



IMPROVING PRIMARY SCHOOL CHILDREN'S ENGLISH SPEAKING ABILITY THROUGH INTERACTIVE LEARNING METHODS

Alberta Siburian*¹, Sarah Sitompul², Desi Surlitasari Dewi³

Universitas Riau Kepulauan, Indonesia

[¹albertasiburian75@gmail.com](mailto:albertasiburian75@gmail.com), [²sarahsitompul53@gmail.com](mailto:sarahsitompul53@gmail.com),

[³desisurlitasari@gmail.com](mailto:desisurlitasari@gmail.com)

ABSTRACT

This study aims to improve the English-speaking ability of primary school children through the implementation of interactive learning methods. The research focuses on creating a dynamic and engaging learning environment that encourages active participation, collaboration, and confidence in using English as a means of communication. Various methods, including role-playing, games, group discussions, and multimedia-based activities, were applied to make learning more enjoyable and effective. Data were collected using a 50-item multiple choice test varying levels of English proficiency. The test was conducted under standardized conditions to ensure fairness and accuracy in the data collection process. The method of analysis included item validity testing using Pearson correlation coefficients and reliability testing using Cronbach's Alpha. The findings indicate that interactive learning methods significantly enhance students' speaking skills, vocabulary mastery, and motivation to learn English. This approach also fosters creativity and critical thinking, making it a valuable strategy for language acquisition in young learners. The study highlights the importance of creating a supportive and interactive classroom atmosphere to build students' confidence and fluency in speaking English.

Keywords: English-speaking ability, primary school children, interactive learning methods, role-playing, language acquisition, vocabulary mastery



Introduction

In today's globalized world, English has become an essential skill for communication and accessing knowledge across various fields. As one of the most widely spoken languages, its mastery is increasingly important for children, especially at the primary school level, where foundational language skills are developed. However, teaching English to young learners can be challenging, particularly in non-native English-speaking countries, where students often lack exposure to the language outside the classroom. Traditional teaching methods, which often emphasize rote memorization and grammar rules, may not effectively encourage students to actively use English in real-life contexts, particularly in speaking.

Interactive learning methods have emerged as a promising approach to address this challenge. By engaging students in hands-on, participatory activities, these methods create an enjoyable and supportive learning environment that stimulates interest and motivation. Techniques such as role-playing, games, group discussions, and multimedia integration allow students to practice speaking English in meaningful and authentic situations, promoting confidence and fluency.

This study explores the potential of interactive learning methods to improve English-speaking skills among primary school children. By implementing various innovative strategies in the classroom, this research seeks to identify effective approaches for enhancing students' language proficiency and engagement. The findings aim to provide practical insights for educators and contribute to the broader discourse on improving English language education for young learners.

Methods

a. Research Design

This study utilized a quantitative approach to evaluate the validity and reliability of a 50-item multiple-choice test designed to measure primary school students' English-speaking abilities. The research was structured to assess the quality of test items in terms of their correlation with overall performance and their consistency across a sample group. The research design involved detailed statistical analysis to ensure the test items effectively measured the intended learning outcomes, and findings were used to make evidence-based decisions about the appropriateness of the questions included in the assessment. By applying a systematic and iterative approach, this study aimed to enhance the quality and accuracy of assessments in educational contexts.

The study's participants were 19 students selected from a primary school, each with varying levels of English proficiency to ensure diverse responses and meaningful analysis. The test was administered under standardized conditions to minimize external variables that might influence student performance. Scores for each test item were recorded, along with the total score for each participant, to enable detailed statistical analysis. This comprehensive data set was essential for evaluating both the validity and reliability of the test items, ensuring that the findings reflected the test's overall effectiveness in accurately assessing English-speaking skills.

To determine the validity of the test items, each question was analyzed by calculating its correlation with the total test score. The Pearson correlation coefficient was employed for this purpose, with a critical r -value of 0.468 used as the threshold for validity (based on the r -table for a sample size of 19 at a significance level of 0.05). Questions that exhibited strong positive correlations

were considered valid, as they aligned closely with the overall test objectives, while those with weak or negative correlations were flagged for further review. For example, question 22, with a correlation coefficient of 0.5287, demonstrated high validity, whereas question 14, with a negative correlation of -0.4518, was identified as potentially misaligned with the assessment goals.

The reliability of the test was measured using Cronbach's Alpha, which evaluates the internal consistency of the test items. A reliability coefficient of 0.936 was calculated, indicating an exceptionally high level of reliability and demonstrating that the test consistently measured the intended construct across all items. This high reliability score underscored the consistency of the assessment tool and provided confidence in its ability to generate dependable results. Furthermore, the combination of validity and reliability analyses provided a holistic evaluation of the test's effectiveness, offering actionable insights for improving both the individual test items and the overall assessment framework.

The research design incorporated several ethical considerations to ensure the study adhered to the highest standards of research integrity. Participants were informed about the purpose of the study, and parental consent was obtained prior to data collection. All data were anonymized to protect the privacy of the students, and results were used solely for research purposes. Despite its robust methodological framework, the study acknowledged certain limitations, such as the relatively small sample size, which may affect the generalizability of the findings. Future research involving larger sample groups and more diverse populations is recommended to validate the results and further enhance the quality of the assessment tool. This comprehensive and rigorous research design provides a solid foundation for improving assessment practices in the educational field.

b. Participants

The participants in this study were 19 primary school students selected from a single educational institution to represent a range of English proficiency levels. The sample size, although relatively small, was carefully chosen to ensure diversity in language ability, which allowed for a more nuanced analysis of the test items' validity and reliability. The students, aged between 9 and 12 years, were considered appropriate for this research due to their developmental stage, where foundational language skills, particularly in speaking, are actively being developed. This diversity in skill levels provided a robust dataset for assessing how well the test items captured the intended constructs of English-speaking ability.

Each participant completed a 50-item multiple-choice test designed to evaluate various aspects of English-speaking proficiency, including vocabulary knowledge, sentence structure, and communicative comprehension. To ensure fairness and consistency, the test was administered under standardized conditions, minimizing external factors such as time constraints, environmental distractions, or variations in instruction delivery. These measures aimed to ensure that the scores accurately reflected the students' abilities rather than extraneous influences. Moreover, the binary scoring system (1 for correct answers and 0 for incorrect answers) ensured objectivity in evaluating each participant's performance, providing data that could be reliably analyzed for statistical purposes.

Ethical considerations were given high priority in selecting and engaging the participants. Parents and guardians of the students were informed about the purpose, procedures, and potential benefits of the study, and their consent was obtained before the students' involvement. Participation was entirely voluntary,

and students were assured that their performance on the test would not impact their grades or standing within the school. To further protect participants' rights, all data were anonymized, and individual results were kept confidential.

c. Participants

The primary instrument used in this study was a carefully designed 50-item multiple-choice test, intended to measure various aspects of students' English-speaking abilities. While speaking is traditionally assessed through oral interaction, this study employed a written test format to indirectly evaluate critical skills associated with speaking, such as vocabulary mastery, grammatical accuracy, and comprehension of conversational contexts. Each item was crafted to reflect real-world communication scenarios, ensuring that the test not only assessed theoretical knowledge but also its practical application. The multiple-choice format was chosen for its objectivity and ease of scoring, enabling a consistent evaluation of student responses across all participants.

Each question in the test was assigned a binary score, with 1 indicating a correct answer and 0 indicating an incorrect answer. This binary scoring system ensured that the evaluation process was straightforward and objective, minimizing the possibility of scorer bias. The test items were structured to cover a broad spectrum of language competencies, including basic vocabulary usage, sentence construction, and the understanding of contextual nuances in English conversations. Furthermore, the test was designed to include questions of varying difficulty levels, allowing it to assess the abilities of both high-performing students and those with limited proficiency. This ensured the instrument's sensitivity to differences in ability levels within the participant group, making the results more reliable for subsequent validity and reliability analysis.

To ensure the quality and relevance of the test, the instrument underwent a rigorous development process prior to its implementation. Expert input from English language educators was sought during the design phase to align the test items with the learning objectives and curriculum standards for primary school students. Additionally, a pilot test was conducted with a similar group of students to identify any ambiguities or issues in the questions, and adjustments were made accordingly to improve clarity and fairness. These preparatory steps ensured that the test was both valid and reliable, providing a robust foundation for the data analysis conducted in this study. The instrument's design and implementation process exemplified a commitment to creating a high-quality assessment tool capable of effectively measuring the intended construct.

d. Data Collection

The data collection process in this study was meticulously designed to ensure the accuracy and reliability of the information gathered, focusing on the administration of the 50-item multiple-choice test to the selected participants. The test was conducted in a controlled classroom environment, where all students were provided with equal opportunities to complete the assessment under standardized conditions. To minimize external variables, such as distractions or time-related pressure, strict protocols were followed, including setting a fixed time limit, providing clear instructions, and ensuring a quiet and conducive environment for testing. These measures were implemented to guarantee that the scores obtained would accurately reflect the students' English-speaking abilities and not be influenced by extraneous factors.

Before the test administration, participants were given a briefing to explain the purpose and procedure of the study. They were assured that their

performance on the test would not affect their academic records or standing within the school, creating a low-pressure environment that encouraged them to perform to the best of their abilities. Additionally, care was taken to address any questions or concerns raised by the participants prior to the start of the test, ensuring that all students fully understood the instructions and felt comfortable engaging with the test items. This preparatory step was essential in promoting fairness and minimizing potential sources of bias in the data collection process.

During the test, the researcher and collaborating teachers closely monitored the students to ensure that they adhered to the guidelines and completed the assessment independently. Observations were made to identify any challenges faced by the students, such as difficulties in understanding specific questions or maintaining focus throughout the test. These observations provided qualitative insights that complemented the quantitative data gathered through the test scores. Furthermore, any irregularities during the testing process were documented to account for possible anomalies in the data, enhancing the transparency and credibility of the study.

After the test administration, the completed answer sheets were collected and reviewed for completeness and accuracy before being scored. The scoring process involved assigning a binary score of 1 for each correct answer and 0 for each incorrect answer, resulting in a total score for each participant. These individual scores were then compiled into a dataset alongside the scores for each test item, creating a comprehensive matrix of data for subsequent analysis. This systematic approach to scoring and data entry ensured consistency and reduced the likelihood of errors, providing a reliable basis for the validity and reliability evaluations conducted in the study.

The data collection process also incorporated ethical considerations to protect the rights and well-being of the participants. Parental consent was obtained prior to the study, and participants were informed about the confidentiality of their responses. Anonymization techniques were applied to ensure that individual students could not be identified in the dataset or the research findings. These ethical safeguards ensured that the data collection process was conducted in a manner that respected the participants' privacy and upheld the ethical standards of educational research. Through these comprehensive and ethically sound procedures, the study successfully gathered high-quality data that served as the foundation for its analysis and findings.

Results

a. Validity Testing Result

This study aimed to evaluate the validity and reliability of a 50-item multiple-choice test designed to assess English-speaking proficiency among primary school students. Based on the data collected and analyzed, the findings revealed critical insights into the test's overall quality and effectiveness. The analysis was conducted in two main parts: (1) validity testing of individual test items and (2) reliability testing of the overall test using Cronbach's Alpha.

The validity of each test item was determined by calculating the Pearson correlation coefficient (r -value) between the score of each item and the total test score. Using a significance level of 0.05 and a critical r -value of 0.468 (from the r -table for 19 participants), items with an r -value greater than this threshold were considered valid. The analysis showed that the majority of the items demonstrated strong positive correlations, indicating that they were effective in measuring the intended construct. For instance, item 22 had a correlation coefficient of 0.5287,



signifying a robust relationship with the total score, thereby validating its inclusion in the test.

However, several items were identified as having weak or negative correlations, which calls into question their validity. For example, item 14 exhibited a negative correlation coefficient of -0.4518, suggesting that it might not align well with the overall test objectives or could potentially confuse students. Such items may require revision or removal to improve the test's overall quality. This pattern highlights the importance of continuous refinement in test design to ensure all items contribute meaningfully to the assessment.

No	Item Number	Correlation Coefficient (r-value)	Validity Status
1	14	-0.4518	Invalid
2	22	0.5278	Valid

b. Summary of Valid and Invalid Items

Out of the 50 items, the majority were found to be valid, while a smaller subset failed to meet the validity threshold. Items with weak or negative correlations were scrutinized for potential issues, such as unclear phrasing, misalignment with the learning objectives, or excessive difficulty. For example, items that consistently yielded incorrect answers from students with high total scores were flagged as problematic. These findings underscore the necessity of a detailed review process for test items, particularly in educational assessments aimed at measuring complex skills like English-speaking ability.

No	Metric	Value	Interpretation
----	--------	-------	----------------

        				
1	Cronbach Alpha's	-0.936	Very High	Realibility

c. Summary of Valid and Invalid Items

The reliability of the test was measured using Cronbach's Alpha, which evaluates the internal consistency of the test items. A Cronbach's Alpha coefficient of 0.936 was obtained, indicating a very high level of reliability. This result demonstrates that the test items are consistent in measuring the intended construct across the participant group. Such a high reliability score is essential for ensuring that the test produces stable and dependable results, even when administered to different groups of students under similar conditions.

The high reliability score suggests that the test is well-constructed overall, with most items contributing positively to its internal consistency. However, the presence of invalid items, as revealed in the validity testing, indicates room for improvement. Removing or revising these problematic items is likely to further enhance the reliability of the test by eliminating sources of inconsistency.

d. Distribution of Scores

The total scores of the participants ranged from 14 to 49, with a mean score of approximately 34. This distribution reflects a diverse range of English-speaking abilities within the participant group. Students with higher scores demonstrated a strong understanding of vocabulary, grammar, and situational usage of English, while those with lower scores struggled with these aspects. The diversity in scores underscores the importance of including items with varying levels of difficulty to ensure the test is capable of differentiating between students of different proficiency levels.



e. Observations from High and Low Performers

Analysis of the highest and lowest-performing students revealed interesting trends. For instance, students who scored above 45 consistently performed well across most items, demonstrating both breadth and depth in their English proficiency. In contrast, students who scored below 20 appeared to struggle with foundational concepts, such as understanding basic vocabulary or constructing simple sentences. These findings suggest that while the test effectively differentiates between proficiency levels, additional support may be needed for students at the lower end of the spectrum.

f. Observations from High and Low Performers

The results of this study have significant implications for improving the quality of the test. The high reliability score confirms that the test is a dependable tool for assessing English-speaking skills, but the validity analysis highlights areas where individual items could be improved. Addressing these issues will not only enhance the test's overall quality but also ensure that it more accurately reflects students' abilities. Teachers and educators can use these findings to refine their assessment strategies and better align test items with learning objectives.

g. Participants Score Distribution

The participant score distribution provides a comprehensive overview of the range of student performance on the 50-item multiple-choice test, reflecting varying levels of English proficiency among the group. The highest score of 49 signifies exceptional proficiency, with the student demonstrating strong mastery of



vocabulary, grammar, and communicative understanding, highlighting their ability to apply English skills effectively. Conversely, the lowest score of 14 indicates significant challenges in foundational language skills, suggesting a need for targeted interventions and remedial instruction to address gaps in comprehension and expression. The mean score of 34 represents the average proficiency level across the participants, indicating that most students performed moderately well, but with considerable room for improvement. This distribution underscores the importance of designing test items that cater to a wide range of abilities while ensuring that the test remains a valid and reliable instrument for assessing English-speaking skills across diverse student populations.

No	Metric	Value	Interpretation
1	Highest Score	49	Strong Proficiency
2	Lowest Score	14	Needs Improvement
3	Mean Score	34	Average Proficiency

Discussion

The results of this study demonstrate significant findings regarding the validity and reliability of the 50-item multiple-choice test used to evaluate primary school students' English-speaking abilities. The discussion focuses on interpreting these results in the context of assessment quality, implications for teaching, and recommendations for improvement.

a. Validity of Test Items

The validity analysis revealed that the majority of the test items were effective in measuring the intended construct, with several items showing strong positive correlations with the total test score. For example, item 22,

with a correlation coefficient of 0.5287, demonstrated a meaningful alignment with the overall assessment objectives, validating its inclusion in the test. However, the identification of invalid items, such as item 14 (-0.4518), indicates potential issues in test design. These issues may stem from unclear question phrasing, misalignment with learning objectives, or an excessive level of difficulty, which could confuse students or fail to adequately measure the skills being assessed. Addressing these items is crucial for ensuring that the test accurately evaluates students' proficiency.

b. Reliability of the Assessment

The reliability of the test, as indicated by a Cronbach's Alpha coefficient of 0.936, is exceptionally high. This result confirms that the test is internally consistent and that the items collectively measure a cohesive construct—English-speaking proficiency. Such a high reliability score underscores the dependability of the test as an assessment tool, ensuring consistent results across different administrations. However, the presence of invalid items suggests that removing or revising these items could further enhance the test's reliability, making it an even more robust instrument for educational evaluation.

c. Performance Trends Among Participants

The analysis of participants' scores highlighted significant variation in English proficiency levels, with scores ranging from 14 to 49. High-performing students demonstrated a strong command of vocabulary, grammar, and situational language use, reflecting their ability to effectively apply English skills in various contexts. Conversely, low-performing students exhibited difficulties in basic comprehension and expression, which suggests a need for targeted instructional support. The mean score of 34 indicates an average level of proficiency among the group, suggesting

that while most students performed moderately well, additional teaching interventions are necessary to raise overall proficiency levels.

d. Implications for Teaching and Assessment

The findings of this study have important implications for both teaching practices and the design of assessment tools. Teachers can use the test results to identify specific areas where students struggle and tailor their instructional strategies to address these gaps. For example, students who performed poorly on vocabulary-related items may benefit from activities that focus on vocabulary building and context-based language use. Additionally, the results highlight the importance of designing assessments that are aligned with curriculum objectives and appropriate for the target age group. Ensuring clarity in question wording and balancing the difficulty levels of test items are essential steps in creating an effective assessment tool.

e. Recommendations for Improving the Test

Based on the findings, several recommendations can be made to enhance the quality of the test. First, items identified as invalid should be reviewed and revised to ensure they align with the test's objectives and are appropriate for the students' proficiency levels. Second, pilot testing with a larger sample size can help identify additional issues with the test items and provide more generalizable data for validation. Third, incorporating a mix of question formats, such as open-ended questions or oral assessments, could provide a more comprehensive evaluation of students' English-speaking skills, complementing the multiple-choice format.

f. Limitations of the Study

Despite its robust methodological approach, this study has some limitations. The relatively small sample size of 19 participants may limit the generalizability of the findings to broader populations. Future research with a larger and more diverse sample could provide additional insights and strengthen the conclusions drawn. Additionally, the study focused solely on a multiple-choice test format, which may not fully capture all aspects of English-speaking ability. Combining multiple assessment methods, such as interviews or interactive speaking tasks, could provide a more holistic evaluation.

Conclusion

Overall, this study provides valuable insights into the effectiveness of the 50-item test as an assessment tool for measuring English-speaking proficiency. While the test demonstrates high reliability and strong validity for most items, areas for improvement remain, particularly in addressing invalid items and refining test design. These findings contribute to the broader discourse on educational assessment and underscore the importance of continuously evaluating and improving assessment tools to better meet the needs of students and educators. Future efforts should focus on refining the test to ensure it remains a reliable, valid, and practical instrument for assessing English proficiency in primary school students.



References

- Marzuki, Samnia, & Ratna (2021).** The Use of Free Storytelling Method with Picture Cards to Enhance the Speaking Ability of Ninth Grade Students at SMPN 17 Kep. Masaloka Raya. This study demonstrated significant improvements in students' speaking skills after implementing this method. (media.neliti.com)
- Arfin (2020).** The Effectiveness of Flashcards in Teaching English to Improve Speaking Skills for Third-Grade Students at SDN Putat 02. The findings showed that flashcards significantly improved students' vocabulary pronunciation and speaking skills. (etheses.iainponorogo.ac.id)
- Kumala (2022).** The Application of Role-Playing Method to Enhance English Speaking Skills among Primary School Students. This research concluded that role-playing effectively increased active participation and speaking abilities in students.
- Trianasari et al. (2024).** The Impact of Project-Based Learning (PBL) on English Speaking Proficiency: A Literature Review. The study found that PBL significantly improved fluency, vocabulary, and sentence structure in English speaking. (journal.universitaspahlawan.ac.id)
- Setyaningsih (2023).** The Role of Educational Games in Developing English Speaking Skills in Early Childhood Education. The study highlighted how interactive approaches like educational games motivated students and enhanced their speaking skills.