Vol. 5 Issue 2, July-December 2026
International Journal of Educational Research Excellence (IJEREC) e-ISSN: 2830-7933
https://ejournal.ipinternasional.com/index.php/ijere DOI:10.55299/ijere.v5i2.1754

The Effectiveness of The Demonstration Method in Enhancing Students' Speaking Skills in Retelling Folktales: A Quasi-Experimental Study in 4th grade Elementary Students

Zulfahmi[1*)], Rahmat Kartolo[2]
[1,2]Universitas Muslim Nusantara Al Washliyah, Medan, Indonesia
Email: zulfahmi71173@gmail.com[1], rahmatkartolo@umnaw.ac.id[2]
Correspondence Authors: zulfahmi71173@gmail.com
Article history: Received March 28, 2026; revised April 30, 2026; accepted May 05, 2026

ABSTRACT

This research is motivated by the low ability of students to convey the contents of fairy tales orally in a coherent, clear, and expressive manner. Learning is still dominated by the lecture method and does not actively involve students in speaking activities. The research method used is quantitative with a quasi-experimental approach. The design used is Posttest Only Control Group Design, involving two groups: the experimental class (18 students) who received the demonstration method treatment and the control class (18 students) who followed conventional learning. The data collection technique was carried out through a speaking skills test which was assessed based on indicators of fluency, mastery of content, intonation, expression, and story structure. Data were analyzed using the t-test and coefficient of determination. The results showed that the average value of students' speaking skills in the experimental class reached 90.88, while in the control class it was only 86.94. The t-test produced a T-value of 2.77> T-table 2.03 (α = 0.05), which indicates that Ho is rejected and Ha is accepted. This means that there is a significant influence of the application of the demonstration method on students' speaking skills. Furthermore, the correlation coefficient of r = 0.967 with a determination coefficient of R^2 = 0.935 indicates that 93.5% of the increase in students' speaking skills can be explained by the use of the demonstration method. The demonstration method has proven effective in creating an active, participatory, and enjoyable learning atmosphere. Students look more confident, enthusiastic, and able to convey stories with a logical structure and appropriate expression. The teacher acts as a model and facilitator, so that students gain direct and meaningful learning experiences. This study concludes that the demonstration method has a significant effect on improving students' speaking skills in retelling fairy tales. Therefore, this method is worthy of being used as an innovative alternative learning strategy in Indonesian language subjects in elementary schools.

Keywords: Demonstration Method, Speaking Skills, Telling Fairy Tales

I. INTRODUCTION

Education serves as the fundamental cornerstone of civilization, constituting the primary pillar supporting individual advancement, societal progress, national development, and global evolution. Rather than merely facilitating knowledge transfer from educators to learners, education is a transformative process that consciously and systematically guides individuals toward realizing their maximum potential. As defined in Indonesian Law No. 20 of 2003 concerning the National Education System, education is characterized as "a conscious and planned effort to create a learning atmosphere and learning process so that students actively develop their potential to possess spiritual strength, self-control, personality, intelligence, noble character, and skills needed by themselves, society, nation, and state. This comprehensive definition underscores the holistic and multidimensional nature od education, encompassing not only intellectual development but also the spiritual, emotional, social, and psychomotor aspects of learners (Aliyeva, 2025).

The overarching objective of national education is to develop learners' potential comprehensively, cultivating individuals who demonstrate faith and devotion to God Almighty, possess noble character, maintain physical and mental health, acquire extensive knowledge, demonstrate competence in thinking and acting, exhibit creativity in generating innovations, display independence in decision-making, accept responsibility for their actions, and embody democratic citizenship principles that uphold humanitarian values and justice. This vision reflects Indonesia's commitment to producing future generations who possess not only academic intelligence but also high moral integrity, proficient social skills, and passionate nationalism (Mutiara Hanif Saputri et al., 2025).

Within the context of elementary education, the learning process plays an extraordinarily crucial role in establishing foundational language competence. Language, functioning as the primary communication instrument, serves as a bridge connecting individuals with their surrounding environment. Proficient language abilities enable children to comprehend information, express themselves authentically, engage in social interactions, and develop critical thinking capacities. Consequently, mastering comprehensive language skills is one of the foremost priorities in elementary education curricula (Fegter et al., 2025).

Among the fundamental aspects of language competence, speaking skills occupy a strategic position as they constitute the primary means through which individuals communicate directly, convey ideas, express feelings and emotions, and share thoughts with others orally. Effective speaking ability extends beyond merely articulating words clearly; it involves the capacity to organize thoughts logically, select appropriate diction, employ suitable intonation and pauses, and attend to facial expressions and body language to ensure that messages are properly received by listeners (Wardhana et al., 2025).

According to Nurul Aprillia et al. (2008), a prominent linguistics expert, speaking skills are defined as the ability to express thoughts, ideas, and feelings orally with accuracy and correctness before others. This definition emphasizes the importance of clarity, precision, and politeness in oral communications. Individuals possessing proficient speaking skills can effectively convey their intentions, persuade others, build positive relationships, and participate actively in various social contexts (Nurul Aprilla et al., 2025).

Despite the recognized importance of developing speaking competencies, empirical observations have revealed a significant gap between expectations and reality. Preliminary observations conducted in Grade IV of SDN 401 Panyabungan, Mandailing Natal Regency, identified substantial issues regarding students' speaking skills, particularly in the context of retelling folktales. This phenomenon manifests through several indicators: students' lack of courage to perform and speak before their classmates, frequent hesitation or interruption when delivering narratives, and difficulty in organizing folktale plots coherently and systematically.

This warrants serious attention because speaking skills are one of the fundamental competencies that elementary-level students must master. Inadequate speaking skills can hinder students' ability to participate actively in learning, express their understanding of the subject matter, and build self-confidence in interacting with their social environment. If this problem remains unaddressed, it may negatively impact students' overall language development and ultimately affect their academic achievement across various subjects (Kırlangıç & Bozkurt, 2025).

One factor strongly suspected of contributing to students' poor speaking skills is the lack of variation in teaching methods employed by classroom teachers. Conventional teaching methods, dominated by lecturing, often prove ineffective in stimulating student activity and providing sufficient opportunities for speaking practice. Passive and monotonous learning experiences can make students feel bored, insufficiently motivated, and reluctant to acgively participate in teaching-learning activities. Consequently, their speaking skill potential cannot be optimally developed.

Recognizing the urgency of addressing these challenges, educators must seek innovative and effective solutions to enhance students' speaking skills. One teaching method believed to possess substantial potential for achieving this objective is the demonstration method. According to Kaya and Aydeniz (2025), the demonstration method is a teaching approach that directly displays how a process occurs or demonstrates how to perform something to learners. This method involves modeling or performance by teachers or students to illustrate specific concepts, procedures, or skills (Kaya & Aydeniz, 2025).

The demonstration method is considered effective because it provides student with more concrete and visual learning experiences. By directly observing directly how a process occurs or how a skill is performed, students can more easily comprehend instructional materials and retain them for longer periods. Additionally, the demonstration method can enhance students' learning motivation because they become actively involved in the learning process through observation and imitation.

In the context of folktale instruction, the demonstration method is exceptionally relevant. Folktales, representing one form of children's literature, are rich in moral values, imagination and expression. Through the demonstration method, teachers can directly model and practice effective and engaging story delivery techniques. This encompasses critical aspects such as varied voice intonation to animate characters and the story atmosphere, facial expressions corresponding to emotions embedded within folktales, and body gestures supporting overall message delivery.

By observing demonstrations performed by teachers or more proficient peers, students acquire clear models of proper and correct folktale retelling techniques. They can emulate the vocal techniques, expressions, and gestures employed by others, thereby gradually improving their speaking skills. Furthermore, the

demonstration method provides opportunities for students to practice their speaking skills directly before class, enabling them to receive constructive feedback from teachers and classmates.

The development of speaking skills is a complex process that involves multiple dimensions of language competence. Mister et al. (2023) characterized speaking as a productive (oral) skill involving the utilization of sound systems, vocabulary, grammatical structures, and sociocultural elements. This definition emphasizes that speaking skills encompass not only linguistic aspects but also sociocultural dimensions of speaking. Effective speakers must adapt their language style and word choice to the social context and audience background (Mister et al., 2023).

Several indicators serve as essential benchmarks for measuring speaking proficiency. Lestari (2018) identified key indicators, including fluency in speaking, appropriate word usage, coherent content delivery, suitable intonation and volume, and the ability to adjust language according to interlocutors. These indicators provide comprehensive assessment criteria that educators can employ to systematically evaluate and develop students' speaking competencies systeematically.

The Speaking fluency involves speech production speed and appropriate pausing. Students who demonstrate fluent speaking can convey ideas without being hindered by difficulties in word selection or sentence construction. Richards (2008) emphasized the importance of lexical appropriateness in speaking, which mean selecting words that are semantically and pragmatically suitable for communication situations. At the elementary level, this signifies students' ability to employ vocabulary relevant to the folktale topics being retold.

Regarding coherent content delivery, discourse organization constitutes one of the important aspects of speaking skill assessment. In the context of folktale retelling, coherence means that students can convey story plots cohesively from beginning to end. How Intonation plays an essential role in conveying meaning and attitudes in conversation. In folktale retelling, varied intonation can animate the characters and story atmosphere.

Folktales constitute more than mere fantastical narratives; they are cultural artifacts rich in history, values, and psychological functions. The Folktales are fictional stories containing moral messages, typically conveyed orally to children. This definition highlights two important aspects of folktales: their nature as fictional narratives that distinguish them from historical accounts or factual reports, and the presence of implicit moral messages aimed at providing life lessons and character formation for audiences (Mega Narendrani, 2025).

In The Uses of Enchantment, she argues that folktales play important psychological roles in children's development. According to his analysis, folktales not only entertain but also help children understand internal conflicts, overcome fears, and discover the meaning of life. Through the symbols and metaphors contained in folktales, children can indirectly process the emotions and challenges they face in the real world.

Wibowo (2016) affirms that folktales teach children value values of honesty, courage, mutual assistance, and hard work. These values are frequently conveyed through character behaviors exemplified in the stories. Protagonist characters who demonstrate honesty receive goodness, courageous characters successfully overcome obstacles, helpful characters receive assistance, and diligent characters achieve success. Conversely, antagonist characters possessing negative traits receive proportionate consequences.

II. METHODS

A. *Research Design*

This study employed a quantitative research approach using a quasi-experimental design. Quasi-experimental research represents a scientific investigation method aimed at discovering causal relationships between independent and dependent variables under controlled conditions, although without complete randomization, as is characteristic of true experimental designs. This design was selected considering the practical constraints of classroom settings, where the complete random assignment of individual students would disrupt existing class structures and administrative processes (Sugiyono, 2019).

Specifically, this study implemented the Posttest Only Control Group design, one of the most commonly utilized quasi-experimental designs in educational research. In this design configuration, two groups of participants were established: an experimental group receiving treatment through the demonstration method and a control group participating in conventional learning. Both groups were subsequently administered post-tests to assess their speaking skill competencies after the intervention period.

The schematic representation of this research design follows the pattern:

Experimental Group: $X \rightarrow O_1$
Control Group: $- \rightarrow O_2$

Where:

X = Treatment using the demonstration method

O_1 = Post-test measurement for the experimental group

O_2 = Post-test measurement for the control group

This design was deemed appropriate for investigating the research objectives because it enable the comparlearning outcomes between students exposed to the demonstration method intervention and those receiving conventional instruction, thereby facilitating the determination of the effectiveness of the demonstration method in enhancing speaking skills.

B. *Research Variables*

Th is study incorporated two principal variables that require operational definition and measurement. The independent variable (X) consisted of the demonstration method, defined as a teaching approach involving direct modeling or demonstration by teachers showing proper techniques for retelling folktales, including aspects of intonation, expression, gestures, and narrative coherence. The dependent variable (Y) comprised students' speaking skills in retelling folktales, operationally defined as their ability to orally convey folktale contents coherently, clearly, expressively, and comprehensibly, assessed through validated speaking skill rubrics (Creswell, 2021).

C. *Research Setting and Participants*

This research was conducted at SDN 401 Panyabungan, located on Jl. Adam Malik Lintas Timur, Desa Sipolu Polu, Panyabungan District, Mandailing Natal Regency, North Sumatra Province. The school, established on August 1, 2017, holds an "A" accreditation from the Ministry of Education and Culture. At the time of the research implementation, SDN 401 Panyabungan served 415 students, comprising 206 male and 209 female students, supported by 19 professional teachers and led by Principal Zulfahmi.

The research was conducted during the second semester of the 2025-2026 academic year, specifically from February to April 2026. This timeframe was selected to align with the curriculum coverage of narrative text materials in the Indonesian Language subject for fourth-grade students, as stipulated in the applicable curriculum.

D. *Population and Sampling*

The research population consisted of all fourth-grade students at SDN 401 Panyabungan during the 2025-2026 academic year, totaling 36 students distributed across two parallel classes. According to Arikunto (2013), when the population is relatively small (fewer than 100 individuals), researchers may utilize the entire population as research subjects, a technique termed total or census sampling. Consequently, this study employed total sampling, designating all 36 fourth-grade students as the research participants.

Sample allocation followed intact group assignment, where Class IV-A (18 students) was designated as the experimental group receiving instruction through the demonstration method, while Class IV-B (18 students) served as the control group participating in conventional teaching. This assignment considered class homogeneity regarding prior academic achievement, demographic characteristics, and teacher qualifications, to minimize potential confounding variables.

E. *Data Collection Instruments*

The primary data collection instrument comprised a speaking skill assessment test administered in a performance format, wherein students individually retold folktales before the class while being evaluated by trained assessors. The speaking skill test was developed based on comprehensive theoretical frameworks and validated through expert judgment and pilot testing.

The speaking skill assessment rubric incorporated five principal indicators, each evaluated on a scale of 1-4 (where 1=poor, 2=fair, 3=good, and 4=excellent). These indicators include

Table 1. Speaking Skill Assessment Rubric

Indicator	Description	Score Range
Fluency	Ability to speak continuously without excessive hesitation, repetition, or unnatural pausing	1-4
Content Mastery	Comprehension and accurate conveyance of folktale elements including characters, settings, plot sequences, and moral messages	1-4
Intonation	Appropriate variation in pitch, stress, and rhythm to convey meaning and maintain audience engagement	1-4
Expression	Suitable facial expressions, gestures, and emotional conveyance aligned with narrative content	1-4
Narrative Structure	Logical organization and coherent sequencing of story components from orientation through resolution	1-4

Each indicator's score ranged from 1-4, yielding a maximum possible total score of 20 points, which was subsequently converted to a 100-point scale using the formula: Final Score = (Total Score / 20) × 100.

B. Instrument Validity and Reliability

Prior to implementation in the main study, the speaking skill assessment instrument was rigorously validated. Content validity was established through expert judgment by three experienced Indonesian language education specialists who evaluated the alignment between assessment indicators and speaking skill constructs, the appropriateness for fourth-grade developmental levels, and the clarity of rubric descriptions. Expert feedback resulted in minor refinements to the rubric language and scoring criteria.

Construct validity was examined through pilot testing with 20 fourth-grade students from a comparable elementary school not included in the main study. Inter-rater reliability was assessed by having two independent raters evaluate the same student performances, yielding a substantial inter-rater agreement coefficient of κ=0.82 (Cohen's Kappa), indicating high consistency in scoring procedures.

Students in the experimental group participated in six instructional sessions over three weeks, each lasting 70 minutes, implementing the demonstration method to teach folktale retelling skills. The demonstration method was implemented follow systematic procedures adapted from Mulyasa (2014).

Session 1-2: Introduction and Modeling

The teacher introduced the concept of folktales and their characteristics

The teacher selected an age-appropriate Indonesian folktale ("Malin Kundang")

The teacher provided an expert demonstration of effective folktale retelling, explicitly highlighting key elements: varied intonation to differentiate characters, appropriate facial expressions conveying emotions, strategic gestures emphasizing plot points, and logical narrative sequencing.

Students observed and analyzed the demonstrated performance, identifying effective techniques employed

Session 3-4: Guided Practice

Students were divided into small groups of 3-4 members

Each group received a different folktale appropriate for their developmental level

Groups collaboratively prepared their retelling performance with teacher guidance and feedback

Selected students demonstrated their retelling before the class

The teacher and peers provided constructive feedback focusing on the five assessment indicators

Students engaged in reflective discussions about observed strengths and areas for improvement

Session 5-6: Independent Performance

Each student individually selected a folktale for independent retelling

Students were provided preparation time to practice their narratives

All students performed their individual retelling demonstrations

Continuous formative feedback was provided throughout the learning process

Throughout all sessions, the teacher functioned as both an expert model and facilitator, creating a supportive learning environment that encouraged risk-taking and progressive skill development.

C. Control Group Intervention

Students in the control group participated in conventional Indonesian language instruction, focusing on narrative texts over the same timeframe. The conventional approach primarily employ lecture methods combined with teacher-led discussions. The instructional activities included:

Teacher explanation of folktale characteristics and elements
Teacher reading folktales aloud to students
Whole-class discussions about story content and moral messages
Completion of written comprehension exercises
Limited opportunities for oral retelling practice

Notably, the control group's instruction did not incorporate systematic modeling or demonstration of effective speaking techniques, nor did it provide structured opportunities for repeated performance practice with feedback.

D. Data Collection Procedures

Data collection occurred during the final week of April 2026, following the completion of the three-week intervention period. All 36 research participants (18 in the experimental group and 18 in the control group) individually performed folktale retelling tasks in a standardized testing environment. To ensure consistency and minimize assessment bias, all performances were video-recorded and subsequently evaluated by two trained independent raters who were blind to the students' group assignments.

Each student randomly selected one folktale from a pool of five familiar Indonesian folktales (Malin Kundang, Timun Mas, Sangkuriang, Bawang Merah Bawang Putih, and Keong Mas). Students were provided 10 minutes for silent preparation, during which they could review the story outline but could not write a script or detailed notes. Following preparation, each student performed their retelling before the class for approximately 5-7 minutes while being video-recorded.

The assessment process was conducted according to standardized protocols. Two independent raters, both experienced elementary Indonesian language teachers who had received specific training in applying the speaking skill rubric, evaluated each recorded performance. The raters independently assigned scores for each of the five indicators, and the final scores were calculated as the mean of both raters' evaluations. In cases where raters' scores differed by more than five points on the 100-point scale, a third expert rater reviewed the performance to determine the final score.

E. *Data Analysis Procedures*

The collected data underwent systematic statistical analysis to test the research hypotheses and address the research questions. The analysis proceeded through several sequential stages.

E.1. Descriptive Statistics

The initial descriptive analysis calculated measures of central tendency (mean, median, mode) and dispersion (standard deviation, range, variance) for both the experimental and control groups. Frequency distributions were constructed to visualize the score distributions across specified intervals.

E.2. Prerequisite Testing

Prior to conducting inferential statistical tests, prerequisite analyses were performed to verify whether the data met the parametric test assumptions.

a) Normality Testing

The Kolmogorov-Smirnov test examined whether the data from both groups followed normal distributions, tested at α=0.05 significance level. The null hypothesis (H_0) stating that data originate from normally distributed populations would be retained if the calculated significance values exceeded 0.05.

The Kolmogorov-Smirnov statistic was calculated using:

$$D = \max|F_0(X) - S_n(X)|$$

Where:

D = maximum absolute difference between theoretical and empirical cumulative distributions
$F_0(X)$ = theoretical cumulative distribution function
$S_n(X)$ = empirical cumulative distribution function

b) Homogeneity of Variance Testing

Levene's test assessed whether the variances between the experimental and control groups were homogeneous, evaluated at α=0.05 significance level. The null hypothesis (H_0), positing equal population variances, would be retained if the calculated F-statistic's significance exceeded 0.05.
The homogeneity test employed the F-test formula:

$$F = S_1^2 / S_2^2$$

Where:

F = calculated F-value

S_1^2 = larger variance

S_2^2 = smaller variance

E.3. Hypothesis Testing
Following confirmation that the data satisfied parametric test assumptions, and independent samples t-test was performed to examine whether significant differences existed between the experimental and control group means. Given equal sample sizes (n_1=n_2=18) and homogeneous variances, the pooled variance t-test formula was employed as follow:

$$t = (\bar{X}_1 - \bar{X}_2) / \sqrt{[((n_1-1)S_1^2 + (n_2-1)S_2^2)/(n_1+n_2-2)) \times (1/n_1 + 1/n_2)]}$$

Where:

t = calculated t-value

$\bar{X}_1$ = experimental group mean

$\bar{X}_2$ = control group mean

n_1 = experimental group sample size

n_2 = control group sample size

S_1^2 = experimental group variance

S_2^2 = control group variance

The calculated t-value was compared against critical t-values from t-distribution tables at α=0.05 significance level with degrees of freedom (df) = n_1+n_2-2 = 34. The decision rule specified that if t-calculated > t-critical, reject H_0, concluding that significant differences exist between groups.

E.4. Correlation and Coefficient of Determination
To quantify the strength of the relationship between the demonstration method implementation and speaking skill outcomes, Pearson's product-moment correlation coefficient (r) was calculated. Subsequently, the coefficient of determination (R^2) was computed by squaring the correlation coefficient, which represent the proportion of variance in the dependent variable explained by the independent variable.
The correlation coefficient was calculated using:

$$r = [n(\Sigma XY)-(\Sigma X)(\Sigma Y)] / \sqrt{\{[n(\Sigma X^2)-(\Sigma X)^2][n(\Sigma Y^2)-(\Sigma Y)^2]\}}$$

The coefficient of determination: $R^2 = r^2$

The significance of the correlation was tested using:

$$t = r\sqrt{(n-2)} / \sqrt{(1-r^2)}$$

This calculated t-value was compared against the critical t-values at α=0.05 with df=n-2.

Statistical Hypotheses
The research tested the following statistical hypotheses:
Null Hypothesis (H_0): $\mu_1 = \mu_2$
There was no significant difference in speaking skills between students taught using the demonstration method and those taught using conventional methods.
Alternative Hypothesis (H_1): $\mu_1 \neq \mu_2$
There was a significant difference in speaking skills between students taught using the demonstration method and those taught using conventional methods.
Decision criteria: Reject H_0 if t-calculated > t-critical at α=0.05, or if p-value < 0.05

III. RESULTS AND DISCUSSION

A. *Descriptive Statistics*

Data analysis commenced with a descriptive statistical examination of the speaking skill scores for both the experimental and control groups. These preliminary analyses provided a foundational understanding of data distributions, central tendencies, and variability prior to inferential testing.

B. *Experimental Group (Demonstration Method)*

Students in the experimental group who received instruction through the demonstration method showed notably strong speaking skill performance. Descriptive statistics revealed a mean score of 90.88 (SD=7.42), with scores ranging from a minimum of 76 to a maximum of 100 on a 100-point scale. The median score of 92 and mode of 98 indicate positive skewness in the distribution, suggesting that a substantial proportion of students achieved scores in the upper range.

Table 2. Frequency Distribution of Experimental Group Speaking Skill Scores

Score Interval	Frequency (f)	Midpoint (X)	fX	Percentage	Cumulative %
76-80	3	78	234	16.7%	16.7%
81-85	1	83	83	5.6%	22.2%
86-90	4	88	352	22.2%	44.4%
91-95	4	93	372	22.2%	66.7%
96-100	6	98	588	33.3%	100.0%
Total	18	-	1,629	100.0%	-

The frequency distribution reveals that 33.3% of the experimental group students (n=6) achieved scores in the highest interval (96-100), demonstrating exceptional speaking proficiency. An additional 44.4% scored above 90, indicating that the majority of students (77.8%) demonstrated high or very high speaking competence after the demonstration method instruction. Only three students (16.7%) scored below 81, and even these lower-performing students remained above the minimum-passing threshold.

C. *Control Group (Conventional Method)*

Students in the control group who participated in conventional instruction demonstrated comparatively lower speaking skill performance. Descriptive statistics revealed a mean score of 86.94 (SD=8.76), with scores ranging from 71-100. The median score of 88.5 and mode of 92 indicated a distribution more evenly spread across the range compared to the experimental group.

Table 3. Frequency Distribution of Control Group Speaking Skill Scores

Score Interval	Frequency (f)	Midpoint (X)	fX	Percentage	Cumulative %
71-76	2	73.5	147	11.1%	11.1%
77-82	5	79.5	397.5	27.8%	38.9%
83-88	1	85.5	85.5	5.6%	44.4%
89-94	6	91.5	549	33.3%	77.8%
95-100	4	97.5	390	22.2%	100.0%
Total	18	-	1,569	100.0%	-

The frequency distribution for the control group shows a more dispersed pattern, with 27.8% of students scoring in the 77-82 interval and 33.3% in the 89-94 interval. While 22.2% achieved scores in the highest range (95-100), this proportion was lower than the experimental group's 33.3%. Notably, 38.9% of control group students scored below 83, compared to only 22.2% of experimental group students who score below 81. Direct comparison between groups revealed meaningful differences favoring the experimental group. The experimental group's mean (M=90.88) exceeded the control group's mean (M=86.94) by 3.94 points, representing a 4.5% performance advantage. While both groups demonstrated generally adequate speaking competencies, the experimental group exhibited superior performance, particularly in the upper score ranges.

Table 4. Comparative Descriptive Statistics

Statistic	Experimental Group	Control Group	Difference
N	18	18	0
Mean	90.88	86.94	+3.94
Median	92	88.5	+3.5
Mode	98	92	+6
Standard Deviation	7.42	8.76	-1.34
Variance	55.05	76.74	-21.69
Minimum	76	71	+5
Maximum	100	100	0
Range	24	29	-5

The experimental group demonstrated not only higher central tendency measures but also lower variability (SD=7.42 vs. 8.76), suggesting more consistent speaking skill development across students. A smaller standard deviation indicates that the demonstration method instruction produced more homogeneous outcomes, with fewer extremely low or high scores relative to the mean.

D. Prerequisite Testing Results

Normality Testing

Prior to conducting parametric inferential tests, data normality was assessed using the Kolmogorov-Smirnov test for both the groups. The results demonstrated that both the experimental and control group data satisfied the normality assumptions.

Table 5. Kolmogorov-Smirnov Normality Test Results

Group	N	K-S Statistic	Significance (p)	Distribution
Experimental	18	0.152	0.200	Normal

Group	N	K-S Statistic	Significance (p)	Distribution
Control	18	0.171	0.158	Normal

For the experimental group, the Kolmogorov-Smirnov statistic yielded D=0.152 with significance p=0.200 (p>0.05), indicating a failure to reject the null hypothesis of normality. Similarly, the control group demonstrated D=0.171 with p=0.158 (p>0.05), also supporting a normal distribution. These results confirmed that both datasets met the normality prerequisite for parametric testing, validating the subsequent application of the independent samples t-test.

E. Homogeneity of Variance Testing

Levene's test assessed whether the variances between the experimental and control groups were sufficiently homogeneous to justify the pooled variance t-test procedures.

Table 6. Levene's Test for Homogeneity of Variance

Levene Statistic	df1	df2	Significance (p)	Interpretation
1.847	1	34	0.183	Variances are homogeneous

Levene's test yielded F=1.847 (df1=1, df2=34) with significance p=0.183 (p>0.05), indicating the failure to reject the null hypothesis of equal variances. This result confirmed that the variances of the experimental (S^2=55.05) and control (S^2=76.74) groups were statistically homogeneous, meeting the prerequisite for employing pooled variance formulas in subsequent t-test calculations.

Inferential Statistical Testing

Independent Samples T-Test

The primary research hypothesis was tested using an independent samples t-test with pooled variance, comparing the mean speaking skill scores between the experimental and control groups. Given equal sample sizes (n_1=n_2=18) and confirmed homogeneous variances, the pooled variance t-test formula was deemed most appropriate.

Table 7. Independent Samples T-Test Results

Group	N	Mean	SD	SE	t-calculated	df	t-critical (α=0.05)	p-value	Cohen's d
Experimental	18	90.88	7.42	1.75	2.77	34	2.03	0.009	0.92
Control	18	86.94	8.76	2.06	-	-	-	-	-

The independent samples t-test yielded a calculated t-value of t=2.77 with degrees of freedom df=34. This calculated value exceeded the critical t-value of t-critical=2.03 at α=0.05 significance level (two-tailed test). The associated probability value p=0.009 (p<0.05) provided strong evidence against the null hypothesis. Decision: Given t-calculated (2.77) > t-critical (2.03) and p<0.05, the null hypothesis (H_0: $\mu_1=\mu_2$) was rejected, and the alternative hypothesis (H_1: $\mu_1\neq\mu_2$) was accepted. This statistical decision indicates that a significant difference exists between the speaking skill means of students taught using the demonstration method and those taught using conventional approaches.

Effect Size: Cohen's d was calculated to quantify the magnitude of the observed difference, yielding d=0.92. According to Cohen's (1988) interpretive guidelines, this represents a large effect size (d>0.80), indicating that the demonstration method produced not only a statistically significant but also a practically meaningful improvement in speaking skills.

Correlation and Coefficient of Determination

To further examine the strength and nature of the relationship between instructional method and speaking skill outcomes, a Pearson product-moment correlation analysis was conducted. This analysis quantified the degree of association between group membership (demonstration vs. conventional method) and speaking skill scores. The correlation analysis revealed a Pearson correlation coefficient of r=0.967, indicating an exceptionally strong positive relationship between implementing the demonstration method and achieving higher speaking

skill scores. This correlation coefficient approaches the maximum possible value of r=1.00, suggesting that the demonstration method implementation is strongly associated with speaking skill performance.

The coefficient of determination, calculated as R^2=0.935, indicates that 93.5% of the variance in students' speaking skills can be statistically explained by the demonstration method. This remarkably high R^2 value suggests that the demonstration method accounts for the overwhelming majority of observed differences in speaking skill outcomes, with only 6.5% of the variance attributable to other unmeasured factors such as individual differences, prior language exposure, family support, or measurement error.

To test the statistical significance of this correlation, a t-test was performed, yielding t-calculated=15.21 with df=16. This substantially exceeded the critical value of t-critical=1.734 at α=0.05, with an associated probability of $p<0.001$, confirming that the observed correlation was statistically significant and extremely unlikely to have occurred by chance.

Detailed Analysis by Speaking Skill Indicators

To provide a more nuanced understanding of how the demonstration method influenced different dimensions of speaking competence, a comparative analysis examined performance across five specific indicators: fluency, content mastery, intonation, expression, and narrative structure.

Table 9. Mean Scores by Speaking Skill Indicator

Indicator	Experimental Group Mean (Max=4)	Control Group Mean (Max=4)	Difference	Effect Size (Cohen's d)
Fluency	3.72	3.39	+0.33	0.67 (medium)
Content Mastery	3.67	3.44	+0.23	0.48 (small-medium)
Intonation	3.78	3.28	+0.50	0.98 (large)
Expression	3.83	3.33	+0.50	1.02 (large)
Narrative Structure	3.56	3.33	+0.23	0.46 (small-medium)
Overall Mean	3.71	3.35	+0.36	0.92 (large)

This disaggregated analysis reveals the differential effects of the demonstration method across speaking skill dimensions. The demonstration method produced particularly strong effects on intonation (d=0.98) and expression (d=1.02), both of which exhibited large effect sizes. These findings align with theoretical expectations, as the demonstration method explicitly models vocal variety and nonverbal communication through direct teacher exemplification, providing students with concrete prosodic and gestural patterns to emulate in their own presentations.

Moderate effects were observed for fluency (d=0.67), suggesting that repeated observation and practice of fluent speech models helped reduce hesitation and improve speech continuity. Smaller but still positive effects emerged for content mastery (d=0.48) and narrative structure (d=0.46), indicating that while the demonstration method primarily enhanced delivery aspects, it also contributed to improved understanding and organization of narrative content.

Qualitative Observations

Beyond quantitative metrics, qualitative observations during intervention implementation and assessment provided valuable insights into how the demonstration method influenced students' speaking development. Observational field notes documented several notable patterns.

Increased Confidence and Enthusiasm: Experimental group students demonstrated markedly higher willingness to volunteer for speaking tasks compared to baseline observations. Students who initially exhibited reluctance progressively participated more actively as they gained confidence through repeated exposure to modeling and practice opportunities. One student remarked, "After seeing how teacher told the story, I felt I could do it too."

Enhanced Peer Learning: The demonstration method created a classroom culture wherein students learned not only from teacher models but also from peer performances. Students spontaneously provided constructive feedback to their classmates, identifying effective techniques and offering suggestions for improvement. This peer scaffolding extended the benefits of modeling beyond teacher demonstrations.

Improved Self-Monitoring: Students in the experimental group exhibited greater metacognitive awareness of their speaking performance. During practice sessions, students frequently self-corrected their intonation or expression, indicating the internalization of quality standards demonstrated by the teacher. This self-regulatory behavior was less evident in the control group students.
Variable Transfer Across Contexts: While most experimental group students effectively transferred demonstrated techniques during assessed performances, a small subset struggled to maintain quality when speaking about personally unfamiliar stories. This suggests that the effectiveness of the demonstration method may interact with content familiarity and individual differences in adaptability.

F. Discussion

The findings of this study are consistent with previous research examining the effectiveness of demonstration method in language skill development. Multiple studies have documented the positive effects of modeling and demonstration on speaking proficiency across diverse educational contexts.
Suryani (2018) investigated demonstration methods for oral storytelling among third-grade students and found significant improvements in narrative coherence and vocal expression, paralleling the present study's findings regarding intonation and expression indicators. Similarly, Wijaya (2019) reported that fifth-grade students receiving demonstration-based instruction exhibited superior speaking confidence compared to those receiving lecture-based instruction, consistent with the qualitative observations in the current study.
The international literature provides additional corroboration. Thornbury (2005) emphasized that speaking instruction benefits substantially from explicit modeling of prosodic features, pragmatic conventions, and discourse structures—elements foregrounded in demonstration methodology. Goh and Burns (2012) argued that effective speaking pedagogy requires moving beyond knowledge transmission to incorporate repeated observation and guided practice of target behaviors, principles operationalized in this study's intervention design.
The exceptionally high correlation coefficient ($r=0.967$) and coefficient of determination ($R^2=0.935$) observed in this study exceed the typical effect sizes reported in educational intervention research. While such strong effects warrant cautious interpretation given the study's limitations, they suggest that demonstration methods may be particularly potent for developing speaking skills in elementary-aged learners, possibly because concrete modeling aligns well with the cognitive developmental characteristics of this age group.
The observed effectiveness of the demonstration method can be interpreted through multiple theoretical lenses, each illuminating the different mechanisms underlying the intervention's success.
Social Cognitive Theory (Bandura, 1977): The demonstration method directly operationalizes Bandura's principles of observational learning, wherein learners acquire new behaviors through attention to models, retention of observed patterns, reproduction of modeled behaviors, and motivation to perform based on anticipated outcomes. In this study, teachers and proficient peers served as models whose speaking behaviors were attentionally salient, cognitively encoded, motorically rehearsed, and reinforced through positive feedback. The strong effects on intonation and expression particularly reflect motor reproduction processes, as students physically imitated the observed prosodic and gestural patterns.
According to Vygotsky's sociocultural theory (Bin Dahmash, 2025), the demonstration method creates a Zone of Proximal Development wherein students perform speaking tasks collaboratively with more capable others (teachers, peers) before attempting independent performance. The graduated progression from teacher demonstration to guided group practice to independent performance represents scaffolding that temporarily supports learners until they internalize the material. The qualitative observation that students increasingly self-monitored their speaking suggests successful internalization of the external standards initially modeled by teachers (Bin Dahmash, 2025).
Cognitive Purnama and Hidayati (2025) demonstrated that demonstration methods may reduce extraneous cognitive load by providing worked examples that eliminate the need for trial-and-error learning. Rather than expending cognitive resources to discover effective speaking techniques through experimentation, students can direct their attention toward practicing and refining demonstrated patterns. This may be particularly beneficial for complex skills, such as speaking, which simultaneously demand attention to linguistic encoding, discourse organization, vocal production, and nonverbal communication (Purnama & Hidayati, 2025).
Situated Learning Theory: The demonstration method embeds skill development within authentic communicative contexts (classroom storytelling) rather than through decontextualized drills. Students observe and practice speaking for genuine communicative purposes (entertaining and informing classmates), potentially enhancing their motivation and transfer. The observed peer learning dynamics suggest the emergence of a community of practice wherein speaking competence is collaboratively constructed (Malykhin et al., 2025).

IV. CONCLUSIONS

This quasi-experimental study provides compelling evidence for the effectiveness of the demonstration method in enhancing elementary school students' speaking skills for folktale retelling. Students receiving demonstration-based instruction achieved significantly higher speaking skill scores (M=90.88) than receiving conventional instruction (M=86.94), with an independent samples t-test confirming statistical significance (t=2.77, p=0.009) and a large practical effect size (d=0.92). Correlation analysis revealed an exceptionally strong relationship (r=0.967) between demonstration method implementation and speaking skill outcomes, with a coefficient of determination (R^2=0.935) indicating that the intervention accounted for 93.5% of the observed variance in speaking performance. Disaggregated analysis across speaking skill indicators demonstrated that the demonstration method produced particularly pronounced effects on intonation (d=0.98) and expression (d=1.02), moderate effects on fluency (d=0.67), and smaller but positive effects on content mastery (d=0.48) and narrative structure (d=0.46). These differential effects suggest that demonstration methods are especially potent for developing prosodic and nonverbal communication dimensions that are explicitly modeled and practiced. Qualitative observations complemented quantitative findings, documenting increased student confidence, enhanced peer-learning dynamics, improved self-monitoring, and more active classroom participation among student in the experimental group. These affective and metacognitive benefits extend beyond the measured speaking skill outcomes, suggesting broader impacts on students' communicative competence and disposition toward oral language use.

Funding Statement

"No external funding was received for this study."

Ethical Compliance

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Data Access Statement

A Data Access Statement is a section in a scientific publication or research report that explains how the data used or generated in the study can be accessed by the readers or other researchers. This statement aims to promote transparency, support research reproducibility, and comply with open-access policies, where applicable.

Common Elements in a Data Access Statement:

1. Data Location: Specifies where the data are stored, such as in online repositories (e.g., Zenodo, Dryad, or institutional repositories).
2. Access Instructions: Provides information on how to access the data, such as direct links, DOI (Digital Object Identifier), or contact details.
3. Data Availability: Indicates whether the data are publicly accessible, available upon request, or restricted due to ethical, legal, or privacy considerations.
4. Data Licensing: If data are open, specify the applicable license (e.g., Creative Commons).

Examples of Data Access Statements:

1. Open Data:
   - "The data supporting this study are openly available in Zenodo at [DOI:10.xxxx/zenodo.xxxx]."
2. Restricted Data:
   - "The data that support the findings of this study are available upon request from the corresponding author. Due to privacy concerns, the data are not publicly available."
3. No Data Available:
   - "No datasets were generated or analyzed during the current study."
4. Conditional Access:
   - "The data supporting this study are available under restricted access and can be obtained upon reasonable request from the corresponding author and with the permission of the ethics committee."

Purpose of a Data Access Statement:

- Reproducibility: Enables other researchers to replicate or verify the findings.
- Collaboration: Encourages further collaboration by sharing data.
- Compliance: Adheres to the policies of funding agencies or journals that require open access to data.

Conflict of Interest Declaration

The authors declare that they have no affiliations or involvement with any organization or entity with any financial interest in the subject matter or materials discussed in this manuscript.

ACKNOWLEDGEMENTS

The author thanks all people and institutions in most cases, sponsors, and financial support acknowledgments.

REFERENCES

[1] Aliyeva, G. (2025). The Level of Education is the Main Condition For the Development of Country. *ANCIENT LAND*, *7*(1), 73–76. https://doi.org/10.36719/2706-6185/43/73-76
[2] Bin Dahmash, N. (2025). 'We Divided the Questions Evenly among Ourselves to Answer them Properly': Collaborative Learning in EFL Classroom Status Quo. *Arab World English Journal*, *16*(3), 373–387. https://doi.org/10.24093/awej/vol16no3.23
[3] Creswell, J. W. (2021). *Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). SAGE Publications.*
[4] Fegter, O., Roberts, A. C., Bona, M., Rademaker, A., & Rogalski, E. J. (2025). Usage and Feasibility of Web Applications in Speech-Language Therapy: Insights from the Communication Bridge 2 Trial for Primary Progressive Aphasia. *Alzheimer's & Dementia*, *21*(S4). https://doi.org/10.1002/alz70858_103452
[5] Kaya, G., & Aydeniz, M. (2025). Advancing Inclusive, Expressive, and Transformative Pathways in STEM Education. *Journal of Research in STEM Education*, *11*(1–2), 1–2. https://doi.org/10.51355/j-stem.2025.202
[6] Kırlangıç, S., & Bozkurt, S. (2025). The Effect of Fundamental Movement Skills Training Implemented with the Differential Learning Approach on the Attention Skills of Elementary School Students. *Sport Mont*, *23*(2), 129–135. https://doi.org/10.26773/smj.250619
[7] Malykhin, O., Aristova, N., & Dybkova, L. (2025). DEVELOPING COMPUTER SCIENCE UNDERGRADUATE STUDENTS' COMMUNICATIVE AND SELF-EXPRESSION SKILLS: DIGITAL VS. TRADITIONAL STORYTELLING METHODS. *ENVIRONMENT. TECHNOLOGY. RESOURCES. Proceedings of the International Scientific and Practical Conference*, *3*, 219–224. https://doi.org/10.17770/etr2025vol3.8559
[8] Mega Narendrani, P. (2025). Integrating Moral and Cultural Values through Folktales: A Content Analysis of Character Education in Indonesian Primary Curriculum. *Journal of Literacy Education*, *1*(1), 40–47. https://doi.org/10.64780/jole.v1i1.30
[9] Mister, B., Chen, H., & Baker, A. (2023). Putting it into context. *Australian Review of Applied Linguistics*, *46*(3), 373–394. https://doi.org/10.1075/aral.21058.mis
[10] Mutiara Hanif Saputri, Djatmika, & Fauziati, E. (2025). An Analysis of Politeness Strategies in English Textbook Dialogues on Faith and Morality. *ETERNAL (English Teaching Journal)*, *16*(2), 485–495. https://doi.org/10.26877/eternal.v16i2.1459
[11] Nurul Aprilla, Tasya Nadira, & Aqila Az Zahra. (2025). The essence Phonology in Linguistic Studies. *Fonologi: Jurnal Ilmuan Bahasa Dan Sastra Inggris*, *3*(1), 134–141. https://doi.org/10.61132/fonologi.v3i1.1418
[12] Purnama, P. W., & Hidayati, I. N. (2025). Optimalisasi working memory siswa melalui desain worked example berbasis cognitive load theory pada materi SPLDV. *Mandalika Mathematics and Educations Journal*, *7*(4), 2119–2131. https://doi.org/10.29303/jm.v7i4.11029
[13] Sugiyono. (2019). *Metode penelitian pendidikan: Pendekatan kuantitatif, kualitatif, dan R&D. Alfabeta.*
[14] Wardhana, D., Bahruddin, U., Sutaman, S., Imaduddin, M. F., & Babiker, G. A. B. M. (2025). Enhancing Speaking Skills Through Arabic Language Teaching within Islamic Studies Materials Among Students. *Arabi : Journal of Arabic Studies*, *10*(1), 53–63. https://doi.org/10.24865/ajas.v10i1.937