IGNATIAN
INTERNATIONAL JOURNAL FOR MULTIDISCIPLINARY RESEARCH
www.icceph.com
ISSN 2984-9942

THE EFFECTS OF ROLEPLAY ACTIVITIES ON STUDENTS' ENGLISH-SPEAKING SKILLS

Kiervey P. Manginsay [1], Rolly D. Pampilo, Jr. [1], Miechaela A. Pactol [1], Frietz G. Oyda [1], Hannah Marie T. Mayormita [1], Jose Calasanz Julysses D. Tabia [2]

[1] *Teacher Education, St. Vincent's College Incorporated, Dipolog City, Philippines*
[2] *Graduate School, St. Vincent's College Incorporated, Dipolog City, Philippines*

https://doi.org/10.5281/zenodo.18618879

ABSTRACT

This study examined the effects of spontaneous roleplay activities on the English-speaking skills of Grade 12 students at St. Vincent's College Incorporated – Basic Education Campus during the school year 2025–2026. Anchored on Vygotsky's Sociocultural Theory and the concept of Communicative Competence, the research explored how interaction, scaffolding, and authentic communication contribute to learners' fluency, pronunciation, vocabulary, and confidence. A quasi-experimental crossover design was utilized involving two intact classes that alternately experienced roleplay and non-roleplay conditions over two days. Data were gathered through a standardized rubric and analyzed using weighted mean, standard deviation, independent sample t-test, and paired sample t-test. Results on Day 1 showed that the control group and experimental group achieved mean scores of 90.8 and 91.3 in fluency, 90.2 and 91.7 in pronunciation, 90.5 and 91.4 in vocabulary, and 90.8 and 92.1 in confidence, respectively. After the crossover on Day 2, both groups recorded similar outcomes across all components, with only minor variations. Findings indicate that spontaneous roleplay activities benefited lower-performing students by improving measurable speaking skills, while advanced learners demonstrated stable performance due to a ceiling effect. Overall, the study concludes that roleplay is an effective strategy for enhancing speaking proficiency and strengthening communicative confidence in senior high school students.

Keywords: *Spontaneous Roleplay, Fluency, Pronunciation, Confidence, Vocabulary, Quasi-experimental Crossover*

INTRODUCTION

Speaking is a fundamental macro skill in second language learning, as it enables learners to express ideas, emotions, and intentions through real-time oral communication. In academic and professional contexts, English-speaking proficiency is essential for participation in discussions, collaboration, and engagement in global interactions (Ghafar & Raheem, 2023). For Grade 12 students, who are preparing for higher education and employment, the ability to speak English fluently and confidently is particularly critical in navigating an increasingly English-dominated global environment.

Despite prolonged exposure to English through formal instruction and media, many senior high school learners continue to struggle with speaking performance. Difficulties in fluency, pronunciation, and vocabulary remain prevalent, often compounded by language anxiety and low confidence (MacIntyre & Gregersen, 2022; Khasbani & Seli, 2021). Traditional classroom instruction, which frequently emphasizes written tasks and controlled drills, often fails to provide sufficient opportunities for authentic and spontaneous oral communication (Rao & Zakaria, 2025). As a result, learners may possess theoretical knowledge of English but lack the confidence and competence to use it effectively in real-life interactions.

Research has shown that interactive and learner-centered strategies are effective in addressing these challenges. Roleplay activities, in particular, allow students to simulate real-life communicative situations, encouraging meaningful language use, peer interaction, and active engagement. Studies have demonstrated that roleplay improves speaking fluency, pronunciation, vocabulary, and overall communicative competence (Dwiyanti & Lolita, 2023; Idham et al., 2022; Hussam et al., 2025). In addition to linguistic gains, roleplay has been found to reduce speaking anxiety and enhance learners' confidence by creating a supportive and low-risk environment for oral practice (Togimin & Jaafar, 2020; Dörnyei & Ryan, 2020).

However, while existing studies affirm the effectiveness of roleplay in general, limited attention has been given to spontaneous or unscripted roleplay, particularly at the senior high school level. Most classroom roleplay activities rely on memorized scripts, which may restrict learners' ability to think, respond, and negotiate meaning in real time. Spontaneous roleplay, on the other hand, requires learners to generate language immediately, adapt to unpredictable responses, and apply linguistic knowledge authentically. Previous studies suggest that unscripted roleplay enhances fluency, creativity, vocabulary use, and learner autonomy, while also fostering confidence and motivation (Aini et al., 2020; Salsabila & Megawati, 2024). Nonetheless, empirical research examining its specific effects on English-speaking skills and confidence among Grade 12 students remains scarce.

Given the increasing demand for communicative competence in academic and professional settings, and the persistent speaking difficulties observed among senior high school learners, this study is both timely and necessary. It addresses an identified research gap by examining the effects of spontaneous roleplay activities on students'

English-speaking skills—specifically fluency, pronunciation, and vocabulary—as well as their confidence in speaking English. Anchored in the principles of communicative language teaching, this study proposes spontaneous roleplay as an instructional innovation that promotes authentic communication, learner engagement, and affective development. The findings of this study aim to contribute to English language teaching practices by providing empirical evidence on the effectiveness of spontaneous roleplay as a pedagogical tool for enhancing speaking proficiency and confidence among Grade 12 students.

Research Questions

This study aimed to assess the students' English-speaking skills and their confidence in English speaking through the integration of spontaneous roleplay. specifically, the study sought to answer the following questions:

1. What is the level of English-speaking skills of the students with and without spontaneous roleplay in terms of:
   1.1 fluency;
   1.2 pronunciation; and
   1.3 vocabulary?
2. What is the level of the students' confidence in English speaking with and without spontaneous roleplay?
3. Is there a significant difference in the level of English-speaking skills of the students between the groups with and without spontaneous roleplay in terms of:
   3.1 fluency;
   3.2 pronunciation; and
   3.3 vocabulary?
4. Is there a significant difference in the level of the students' confidence in English speaking between groups with and without spontaneous roleplay?
5. Is there a significant difference on the students' English-speaking skills with and without the implementation of spontaneous roleplay?
6. Is there a significant difference on the students' confidence with and without the implementation of spontaneous roleplay?

METHODOLOGY

This study employed a quasi-experimental crossover design to investigate the effects of spontaneous roleplay activities on the English-speaking skills and the confidence in English speaking of Grade 12 students. This specific design was purposefully selected because the research involved two pre-existing, intact classes, which made the random assignment of individual students impractical. The crossover approach was particularly advantageous in such a setting as it addressed a primary challenge of quasi-experimental research: the potential for baseline inequality between the two groups. By ensuring that each class participated in both the experimental and control conditions, the design effectively minimizes the influence of confounding

variables, such as pre-existing differences in baseline language proficiency, learner motivation, or classroom dynamics.

The design allowed within-subject comparison, such that the researchers could compare the performance scores of students when they experienced the roleplay intervention against their own scores when receiving standard instruction. This direct comparison provided a more precise and rigorous analysis of the intervention's specific impact by isolating the effect of the instructional strategy from the inherent differences between the two classes.

The research was conducted at St. Vincent's College Incorporated – Basic Education Campus (SVCI – BEC), located in Padre Ramon Street, Estaka, Dipolog City, Zamboanga del Norte, during the school year 2025-2026. This institution provides education from Kindergarten to Senior High School and is known for its commitment to academic excellence and values-based instruction.

SVCI-BEC is considered the most suitable environment for this study because it offers a diverse student population with varying levels of English proficiency, especially in the Senior High School Department. English is used as the primary medium of instruction in most subjects, and the institution encourages students to actively participate in oral communication tasks such as classroom reporting, oral recitation, and interactive group work. These practices aligned well with the objectives of the study, which focused on enhancing students' English-speaking skills through communicative activities like roleplay.

In recent classroom observations and feedback from teachers, many students at SVCI-BEC demonstrated a foundational understanding of grammar and vocabulary but often struggle with fluency, confidence, and spontaneous use of English. This indicated a need for more interactive and performance-based strategies, such as roleplay, to improve students' communicative competence. Therefore, conducting the study in this academic setting provided meaningful insights into how roleplay can be effectively integrated into the curriculum to support the development of speaking.

The respondents of this study consisted on Grade 12 STEM students enrolled at St. Vincent's College Incorporated – Basic Education Campus during the school year 2025-2026. These students, regardless of their academic strand, were selected for their active participation in English language instruction and communicative activeness across English subjects that involved oral communication.

The study focused on Grade 12 students as they are in the final phase of senior high school and are expected to have developed a foundational proficiency in the English language. Their coursework included opportunities to apply speaking skills in academic and practical contexts, making them appropriate participants for examining the effects of roleplay activities.

A purposive sampling technique was employed to ensure that only students currently enrolled in English language courses are included in the study. Specifically employing a total population sampling approach. This method was deemed the most appropriate as the researchers did not select a sample from a larger pool. Instead, the study included the entire target population available, which consisted of the existing two sections of students enrolled in the English for Academic and Professional Purposes (EAPP) subject. Because the study aimed to assess an intervention within this specific academic context, these two classes were the only relevant and eligible groups for participation.

The use of these two intact classes was a key component of the study's quasi-experimental crossover design, which was designed for real-world educational settings where random assignment is not feasible. The validity of this participant group was strongly supported by their relevance and homogeneity. As all participants were Grade 12 STEM students in the same subject, they were assumed to share similar academic foundation, course objectives, and language learning needs. This homogeneity was an advantage, as it reduced potential extraneous variables related to differing academic tracks of subjects, thereby allowing a more focused and clearer analysis of the intervention's direct impact on their speaking skills.

The data for this study were gathered using a set of researcher-made standardized rubrics specifically curated to assess students' English-speaking skills and their confidence in English speaking. A separate, detailed rubrics were utilized for each of the four key indicators: fluency, pronunciation, vocabulary, and confidence. Each rubric provided specific criteria and performance-level descriptors (Advanced, Proficient, Developing, and Beginning), which allowed a qualified English teacher evaluator to conduct a systematic and objective evaluation of each component of the students' speaking abilities.

To ensure the validity of these instruments, a thorough expert validation process was conducted. The set of rubrics was submitted to a panel of professionals and experts in the field of English language teaching. These experts were tasked with meticulously reviewing the rubrics to evaluate their content validity. Their evaluation focused on the rubrics' comprehensiveness, clarity, and relevance as well as their clear agreement with the study's goals.

The feedback, suggestions, and recommendations provided by the expert validators were carefully considered and integrated to revise and refine the rubrics.

For this quasi-experimental research, the data collection process did not utilize any questionnaire. Instead, the researchers used a standardized assessment rubrics based on the K-12 Basic Education Curriculum Guide, which have been validated by English instructors of St. Vincent's College Incorporated.

The implementation of the study was structured across two consecutive instructional days. On the first intact class (Group A) received the standard instructional

method, acting the control group. Simultaneously, the second class (Group B) was exposed to spontaneous roleplay intervention, thereby serving as the experimental group. On the second day, the conditions were "crossed over": Group A was then exposed to spontaneous roleplay intervention, thereby serving as the experimental group while Group B, now the control group, reverted to the standard instruction. This procedure ensured that all participants experienced both instructional methods, allowing each group to serve as its own control.

The effectiveness of the intervention was measured by assessing students' speaking skills, specifically their fluency, pronunciation, vocabulary, and their confidence in English speaking, at the conclusion of each day's session. The data were collected using a validated rubric.

A formal letter requesting permission to conduct the study and administer the assessment rubric was submitted to the school principal of St. Vincent's College Incorporated – Basic Education Campus in Dipolog City. Upon approval, the researchers coordinated with the designated teacher who served as the sole evaluator of the study.

The selected teacher then validated standardized rubric – based on the K-12 Basic Education Curriculum Guide and approved by the English instructors of St. Vincent's College Incorporated – to evaluate the English-speaking skills of the Grade 12 students. Assessments were conducted on a two-day timeframe of instructional format. On Day 1, Group A (control) received the standard instructional method, while the second class, Group B (experimental) was exposed to spontaneous roleplay intervention. At the conclusion of this session, each participant's speaking performance was individually assessed by the teacher-evaluator using the validated standardized rubrics provided by the researchers. Scores for fluency, pronunciation, vocabulary, and confidence were recorded for every student. On the following day, the conditions were reversed: Group A, now experimental group, was exposed to spontaneous roleplay while Group B, now the control group, received the standard instructional method. Following this session, the assessment was repeated, with every participant's speaking skill and confidence rated using the same rubrics.

The scores from the rubrics were meticulously tallied and collated to ensure accuracy and prepared the data for analysis. Throughout this process, ethical protocols were strictly maintained. The participants' data were anonymized using codes to protect their identities. The collected scores were handled with confidentiality by the researchers and were secured to be used solely for the statistical analysis required by the study.

To analyze the effects of spontaneous roleplay on Grade 12 students' English-speaking skills and their confidence in English speaking, both descriptive and inferential statistical tools were employed:

To describe the performance of the entire group and to prepare the data for the inferential statistical test, the arithmetic mean was calculated. The mean was computed for each indicator by summing the individual scores of all students and dividing it by the

total number of students. This calculation was done separately for the control and experimental conditions. The resulting value represents the overall level of the entire class for that specific indicator.

In evaluating the effectiveness of the roleplay intervention, the standard deviation served as a critical analytical tool that provided a much deeper insight than the mean scores alone. It described how spread-out students' speaking skills were at sessions with and without roleplay implementation. By calculating the standard deviation without the intervention, we can understand the initial spread of abilities, and by comparing it to the results with the implementation's standard deviation, we can determine if the spontaneous roleplay had an equalizing effect, which would result in a decreased standard deviation, or if they widened the skill gap, an increased standard deviation will be evident. More importantly, the standard deviation of the difference scores, which will be calculated by subtracting control scores from experimental scores, is fundamental to achieving the study's goal. This specific measure quantifies the consistency of the intervention's effects across all participants. A small standard deviation of the differences would indicate that most students improved by a similar amount, suggesting a reliable and predictable pedagogical effect. Conversely, a large standard deviation would imply the intervention's impact was erratic, with some students benefiting far more than others.

Given the study's crossover design, the analysis was conducted in two distinct phases rather than as a single repeated-measures analysis. First, an independent samples t-test was performed on the Day 1 scores, comparing the performance of the Control Group (A) against the Experimental Group (B) to determine the intervention's initial effect. Second, a separate independent sample t-test was conducted on the Day 2 scores, comparing the Experimental Group (A) against the Control Group (B) after the crossover. This two-phase approach allowed the analysis of the intervention's effectiveness at both time points and it provided insight into potential carryover effects. Prior to conducting the t-test, the data were screened to ensure that all statistical assumptions were met. The Shapiro-Wilk test was used to assess the normality of the data distribution for each group, and Levene's Test for equality of variances was applied to confirm the homogeneity of variances between the groups. The appropriate t-test results ("Equal variances assumed" or "Equal variances not assumed") were interpreted based on the outcome of Levene's test. For all inferential analyses, the alpha level was set at .05, with p-values less than or equal to this threshold is considered to be statistically significant.

To determine whether the roleplay intervention led to a significant difference in the students' English-speaking skills and confidence in English speaking, a paired sample t-test is the intended statistical analysis. This test is appropriate for this study's quasi-experimental crossover design, as it is specifically designed to compare the mean scores of the same group of participants at two different points in time. However, the validity of the paired-samples t-test is contingent upon the satisfaction of several key assumptions. Therefore, before analysis, a thorough diagnostic check was conducted on the data. While the assumptions that the dependent variable is continuous and that observations

are independent are met by the study's design, particular scrutiny was applied to the distributional assumptions.

This study focused on examining the effects of spontaneous roleplay on Grade 12 students' English-speaking skills and confidence in English speaking. The intervention was implemented at St. Vincent's College Incorporated – Basic Education Campus. The participants are Grade 12 STEM students enrolled in the subject English for Academic and Professional Purposes (EAPP), during the school year 2025-2026. The study utilized a quasi-experimental crossover design. The study measured speaking skills of students from two purposively picked classes, which means the findings are context-specific to these participants and are not intended for broad generalization. Moreover, a teacher-evaluator is utilized in this study – the same teacher of the course for both classes. The teacher was selected because of their experience with the students within the school year. This ensures that the teacher already has a foundational knowledge of the speaking skills of the students involved. Furthermore, the teacher-evaluator utilized a standardized rubric based on the K-12 curriculum containing criteria for the following: fluency, pronunciation, vocabulary, and confidence. The study was also deliberately narrowed its focus to the four aforementioned components, excluding other aspects of communicative competence like grammatical accuracy.

The scope is limited to performance in English within classroom-based, academic, and semi-authentic scenarios designed by the teacher. Other language skills, such as listening comprehension, reading, and writing were not measured. The study did not account for external factors such as students' prior English proficiency outside the school setting, language exposure in the home or community, or individual learning preferences. The variables including the confidence variable, were measured through a validated rubric assessed by the teacher. while this method avoids self-perception bias from the participants, the results may be subject to observer bias, as they depend on the teacher's professional judgement and interpretations.

RESULTS

Table 1 presents the descriptive statistics for the English-speaking skills of the students in both the control and experimental groups across three categories: fluency, pronunciation, and vocabulary, on Day 1.

For fluency, the experimental exhibited a higher mean (91.3) compared to the control group (90.8). The experimental group had a standard deviation of 2.25, while the control group had a slightly smaller standard deviation at 2.04. This indicated that while the experimental group had a slightly greater range in fluency scores, it still maintained a relatively high level of consistency.

For pronunciation, the experimental outperformed the control group, with means of 91.7 and 90.2, respectively. The experimental group also had a smaller standard

deviation with 1.67, indicating less variability in their scores, while the control group had a standard deviation of 2.56.

In terms of vocabulary, the mean score for both groups were similar, with the experimental group scoring 91.4 and the control group scoring 90.5. the standard deviation for the experimental group was 2.58, which is slightly higher than that of the control group with 2.12, indicating that there was a slightly more variability in the experimental group's vocabulary scores.

Although all both groups exhibit an advanced level of English-speaking skills, the experimental group had slightly higher mean scores and less variability across all three categories. This suggests that the spontaneous roleplay might have been slightly more effective in improving students' English-speaking skills compared to the control group. However, the differences were small, indicating that both groups performed similarly in overall English-speaking skills.

Table 1

The level of English-speaking skills of the students with and without spontaneous roleplay implementation in terms of fluency, pronunciation, and vocabulary on Day 1

	DAY 1	Mean	SD	Interpretation
Fluency	**Control**	90.8	2.04	Advanced
	Experimental	91.3	2.25	Advanced
Pronunciation	**Control**	90.2	2.56	Advanced
	Experimental	91.7	1.67	Advanced
Vocabulary	**Control**	90.5	2.12	Advanced
	Experimental	91.4	2.58	Advanced

Table 2 presents the descriptive statistics for the English-speaking skills of students in both the control and experimental groups on Day 2, across three categories: fluency, pronunciation, and vocabulary.

For each category, the mean scores of both groups were very close to each other, with the control group having slightly higher mean scores compared to the experimental group. In pronunciation, the control group had a mean score of 90.4 while the experimental group had a 90.2. Similar results were observed for vocabulary (both groups had 90.4) and fluency (each had 90.2). These findings suggest that there were minimal differences in the overall performance of the students between two groups on Day 2.

The standard deviations were similar for both groups across three categories. The control group had slightly lower standard deviations (2.25 for fluency, 2.02 for pronunciation, and 2.57 for vocabulary) compared to the experimental group (2.43 for fluency, 2.10 for pronunciation, and 2.71 for vocabulary). This indicated that the control group exhibited slightly less variation in their scores, meaning their performance was more consistent across students. These results indicate that by Day 2, both the control

and experimental groups had very similar levels of English-speaking skills, suggesting that any initial differences between the groups might have narrowed by this point.

Table 2

The level of English-speaking skills of the students with and without spontaneous roleplay implementation in terms of fluency, pronunciation, and vocabulary on Day 2

	DAY 2	Mean	SD	Interpretation
Fluency	**Control**	90.2	2.25	Advanced
	Experimental	90.2	2.43	Advanced
Pronunciation	**Control**	90.4	2.02	Advanced
	Experimental	90.2	2.10	Advanced
Vocabulary	**Control**	90.4	2.57	Advanced
	Experimental	90.4	2.71	Advanced

Table 3 presents the descriptive statistics for the confidence in English speaking of students in both control and experimental groups on Day 1. The experimental group had a higher mean score (92.1) compared to the control group (90.8), suggesting that students in the experimental group felt more confident in their English-speaking abilities on average. Additionally, the experimental group had a slightly lower standard deviation (2.10), indicating that the confidence levels in this group were more consistent compared to the control group, which had a standard deviation of 2.30.

Table 3

The level of the students' confidence in English speaking with and without spontaneous roleplay on Day 1

	DAY 1	Mean	SD	Interpretation
Confidence	**Control**	90.8	2.30	Advanced
	Experimental	92.1	2.10	Advanced

Table 4 shows the confidence levels in English speaking for students in both control and experimental groups on Day 2. Both groups exhibited similar mean scores, with the control group scoring a mean of 91.1 while the experimental group is at 91.0, indicating comparable overall confidence levels. The standard deviation is slightly higher for the experimental group (2.68) compared to the control group (2.10), suggesting that there is more variability in confidence scores within the experimental group.

Table 4

The level of the students' confidence in English speaking with and without spontaneous roleplay on Day 2

	DAY 2	Mean	SD	Interpretation
Confidence	**Control**	91.1	2.10	Advanced
	Experimental	91.0	2.68	Advanced

Table 5 presents the results of the Independent Sample t-test comparing the English-speaking skills between the experimental and control groups on Day 1, focusing on fluency, pronunciation, and vocabulary. Prior to conducting the t-test, the assumption of normality and homogeneity of variants were tested.

For fluency, the assumption of normality was met (Shapiro-Wilk p = 0.569), and Levene's test showed no significant difference in variance (p = 0.359). The student's t-test statistic was -0.923 and a p-value of 0.359, indicating no significant difference.

For pronunciation, the assumptions were not met: the Shapiro-Wilk test indicated normality (p = 0.703), but the Levene's test for homogeneity of variances showed significant difference (p = 0.045). Thus, Welch's t-test is presented with the t-statistic (-2.775), and p-value of 0.007, indicating a significant difference between experimental and control group.

For vocabulary, the assumption of normality was violated (Shapiro-Wilk p = 0.002), but Levene's test for homogeneity showed significant difference (p = 0.418). The Mann-Whitney U test indicated significant difference at p= 0.029, respectively.

Table 5

Independent Samples t-test for English-speaking skills on Day 1 (A-Control versus B-Experimental)

		Statistic	p	Interpretation
Fluency	Student's t	-0.923	0.359	Not Significant
Pronunciation	Welch's t	-2.775	0.007	Significant
Vocabulary	Mann-Whitney U	400	0.029	Significant

Table 6 presents the results of the independent sample t-test conducted to compare the English-speaking skills (fluency, pronunciation, and vocabulary) between the experimental and control groups on Day 2. The table reports the student's t-test and Mann-Whitney U test statistics, along with the corresponding p-values for each skill.

For fluency, the student's t-test produced a statistic of 0.11220 with a p-value of 0.911, indicating no significant difference in fluency scores between the experimental and control groups ($p > 0.05$).

For pronunciation, the student's t-test yielded a statistic of 0.43491 with a p-value of 0.665, indicating no significant difference between the experimental and control groups in terms of pronunciation skills ($p > 0.05$). For vocabulary, the Mann-Whitney U test yielded a U statistic of 572 with a p-value of 0.941, also suggests no significant difference in vocabulary skills between the groups.

Table 6

Independent Samples t-test for English-speaking skills on Day 2 (A-Experimental versus B-Control)

		Statistic	p	Interpretation
Fluency	Student's t	0.11220	0.911	Not Significant
Pronunciation	Student's t	0.43491	0.665	Not Significant
Vocabulary	Mann-Whitney U	572	0.941	Not Significant

Table 7 presents the results of the Independent Sample t-test for confidence in English speaking on Day 1, comparing the control and experimental groups. Prior to conducting the t-test, the assumptions of normality and homogeneity of variances were tested.

For confidence, the assumption of normality was not met. The student's t-test statistic was -2.51 and a p-value of 0.015, indicating a significant difference between the experimental and control groups.

Table 7

Independent Sample t-test for Confidence in English speaking on Day 1(A-Control versus B-Experimental)

		Statistic	p	Interpretation
Confidence	Mann-Whitney U	325	0.006	Significant

Table 8 presents the results of the Independent Sample t-test for confidence in English speaking on Day 2, comparing the experimental and control groups. The table reports the statistics for the student's t-test, along with the corresponding p-value.

For confidence, the student's t-test yielded a statistic of 0.164 with a p-value of 0.870, indicating that there is no significant difference in confidence between the experimental and control groups ($p > 0.05$). This indicates that there is no significant difference in confidence in English speaking between the experimental and control groups on Day 2.

Table 8

Independent Sample t-test for Confidence in English speaking on Day 2 (A-Experimental versus B-Control)

		Statistic	p	Interpretation
Confidence	Student's t	0.164	0.870	Not Significant

Table 9 presents the results of the paired samples t-test for the English-speaking skills of Group A, comparing Day 1 and Day 2. The table reports the t-statistic and p-value for each skill: fluency, pronunciation, vocabulary.

For fluency, the t-test statistic was 2.3288 with a p-value of 0.026, which is statistically significant ($p < 0.05$). This suggests that fluency in English improved significantly from Day 1 to Day 2.

For pronunciation, the t-test yielded a statistic of 0.0455 with a p-value of 0.964, indicating that there is no significant difference in pronunciation between Day 1 and Day 2 ($p > 0.05$). This suggests that the Group A participants' pronunciation scores remained consistent across both days.

For vocabulary, the t-test statistic was 0.4040 with a p-value of 0.689, which also indicates no significant difference in vocabulary skills between Day 1 and Day 2 ($p > 0.05$). Therefore, vocabulary scores did not show a notable change over the two days.

Table 9

Paired Sample t-test for English-speaking skills of Group A (Day 1 versus Day 2)

		Statistic	p	Interpretation
Fluency	Student's t	2.3288	0.026	Significant
Pronunciation	Student's t	0.0455	0.964	Not Significant
Vocabulary	Student's t	0.4040	0.689	Not Significant

Table 10 presents the results of the paired samples t-test for English speaking skills of Group B on Day 1 and Day 2. The analysis aimed to examine whether there were significant differences in fluency, pronunciation, and vocabulary between the two days.

For Fluency, demonstrated a significant change, with a t-value of 41.8 and a p-value of less than 0.001. For Pronunciation, the t-value was 14.8 with a p-value of less than 0.001, indicating a statistically significant difference between Day 1 and Day 2. Vocabulary also showed a significant difference, with a t-value of 49.2 and a p-value less than 0.001.

These results suggest that all three skills, fluency, pronunciation, and vocabulary, experienced significant improvements between Day 1 and Day 2, with the changes observed being unlikely to have occurred by chance. This highlights the importance of these differences in English speaking skills across the two measurement days.

Table 10

Paired Sample t-test for English-speaking skills of Group B (Day 1 versus Day 2)

		Statistic	p	Interpretation
Fluency	Student's t	41.8	<.001	Significant
Pronunciation	Student's t	14.8	<.001	Significant
Vocabulary	Student's t	49.2	<.001	Significant

Table 11 presents the results of the paired sample t-test conducted to compare Group A's the confidence in English speaking on Day 1 and Day 2. The assumption of normality was verified using the Shapiro–Wilk test was met; thus, a paired samples t-test

was reported with a t-statistic of –0.733 with a p-value of 0.469. Since the p-value was greater than the significance level of .05, the result suggests that there was no statistically significant difference in Confidence in English speaking between Day 1 and Day 2 for Group A.

Table 11
Paired Sample t-test for Confidence in English speaking of Group A (Day 1 versus Day 2)

		Statistic	p	Interpretation
Confidence	Student's t	-0.733	0.469	There is no significant difference

Table 12 presents the results of the paired samples t-test for Confidence in English speaking skills of Group B on Day 1 and Day 2. The analysis aimed to determine whether there was a significant difference in Confidence between the two days. The t-value for Confidence was 137, and the p-value was less than 0.001.

This result indicates a statistically significant difference between Day 1 and Day 2, suggesting that the confidence scores in English speaking skills of Group B changed significantly over the course of the two days. The p-value being less than 0.001 indicates that this difference is highly unlikely to have occurred by chance, reinforcing the importance of the observed change in confidence.

Table 12
Paired Sample t-test for Confidence in English speaking of Group B (Day 1 versus Day 2)

		Statistic	p	Interpretation
Confidence	Student's t	137	<.001	Significant

DISCUSSION

The analysis of English-speaking skills and confidence levels across Day 1 and Day 2 reveals a crucial dynamic defined by the participants' high baseline proficiency. Based on the provided scoring rubric, all students in both the control and experimental groups were already performing at the "Advanced" (90–100) level on Day 1 across all metrics. The experimental group demonstrated mean scores of 91.7 in pronunciation, 91.4 in vocabulary, and 92.1 in confidence, while the control group scored 90.2, 90.5, and 90.8, respectively.

While the experimental treatment correlates with slightly higher initial scores on Day 1, these differences are variations within the "Advanced" level rather than a shift from a lower category. This finding is consistent with intervention studies that show initial gains from new strategies (Tuan, 2023), but the high proficiency of the students is a critical factor. Furthermore, the smaller standard deviations in the experimental group on Day 1

(e.g., 1.67 in pronunciation vs. 2.56 for the control) suggest the intervention may have produced a more consistent 'Advanced' performance.

From a theoretical perspective, this result aligns with Vygotsky's Sociocultural Theory (1978), which emphasizes that language learning occurs through social interaction and scaffolding within the Zone of Proximal Development (ZPD). The roleplay activities provided opportunities for collaborative learning, where peers supported each other's communicative performance through dialogue and feedback. The interpretation is reinforced by the Day 1 inferential results, where significant differences in pronunciation (p=.007), vocabulary (p=.029), and confidence (p=.006) favored the experimental group, showing that roleplay created immediate linguistic and affective benefits. This social interaction likely helped maintain students' already advanced English-speaking proficiency, but because their performance was near the ceiling, the roleplay functioned more as reinforcement than as new skill acquisition.

The alignment of scores by Day 2, with both groups scoring virtually identically and remaining firmly in the 'Advanced' category—91.1 for the control group and 91.0 for the experimental group in confidence—strongly suggests a "ceiling effect." A ceiling effect occurs when participants are already high-performing, such that there would be little to no room for a treatment to demonstrate significant improvement (Bower et al., 2024). The students, having already achieved the 'Advanced' benchmark, had limited potential for measurable gains. his interpretation is consistent with the Day 2 inferential tests, which showed no significant differences in fluency (p = .911), pronunciation (p = .665), vocabulary (p = .941), or confidence (p = .921), confirming that both groups converged at the same advanced level.

According to Swain's Output Hypothesis (1985), learners improve their fluency and accuracy through opportunities for pushed output—that is, when they are required to produce language spontaneously under communicative pressure. The results suggest that while students did not show large numerical gains, the roleplay intervention may have maintained their advanced proficiency by continually "pushing" them to use English actively and creatively. The communicative nature of roleplay allowed them to notice gaps in their vocabulary and fluency, consistent with Swain's claim that output promotes interlanguage development and self-monitoring.

Therefore, the alignment of scores on Day 2 does not necessarily imply that the intervention's effects were not sustained. Rather, it suggests that the treatment's primary, measurable impact may have been an initial boost in confidence (Hasanah, 2021) and performance consistency. However, the pre-existing high proficiency of all participants created a performance ceiling, making it difficult to detect significant, lasting differentiation between the two groups. Both groups simply started at, and returned to, the "Advanced" level.

This study's quasi-experimental crossover design revealed a complex but compelling narrative regarding the intervention's effectiveness. The findings show that the two groups, which began at significantly different skill levels, converged to a state of

statistical equivalence by Day 2. This convergence was not due to a simple equalization, but rather a powerful dual-action effect: the intervention modestly elevated the skills of the lower-performing group, while its absence led to a significant and rapid decline in the higher-performing group.

The analysis of Day 1 confirmed a significant challenge inherent to the quasi-experimental methodology: baseline non-equivalence. Quasi-experimental designs are often employed in educational settings when random assignment is not feasible (Shadish et al., 2002), but as seen here, this can result in groups with pre-existing differences. The findings from Day 1, where Group B (Experimental) significantly outperformed Group A (Control) in pronunciation ($p = .007$), vocabulary ($p = .029$), and confidence ($p = .006$), established this inequality as a major variable. Following the crossover, however, the Day 2 analysis showed that all significant differences between the groups vanished, with no significant differences in pronunciation ($p = .665$), fluency ($p = .911$), vocabulary ($p = .941$), or confidence ($p = .921$). The central purpose of the crossover design using participants as their own control to increase statistical power (StudySmarter, 2024), was achieved, allowing the analysis to shift from if the groups were different to why they had converged.

The paired-samples analyses provided a clear explanation. For the lower-performing Group A, the intervention's primary measurable impact was a significant increase in fluency ($p = .026$), while their vocabulary and confidence remained stable. This supports De Luna (2024), who found that interactive speaking tasks primarily enhance fluency, while other skills may remain unchanged without extended practice. This suggests the intervention was particularly effective at developing the procedural skill of fluency. The most dramatic finding, however, came from the higher-performing Group B. As shown by the non-parametric Wilcoxon test, which was required due to severe violations of normality, this group experienced a highly significant decrease in confidence ($p < .001$) after the intervention was removed. The corresponding t-test p-values of $p < .001$ for all speaking skills strongly imply this "fadeout" effect was pervasive.

This "skill decay," or the loss of acquired skills after periods of non-use, is a documented phenomenon in training and education (Australian Disaster Resilience Knowledge Hub, 2021). The results for Group B suggest that the control condition was not neutral but was an environment in which skills and confidence could not be maintained. This aligns with research on intervention "fadeout," where the benefits of a program diminish after it ends, especially if the subsequent environment fails to support the skills that were developed (National Institutes of Health, 2020).

From the perspective of Long’s Interaction Hypothesis (1996), this pattern is expected. Long posits that interaction and negotiation for meaning—through clarification requests, confirmation checks, and conversational adjustments—enhance comprehension and facilitate language development. The removal of roleplay meant students no longer engaged in these negotiation processes, leading to reduced opportunities for interactional feedback and language adjustment. This explains why

Group B's confidence and fluency declined sharply, consistent with Aubrey et al. (2020), who warned that negative or absent peer interaction can erode willingness to speak.

Moreover, the improvements in fluency for the lower-performing group can also be interpreted through Hymes's (1972) Communicative Competence framework. Roleplay activities required students to communicate in realistic, social situations that demanded not only grammatical accuracy but also sociolinguistic and strategic competence. The spontaneous use of English during these roleplays allowed learners to practice fluency and vocabulary use in authentic contexts, consistent with Communicative Language Teaching (CLT) principles that prioritize meaning-focused interaction over rote memorization (Rao & Zakaria, 2025).

Finally, the observed fadeout or skill decay after the removal of roleplay also aligns with both Swain's Output Hypothesis and Vygotsky's Sociocultural Theory. Swain's framework highlights the need for continuous language production to sustain fluency, while Vygotsky's theory emphasizes that learning is socially mediated. Once the collaborative and interactive elements of roleplay were withdrawn, students lost the communicative support necessary to sustain their advanced-level speaking performance within their Zone of Proximal Development (ZPD).

In conclusion, this study's findings suggest the intervention is a powerful force, not only for modestly building skill in lower-performing students but, more critically, for maintaining skills and confidence in higher-performing students, who demonstrated a significant decline in its absence. The results are consistent with Vygotsky's Sociocultural Theory, Swain's Output Hypothesis, Hymes's Communicative Competence, and Long's Interaction Hypothesis, which together explain how socially interactive, output-driven, and communicative learning environments—like roleplay—foster sustained fluency, vocabulary growth, and confidence in English-speaking performance.

Conclusions

This study examined how spontaneous roleplay influences Grade 12 students' fluency, pronunciation, vocabulary, and confidence using a quasi-experimental crossover design. Findings revealed that students in the control group, after participating in roleplay, showed significant improvements in fluency and confidence, highlighting the benefits of interactive, communicative strategies over conventional methods, particularly for lower-performing learners. The experimental group began with higher mean scores and consistently displayed "Advanced" performance; while their fluency and confidence were maintained, measurable improvement was limited due to a ceiling effect, indicating that quantitative measures may not fully capture subtle progress among already proficient students. The crossover phase further demonstrated the intervention's role in preventing skill decay, as students who no longer engaged in roleplay experienced notable declines in confidence and speaking performance. Overall, spontaneous roleplay is an effective pedagogical tool that fosters communicative competence, sustains confidence, and preserves language fluency, emphasizing the importance of continuous, interactive practice even among advanced learners.

Recommendations

For Teachers and Language Instructors

1. Integrate spontaneous roleplay regularly in English language classes to promote authentic communication and prevent decline in speaking confidence, as the study showed confidence dropped when roleplay was removed.
3. Provide differentiated roleplay scenarios that suit varying proficiency levels, ensuring that both advanced and developing learners remain challenged and engaged, addressing the ceiling effect observed among advanced learners.
4. Facilitate reflection and feedback sessions after each roleplay to help students become more conscious of their fluency, pronunciation, and vocabulary usage, consistent with Swain's Output Hypothesis.

For School Administrators and Curriculum Developers

1. Include roleplay activities as a standard component of oral communication modules to strengthen communicative competence in the curriculum.
2. Conduct teacher training workshops that focus on designing effective, spontaneous roleplay strategies that support language skill retention and address performance anxiety.
3. Encourage collaborative classroom environments that allow students to practice real-world conversational contexts in a low-anxiety setting.

For Students

1. Actively participate in spontaneous speaking activities to build both confidence and fluency beyond the classroom, recognizing that consistent practice prevents skill decline.
2. Engage in peer-to-peer communication and practice speaking English in informal contexts to maintain skill consistency and confidence.

For Future Researchers

1. Replicate the study in different academic strands or schools to determine whether similar outcomes occur among students with varying proficiency levels.
2. Extend the duration of the intervention to examine long-term effects of sustained roleplay practice on English-speaking development, particularly whether confidence gains can be maintained over time.
3. Incorporate qualitative methods such as interviews, observations, or student journals to capture deeper insights into learners' confidence, motivation, and attitudes toward roleplay, since quantitative scores may not fully reflect subtle progress.
4. Explore other interactive learning strategies that can complement roleplay, such as debates, improvisation, or drama-based learning, to further enrich communicative competence.

5. Investigate the ceiling effect observed in this study by using larger samples or alternative assessment tools that can capture incremental progress among already proficient learners.

Compliance with ethical standards

This study adhered to established ethical standards throughout its conduct. A transmittal letter was delivered to the principal, an informed permission from the teacher was obtained, ensuring their full awareness of the study's goals, procedures, potential risks, and anticipated benefits. An informed consent was given to the teacher-evaluator and participants of the intact classes, ensuring autonomy and freedom to withdraw from the study at any time. The researchers guaranteed that no participant would be subjected to any physical abuse or harm for this study. Confidentiality and anonymity were maintained by ensuring that all data were secured and individual identities were not disclosed in any way or form. Data were handled with strict caution, the study was undertaken with respect for the dignity and rights of all those involved, and the findings were used purely for research with no bias whatsoever. Additionally, artificial intelligence (AI) tools were employed to assist in data organization and the refinement of written outputs. These tools were used strictly as aids to enhance clarity and accuracy, with all final analyses and interpretations remaining under the researchers' critical judgement and accountability.

REFERENCES

Aini, F. F., Anwar, K., & Asmara, C. H. (2020). THE EFFECT OF ROLE PLAY FOR ENGLISH STUDENTS SPEAKING IN JUNIOR HIGH SCHOOL OF SMP MUHAMMADIYAH 11 SUKODADI. Journal of English Teaching Literature and Applied Linguistics, 4(2), 102. https://doi.org/10.30587/jetlal.v4i2.2275

Aubrey, S., King, J., & Almukhaild, H. (2020). Language learner engagement during speaking tasks: a longitudinal study. RELC Journal, 53(3), 519–533. https://doi.org/10.1177/0033688220945418

Australian Disaster Resilience Knowledge Hub. (2021). Understanding skill decay and skill maintenance in first responders. https://knowledge.aidr.org.au/media/9069/ajem_15_2021-04.pdf

Bower, I. S., Broadbent, J., Coussens, S., & Enticott, P. G. (2024). RETRACTED: Elevated ceiling heights reduce the cognitive performance of higher-education students during exams. Journal of Environmental Psychology, 97, 102367. https://doi.org/10.1016/j.jenvp.2024.102367

De Luna, A. M. B. (2024). Language Acquisition Games: Its effects on the students' oral Language skills. International Journal of Innovative Science and Research Technology (IJISRT), 2586–2594. https://doi.org/10.38124/ijisrt/ijisrt24aug773

Dornyei, Z., & Ryan, S. (2015). The Psychology of the language learner revisited. https://doi.org/10.4324/9781315779553

Dwiyanti, A., & Lolita, Y. (2023). The effectiveness of role play in improving speaking skill of EFL students. Prosodi: Jurnal Ilmu Bahasa dan Sastra, 17(2), 123–132. https://doi.org/10.21107/prosodi.v17i2.17064

Ghafar, Z. N., & Raheem, B. R. (2023). Factors Affecting Speaking Proficiency in English Language Learning: A general overview of the speaking skill. Journal of Social Science (JoSS), 2(6), 507– 518. https://doi.org/10.57185/joss.v2i6.107

Hasanah, H. (2021). The Impact of Confidence Building on Student Performance in Language Learning. International Journal of Language Education, 32(3), 122-134. https://doi.org/10.1177/13621688221106832

Hussam, A. A., & Saeed, A. (2025). The impact of Role-Playing techniques on the communication skills of junior high school students in Ghana. International Journal of Research and Innovation in Applied Science, X(V), 655–666. https://doi.org/10.51584/ijrias.2025.100500060

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), Sociolinguistics (pp. 269–293). Penguin. https://files.eric.ed.gov/fulltext/EJ1290295.pdf

Idham, S. Y., Subramaniam, I., Khan, A. B. B. M. A., & Mugair, S. K. (2022). The effect of Role-Playing techniques on the speaking skills of students at university. Theory and Practice in Language Studies, 12(8), 1622–1629. https://doi.org/10.17507/tpls.1208.19

Khasbani, I., & Seli, F. Y. (2021). The impact of role-play on students' speaking performance. Linguists Journal of Linguistics and Language Teaching, 7(2), 1. https://doi.org/10.29300/ling.v7i2.5216

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), Handbook of second language acquisition (pp. 413–468). AcademicPress https://www.academia.edu/16404697/The_Role_of_Negotiation_of_Meaning_in_L2_Interactions_An_Analysis_from_the_Perspective_of_Long_s_Interaction_Hypothesis

MacIntyre, P. & Gregersen, T. (2022). The idiodynamic method: willingness to communicate and anxiety processes interacting in real time. International Review of Applied Linguistics in Language Teaching, 60(1), 67-84. https://doi.org/10.1515/iral-2021-0024

MacIntyre, P. D., & Gregersen, T. (2022). The idiodynamic method: Willingness to communicate and anxiety processes interacting in real time.International Review of Applied Linguistics in Language Teaching, 60(1), 67–84. https://doi.org/10.1515/iral-2021-0024

National Institutes of Health. (2020). Persistence and fade-out of educational-intervention effects: Mechanisms and potential solutions. https://pmc.ncbi.nlm.nih.gov/articles/PMC7787577/

Rao V. R. K., & Zakaria N. Y. K. (2025) Role-Playing as a Tool for Improving ESL Learners' Communication Skills: A Systematic Review. Jurnal Pendidikan Malaysia 50(1) (2025): 18-34 http://dx.doi.org/10.17576/JPEN-2025-50.01-02

Salsabila, N., & Megawati, F. (2024). Discovering EFL Learners' Perspectives in Unscripted Role-Play. Journal of English for Academic and Specific Purposes (JEasp), 7(1), 42-54. https://ejournal.uin-malang.ac.id/index.php/jeasp/article/view/26765https://ejournal.uin-malang.ac.id/index.php/jeasp/article/view/26765

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

StudySmarter. (2024). Crossover studies: Design & definition. Retrieved from https://www.studysmarter.co.uk/explanations/medicine/pharmacy/crossover-studies/

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), Input in second language acquisition (pp. 235–253). Newbury House. Retrieved from http://www.isaacpub.com/images/PaperPDF/JAER_100154_202070903081702850.pdf

Togimin, N., & Jaafar, H. (2020). Role Play activities: an innovative approach to teaching speaking skills in an ESL classroom. LSP International Journal, 7(2), 19–36. https://doi.org/10.11113/lspi.v7.16310

Tuan, L. T. (2023). Initial Gains in Educational Interventions: Evidence from English Language Learning Programs. TESOL Quarterly, 57(1), 200-220. https://doi.org/10.1016/j.ijme.2019.04.002

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press. Retrieved from https://www.simplypsychology.org/zone-of-proximal-development.html

APA Citation:

Manginsay, K. P., Pampilo, Jr., R. D., Pactol, M. A., Oyda, F. G., Mayormita, H. M. T., & Tabia, J. C. J. D. (2026). THE EFFECTS OF ROLEPLAY ACTIVITIES ON STUDENTS' ENGLISH-SPEAKING SKILLS. Ignatian International Journal for Multidisciplinary Research, 4(2), 349–369. https://doi.org/10.5281/zenodo.18618879

Kierveymanginsay12@gmail.com
frietzoyda25@gmail.com
rollypams@gmail.com
mayormitahannahmarie@gmail.com
miechachu@gmail.com
jcjulyssestabia@gmail.com