.66% (Low Engagement). The most pronounced differences were observed in Task Persistence (TP) and Phonetic/Oral Initiative (POI), where the experimental class peaked at 93.75% and 87.50%, respectively. This demonstrates that the gamified short video units and real-time AI microphone feedback within the Cake application effectively maintained student focus and encouraged continuous verbal production. In contrast, the traditional,

teacher-centered control group recorded lower enthusiasm (31.25%) and passive oral engagement (46.88%), with students frequently remaining silent during textbook exercises. During the implementation of the Cake application, students demonstrated active participation throughout the speaking learning activities. Most students showed enthusiasm when practicing pronunciation, repeating conversation exercises, and participating in speaking discussions. The observation results showed that students became more confident in expressing ideas in English after several treatment meetings. Students also appeared more motivated because the application provided authentic conversation videos and interactive speaking exercises which encouraged active participation. Compared to conventional classroom activities, the use of the Cake application created a more enjoyable and student-centered learning atmosphere. Students participated more actively during classroom interaction and speaking performances. The findings indicate that the implementation of the Cake application not only improved students' speaking scores but also increased students' classroom participation and learning motivation during speaking activities.

Discussion

The primary objective of this empirical study was to critically evaluate the causal impact of the Cake application on the oral speaking proficiency and classroom participation of ninth-grade students at SMP Negeri 1 Kabupaten Sorong. The inferential statistical results definitively answered the first research question, demonstrating that the integration of this Mobile-Assisted Language Learning (MALL) tool yielded a profound and statistically significant improvement in students' comprehensive speaking skills. This is evidenced by the dramatic post-test mean divergence, where the experimental cohort ($\mu = 83.75$) substantially outpaced the traditional, textbook-bound control group ($\mu = 70.88$). The rejection of the null hypothesis ($t(62) = 6.171$, $p < .001$) with a remarkably large mean difference of 12.88 points highlights that digital intervention in speaking pedagogy is not merely statistically valid but educationally transformative. Furthermore, the calculation of an exceptionally high effect size ($d = 1.57$, $g = 1.55$, $\backslash \eta^2 = 0.381$) reveals that nearly 38.1% of the variance in the post-test oral performance was directly determined by the technological treatment. These mathematical indicators prove that systematic, mobile-mediated oral drilling can radically shift student performance from a baseline of limited, fragmented speech to an advanced state of structured and authentic oral production.

From a theoretical standpoint, these empirical breakthroughs validate the symbiotic intersection of Constructivism, Krashen's Input Hypothesis, and Vygotsky's Sociocultural framework within a digital architecture. According to Constructivist principles, language acquisition occurs when learners actively manipulate symbols and independently construct

cognitive schemas (Rowland et al., 2026). The Cake application operationalizes this by forcing students out of passive listening into active verbal production through automated real-time loops. This technological mechanism functions as an immediate digital scaffold, directly corresponding to Vygotsky's *Zone of Proximal Development* (ZPD). When the application's AI feedback engine instantly evaluates a student's microphone input against native standards, it creates a tailored, low-stakes cognitive bridge that helps the learner close the gap between their current phonetic capacity and their potential competence. Furthermore, the app's curated native-speaker short videos successfully deliver what Krashen terms comprehensible input (i+1). By compressing real-world expressions into digestible, context-rich segments, the application lowers the affective filter—minimizing language anxiety and boosting self-efficacy—allowing linguistic input to successfully transform into permanent cognitive intake.

When juxtaposed with previous literature, the findings of this study both substantiate and significantly expand the existing boundaries of MALL research. While prior investigations by Chaniago (2022) only went as far as documenting positive student perceptions via subjective questionnaires, this study provides rigorous, objective empirical data demonstrating actual oral score gains across five core language components. Similarly, whereas (Hermawati et al. (2023) utilized a restricted pre-experimental design that focused exclusively on pronunciation, this quasi-experimental investigation captured a comprehensive picture of speaking ability, proving that mobile interaction simultaneously enhances grammatical accuracy, lexical range, listening comprehension, and conversational fluency. Most importantly, this study challenges the contextual limitations of Oktaviani & Kusumajati (2022). Their classroom action research was bounded within technologically saturated urban centers with strong institutional readiness. By proving that the Cake application can drive massive oral proficiency gains even within a frontier educational setting with limited infrastructure, this study establishes the cross-contextual adaptability and universal power of mobile-assisted instruction. The pedagogical success of this intervention is particularly remarkable when evaluated against the unique sosiolinguistik and geopolitical realities of Southwest Papua. As documented in the preliminary classroom observations, students in this region battle severe cross-linguistic interference, as English represents their third or fourth language (L3/L4) after their respective indigenous Papuan vernaculars and regional Papuan Malay (Yelvia, 2023). In a traditional classroom, this complex multilingualism typically induces extreme foreign language anxiety, causing the majority of the class to remain completely passive to avoid public errors. The classroom observation data (Table 6) reveals how the mobile application successfully disrupted this defensive silence. The experimental class achieved a striking composite participation rate of 82.03% (High

Engagement) compared to the control group's stagnant 47.66% (Low Engagement). The dramatic spike in Phonetic/Oral Initiative (POI) to 87.50% proves that shifting the conversational partner from a intimidating classroom space to a private, non-judgmental smartphone interface effectively neutralized the students' affective filters, enabling them to practice oral production with newfound confidence. These findings offer crucial pedagogical implications for EFL educators and policymakers striving to modernize language instruction in developing remote regions. The traditional, teacher-centered paradigm—dominated by monotonous textbook recitation and rigid grammar-translation drills—has historically failed to produce functional oral communicators in Indonesian public schools (Rusdin & Purwati, 2023). The results of this study suggest that the strategic integration of gamified mobile platforms can effectively decentralize the classroom, transforming the teacher from an absolute authority figure into an active facilitator of technological scaffolding. The high percentage of Task Persistence (TP) at 93.75% in the experimental group indicates that digital interactive media can naturally sustain student motivation without forcing compliance. Therefore, educational authorities in peripheral regions should actively invest in integrating mobile technology into formal curricula. Smartphones, which are already widely owned by rural students for recreation, can be repositioned as powerful learning devices to bridge the educational quality gap between urban centers and rural schools. Despite its rigorous design and strong statistical outcomes, several limitations must be acknowledged to guide future research. First, due to institutional constraints and rigid administrative policies at SMP Negeri 1 Kabupaten Sorong, the random assignment of individual participants was impossible, requiring the use of non-equivalent intact classes. Although pre-test statistics confirmed baseline homogeneity, the quasi-experimental nature of the design means that unidentified confounding variables, such as varying levels of personal motivation or outside media exposure, could not be completely controlled. Second, the intervention was restricted to a brief instructional timeline of four sessions, which prevents this study from assessing the long-term retention of the oral gains or the permanent stabilization of behavioral engagement. Third, the scope of this study was strictly quantitative, relying on performance scores and frequency-based observation metrics. Future researchers should implement longitudinal, mixed-methods designs that incorporate qualitative depth—such as semi-structured interviews and thematic diaries—to capture the nuanced, lived experiences of indigenous Papuan students as they navigate the digital shift in language learning.

Conclusions

This study demonstrated that the implementation of the Cake application positively influenced students' speaking ability at SMP Negeri 1 Kabupaten Sorong. The findings indicate that students who learned through the application showed better speaking

performance and more active participation during speaking learning activities than those who received conventional instruction. The interactive features of the application provided students with greater opportunities to practice speaking, improve pronunciation, expand vocabulary, and develop confidence in using English. The study contributes to the growing body of MALL research by providing empirical evidence from a Papuan educational context, which has received limited attention in previous studies. The findings suggest that integrating mobile learning applications into English instruction can support the development of speaking skills and create a more engaging learning environment for students. Practically, the Cake application may serve as an alternative learning medium for English teachers seeking to enhance students' speaking ability and participation in classroom activities. Future research is recommended to investigate the long-term effects of the application on speaking development, involve larger and more diverse samples, and compare the effectiveness of Cake with other mobile-assisted language learning platforms in different educational contexts.

References

Asratie, M., Wale, B. D., & Aylet, Y. T. (2023). Effects of using educational technology tools to enhance EFL students' speaking performance. *Education and Information Technologies*, 1–21. https://doi.org/10.1007/s10639-022-11562-y

Brown, H. D., & Lee, H. (2025). Principles of Language Learning and Teaching. In *Principles of Language Learning and Teaching*. https://doi.org/10.4324/9781003494294

Chaniago, A. D. (2022). the Effect of Cake Usage As Mobile Assisted Language Learning on Students' English Speaking Skill in Smp Hang Tuah 1 Belawan. *ETERNAL (English, Teaching, Learning, and Research Journal)*, *8*(1), 131–144. https://doi.org/10.24252/eternal.v81.2022.a9

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. In *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Inc. https://doi.org/10.4324/9780203771587

Creswell, J. W. (2014). Conducting, and evaluating quantitative and qualitative research. In *Educational research*. Pearson Education.

Field, A. (n.d.). *Discovering_Statistics_Using_IBM_SPSS_St* (5th ed.). SAGE Publications, 2018.

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, *10*(2), 486–489. https://doi.org/10.5812/ijem.3505

Glass, G. V. (1966). Testing Homogeneity of Variances. *American Educational Research Journal*, *3*(3), 187–190. https://doi.org/10.3102/00028312003003187

Hermawati, S., Bte Abdul, N., & AM, S. A. (2023). Effectiveness of Using Cake Application on Students' Pronunciation Skills At Smp Unismuh Makassar. *English Language Teaching Methodology*, *3*(3), 309–316. https://doi.org/10.56983/eltm.v3i3.536

Indah Sri Redjeki, & R. Muhajir. (2022). Supporting Students' Self-Directed Learning in EFL Speaking Classroom by Using Cake Application. *Journal of English Language Learning*, *6*(2), 212–219. https://doi.org/10.31949/jell.v6i2.3421

Kim, T. K. (2015). T test as a parametric statistic. *Korean Journal of Anesthesiology*, *68*(6), 540–546. https://doi.org/10.4097/kjae.2015.68.6.540

Mahu, F., Hasbullah, & Suharmoko. (2024). Improving Students' Speaking Skills With Greeting Materials Using Ice Breaking Technique At The 7th Class B Smp Quba Sorong City. *Transformasi : Jurnal Kepemimpinan & Pendidikan Islam*, *7*(2), 74–91. https://doi.org/10.47945/transformasi.v7i2.1551

Metruk, R. (2024). Mobile-assisted language learning and pronunciation instruction: A systematic literature review. *Education and Information Technologies*, *29*(13), 16255–16282. https://doi.org/10.1007/s10639-024-12453-0

Mokhamad Azis Aji Abdilah. (2024). The Use of Cake Mobile Application in Learning English Speaking Skills for Young Learners. *Journal of Education and Culture*, *4*(3), 79–85. https://doi.org/10.58707/jec.v4i3.980

Nayakam, X. A. M., & Aksharagovind, U. S. (2022). Critical Study on Enchaning Speaking Skills. *International Journal of Health Sciences*, 2517–2523. https://doi.org/10.53730/ijhs.v6ns8.12730

Oktaviani, T., & Kusumajati, W. K. (2022). Honing the Speaking Skill with Cake Application in Classroom Action Research. *Proceeding of International Conference on Education-01*, 85–92. https://doi.org/10.37640/ice.01.107

Oppong, F. B., & Agbedra, S. Y. (2016). Assessing Univariate and Multivariate Normality , A Guide For Non - Statisticians. *Mathematical Theory and Modeling*, *6*(2), 26–33. www.iiste.org/Journals/index.php/MTM/article/download/28571/29333

Paramita, I. G. A. D., Ardika, I. W. D., Setyono, E. Y., Yuliantini, N. N., & Suciani, N. K. (2022). The use of cake application to improve speaking ability. *Journal of Applied Studies in Language*, *6*(2), 214–220. https://doi.org/10.31940/jasl.v6i2.345

Pardede, E., Sitanggang, M. M., & Saragih, S. S. (2022). Speaking Problem of Junior High School Students. *Journal of Humanities, Social Sciences and Business (Jhssb)*, *1*(2), 43–48. https://doi.org/10.55047/jhssb.v1i2.98

Pauzan, P. (2024). Theory in Second Language Acquisition (Recognition of Concepts Toward Krashen's Second Language Acquisition Theory for Five Main Hypotheses). *Journal on Education*, *6*(4), 20876–20888. https://doi.org/10.31004/joe.v6i4.6210

Qasim, A. Y. (2021). A Study of the Factors Affecting the Learning of English Speaking Skills.

Academic Journal of Nawroz University, *10*(1), 193. https://doi.org/10.25007/ajnu.v10n1a1078

Rowland, C. F., Westermann, G., Theakston, A. L., Pine, J. M., Monaghan, P., & Lieven, E. V. M. (2026). Constructing language: a framework for explaining acquisition. *Trends in Cognitive Sciences*, *30*(1), 26–39. https://doi.org/10.1016/j.tics.2025.05.015

Rusdin, R., & Purwati, D. (2023). Speaking Up: A Comprehensive Investigation of EFL Secondary Students' Speaking Skill Problems in Indonesia's Islamic School Context. *Elsya : Journal of English Language Studies*, *5*(2), 236–249. https://doi.org/10.31849/elsya.v5i2.11911

Sari, I. P., Muzammil, L., & Mafulah, S. (2025). *Prodi Magister Pendidikan Bahasa Inggris, Universitas PGRI Kanjuruhan Malang Revised: 18*. *10*(1), 98–107.

Slavuj, V. (2023). Using mobile applications for language learning as part of language classes. *Zbornik Veleučilišta u Rijeci*, *11*(1), 213–234. https://doi.org/10.31784/zvr.11.1.12

Smith, A. M. (2013). Research Methods in Education. In *Nurse Education in Practice* (Vol. 13, Issue 3, p. e3). https://doi.org/10.1016/j.nepr.2012.06.001

Thomas, M. P. (2026). *Mobile-Assisted Language Learning (MALL) and English Proficiency Development among ESL College Students : A Quasi-Experimental Mixed-Methods Study*. 15–22.

Tzuriel, D. (2021). The Socio-Cultural Theory of Vygotsky. *Mediated Learning and Cognitive Modifiability*, 53–66. https://doi.org/10.1007/978-3-030-75692-5_3

Yassin, H., Nambi, R. S. K., & Najjemba, J. (2024). Rethinking English Language Pedagogy: An Exploration of the gaps in Traditional Language Teaching Methods. *International Journal For Multidisciplinary Research*, *6*(6), 1–10. https://doi.org/10.36948/ijfmr.2024.v06i06.32444

Yelvia, E. (2023). *The Challenges of English Language Teaching in Multilingual Classroom*. *03*(December), 84–88. https://proceeding-icolp.fbs.unp.ac.id/index.php/icolp

Zhang, M., Yi, N., & Zhou, D. (2023). The Effects of Task Repetition Schedules on L2 Fluency Enhancement. *Languages*, *8*(4), 1–16. https://doi.org/10.3390/languages8040252