

The Effect of Group Work on Students' Speaking Skills

 <https://doi.org/10.31004/jele.v11i1.1745>*Andika Firdaus Hidayat, Mochtar Marhum, Nadrun, Abdul Kamaruddin^{abcd}¹²³⁴Universitas Tadulako, Indonesia*Corresponding Author: andikafirdauslakwood@gmail.com

ABSTRACT

Many seventh-grade students struggle with speaking skills due to limited linguistic knowledge and speaking anxiety, despite the expectations of the Merdeka Curriculum for junior high school learners to communicate effectively in English. To address these challenges, group work is often recommended as a strategy to reduce anxiety and increase student participation. However, most previous studies were conducted at higher education levels, leaving limited evidence for younger learners. Therefore, this study aims to investigate whether group work significantly improves the speaking skills of seventh-grade students at SMP Negeri 18 Palu. A quasi-experimental design was used, involving two intact classes of 28 students each. Both groups completed a pre-test and post-test consisting of three oral questions assessing fluency and comprehensibility. The treatment will be administered over six sessions, from the second to the seventh meeting. The experimental group will use the group work technique, while the control group will receive conventional instruction. The data were analyzed using the Mann-Whitney U Test because the scores were not normally distributed. The findings showed that although both groups demonstrated increased mean scores from pre-test to post-test, the significance value ($p = 0.113$) exceeded the 0.05 threshold. Thus, the null hypothesis (H_0) was accepted, indicating that group work did not produce a statistically significant improvement in students' speaking skills. These results suggest that group work may require longer implementation or additional support to yield stronger effects for junior high school learners.

Keywords: *Group Work, Speaking Skills, Merdeka Curriculum***Article History:**Received 20th November 2025Accepted 15th January 2026Published 18th January 2026

INTRODUCTION

Speaking is an essential skill that enables students to express ideas clearly and reduce misunderstandings. (Heaton, 1988) states that the components of speaking include accuracy, fluency, and comprehensibility, all of which influence the quality of communication. These components help speakers convey messages clearly and ensure that the messages are easily understood. In Indonesia, under the Merdeka Curriculum, junior high school students are expected to use English for interaction and information exchange. However, many seventh-grade students at SMP Negeri 18 Palu still struggle with speaking, particularly in terms of fluency and comprehensibility. Preliminary observations and interviews indicate that limited linguistic knowledge often leads to speaking anxiety, reduced confidence, and difficulty conveying messages effectively.

One instructional technique that may address this issue is group work, a collaborative approach that encourages students to exchange ideas and use language actively. (Harmer, 2007) describes group work as a method in which learners collaborate and use self-initiated language to complete tasks. Furthermore, (Ur, 1996) emphasizes that group work increases the volume of learners' conversations in a limited amount of time and reduces the barriers for learners who are reluctant to speak in front of the whole class. In addition (Harmer, 2007) states that group work offers several advantages in classroom implementation, including increased opportunities for students to speak, reduced interpersonal tension due to the presence of more group members, and the development of cooperative and negotiation skills.

This view is further supported by Cooperative Learning Theory proposed by (Johnson & Johnson, 1999) who argue that positive interdependence, individual accountability, promotive interaction, and group processing contribute to increased learner engagement and improved learning outcomes. When students work cooperatively, they support one another, negotiate meaning, and participate more equally, factors that can enhance speaking performance.

Furthermore, Vygotsky's Sociocultural Theory emphasizes that learning is a social process strongly influenced by interaction, culture, and language. Cognitive development occurs when learners collaborate with more knowledgeable others within their Zone of Proximal Development (ZPD), and gradually internalize the knowledge gained through these interactions (as cited in Mcleod, 2025).

Through group work, students have opportunities to scaffold each other's learning, exchange linguistic knowledge, and co-construct meaning, which can reduce anxiety and facilitate language development. These theoretical perspectives highlight the potential of group work to improve speaking skills, particularly in providing a low-pressure environment for practice and peer support.

Several previous studies support the effectiveness of group work in enhancing speaking skills. (Rospinah et al., 2021) found that group work activities significantly improved students' accuracy and fluency. (Muhammad et al., 2024) also reported notable improvements when group work was aligned with students' interests. Similarly, research by (Fahrurrozi et al., 2019) demonstrated that group work had a significant positive impact on students' speaking performance. Although previous studies focused on higher education, their findings may not fully apply to junior high school contexts due to differences in students' linguistic development and learning needs. Therefore, this research aims to examine whether group work significantly improves the speaking skills of seventh-grade students at SMP Negeri 18 Palu, with specific attention to fluency and comprehensibility.

METHOD

This research employs a quasi-experimental method using a Nonequivalent Control Group Design. Two groups are involved in this design: an experimental group and a control group. Both groups received a pre-test and a post-test. The experimental group received a treatment, which involved the application of the group work technique during the teaching process, whereas the control group was taught using a conventional method.

The treatment was administered after the pre-test and conducted over six meetings across three weeks. Each meeting lasted 2×40 minutes (80 minutes) and covered the topics Greetings & Partings, Introducing Yourself, and Likes & Dislikes. During the lessons, students in the experimental group engaged in group discussions, collaborated during the learning activities, and worked together to complete assignments and present their work. The teacher's role was limited to guiding and supervising the class.

The population of this research consisted of the seventh-grade students of SMP Negeri 18 Palu, comprising five parallel classes. Each class contained 32 students, resulting in a total population of 160 students. This research used the purposive sampling technique to select the sample. According to (Sugiyono, 2013), purposive sampling is a method of selecting samples based on specific criteria or considerations that have been determined in advance. Based on the English teacher's recommendation, classes VII A and VII D were selected as the experimental and control groups because they had comparable English proficiency and demonstrated low speaking skills.

This research uses tests as the instrument for data collection. During both the pre-test and post-test, each student is given three oral questions. The questions are open-ended and functional, requiring students to produce simple communicative responses. The oral test instrument has construct validity because each test item was developed based on speaking skill theory. (Heaton, 1988) states that speaking ability consists of several components such as accuracy, fluency, and comprehensibility. These components serve as the foundation for developing the scoring system. In its assessment, this instrument specifically evaluates the

aspects of fluency and comprehensibility in accordance with Heaton's scoring concept. Furthermore, the instrument was validated through expert judgment by asking an English language expert to review each test item and the scoring system to ensure that both accurately reflect the speaking construct being measured.

Inter-rater reliability (IRR) was used to assess the reliability of the instrument through the Intraclass Correlation Coefficient (ICC), which aimed to determine the level of consistency between two raters in scoring the same test. Based on the results of the analysis, the ICC values of the pre-test for the experimental and control groups were 1.000 and 0.975, respectively, and the ICC values of the post-test for the experimental and control groups were 0.997 and 0.981, indicating that the instrument used in this research demonstrated good reliability."

Table 1. ICC Pre-Test of Experimental Group

	Intraclass Correlation ^b	Intraclass Correlation Coefficient		F Test with True Value 0			
		95% Confidence Interval		Value	df1	df2	Sig
		Lower Bound	Upper Bound				
Single Measures	1,000 ^a	.	.	.	27	.	.
Average Measures	1,000 ^c	.	.	.	27	.	.

Table 2. ICC Pre-Test of Control Group

	Intraclass Correlation ^b	Intraclass Correlation Coefficient		F Test with True Value 0			
		95% Confidence Interval		Value	df1	df2	Sig
		Lower Bound	Upper Bound				
Single Measures	,951 ^a	,896	,977	41,452	27	27	,000
Average Measures	,975 ^c	,945	,988	41,452	27	27	,000

Table 3. ICC Post-Test of Experimental Group

	Intraclass Correlation ^b	Intraclass Correlation Coefficient		F Test with True Value 0			
		95% Confidence Interval		Value	df1	df2	Sig
		Lower Bound	Upper Bound				
Single Measures	,993 ^a	,986	,997	314,923	27	27	,000
Average Measures	,997 ^c	,993	,998	314,923	27	27	,000

Table 4. ICC Post-Test of Control Group

	Intraclass Correlation ^b	Intraclass Correlation Coefficient		F Test with True Value 0			
		95% Confidence Interval		Value	df1	df2	Sig
		Lower Bound	Upper Bound				
Single Measures	,962 ^a	,919	,982	54,974	27	27	,000
Average Measures	,981 ^c	,958	,991	54,974	27	27	,000

Furthermore, the researcher records the students' responses using a voice recording application. These recordings are used for scoring purposes, particularly to assess each student's speaking skill with a focus on fluency and comprehensibility. The researcher will use the following scoring system:

Table 5. Scoring system

Rating	Fluency	Comprehensibility
4	Although he has to make an effort and search for words, there are not too many unnatural pauses. Fairly smooth delivery mostly. Occasionally fragmentary but succeeds in conveying the general meaning. Fair range of expression.	Most of what the speaker says is easy to follow. His intention is always clear but several interruptions are necessary to help him to convey the message or to seek clarification.
3	Has to make an effort for much of the time. Often has to search for the desired meaning. Rather halting delivery and fragmentary. Range of expression often limited	The listener can understand a lot of what is said, but he must constantly seek clarification. Cannot understand many of the speaker's more complex or longer sentences.
2	Long pauses while he searches for the desired meaning. Frequently fragmentary and halting delivery. Almost gives up making the effort at times. Limited range of expression.	Only small bits (usually short sentences and phrases) can be understood - and then with considerable effort by someone who is used to listening to the speaker.
1	Full of long and unnatural pauses. Very halting and fragmentary delivery. At times gives up making the effort. Very limited range of expression.	Hardly anything of what is said can be understood. Even when the listener makes a great effort or interrupts, the speaker is unable to clarify anything he seems to have said.

The researcher will apply the scoring range adopted from the KTTP of the Merdeka Curriculum to rank students' scores as follows:

Table 6. The Scoring range

Rating	Category	Scoring range	Qualification
4	Excellent	85-100	Successfull
3	Good	71-85	Successfull
2	Fair	60-70	Failed
1	Poor	0-59	Failed

The score obtained will be converted as seen below:

$$\text{Student score} = \frac{\text{Obtained score}}{\text{Maximum score}} \times 100$$

Statistical analysis was conducted to examine the pre-test and post-test data, including normality and hypothesis testing. Since the results of the Shapiro-Wilk test indicated that the data were not normally distributed, the Mann-Whitney U test was used to test the hypothesis. All analyses were performed using SPSS version 24.

FINDINGS AND DISCUSSION

Findings

The research findings present the results of the pre-test and post-test, which were administered and analyzed statistically. Descriptive statistical analysis was conducted to determine the range, minimum, maximum, mean, and standard deviation, and the results are shown in the table below.

Table 7. Descriptive Statistics of Pre-test Result

	N	Range	Minimum	Maximum	Mean	Std. Deviation
Pre-Test Eksperimental	28	62	25	87	39.14	21.258
Pre-Test Control	28	62	25	87	45.39	23.478

The descriptive statistics of the pre-test scores for both groups show that the experimental and control groups consisted of 28 students each. Both groups had the same score distribution, with a range of 62, a minimum score of 25, and a maximum score of 87, indicating a wide spread of students' initial speaking abilities. The mean score of the experimental group was 39.14, which was lower than that of the control group (45.39). This suggests that the control group had slightly better initial performance prior to the treatment. The standard deviations were relatively high in both groups (21.258 for the experimental group and 23.478 for the control group), showing substantial variability in students' initial abilities. Overall, the descriptive data indicate that the two groups were comparable at the beginning of the research, although the control group demonstrated a somewhat higher average score.

Table 8. Descriptive Statistics of Post-test Result

	N	Range	Minimum	Maximum	Mean	Std. Deviation
Post-Test Eksperimental	28	75	25	100	58.43	28.637
Post-Test Control	28	75	25	100	69.89	20.165

The descriptive statistics of the post-test scores show that both the experimental and control groups experienced an increase in their scores. Both groups consisted of 28 students and shared the same score distribution, with a range of 75, a minimum score of 25, and a maximum score of 100, indicating a wide spread of students' speaking performance after the treatment. The mean score of the experimental group was 58.43, whereas the control group achieved a higher mean score of 69.89. This suggests that the control group performed better in the post-test compared to the experimental group. The standard deviation of the experimental group (28.637) was higher than that of the control group (20.165), showing that the experimental group had greater variability in their post-test performance. Overall, the descriptive data indicate that although both groups exhibited similar score ranges, the control group demonstrated a higher and more consistent level of achievement in the post-test.

Based on the comparison of the pre-test and post-test scores, it can be concluded that both groups experienced improvement in learning outcomes, as indicated by the increase in their scores shown in the pre-test and post-test tables. However, the control group showed a higher

mean increase and a more stable distribution of scores. Although the experimental group also improved, the wider score range and higher standard deviation indicate that the gains were uneven, with some students making substantial progress while others remained at lower levels. To determine whether the difference in improvement between the two groups is statistically significant, further analysis using an appropriate significance test is required.

Normality testing was conducted to determine whether the pre-test and post-test data of both groups were normally distributed or not. The Shapiro-Wilk test was used with the following criteria: if the significance value (Sig.) ≥ 0.05 , the data are normally distributed; whereas if the significance value ≤ 0.05 , the data are not normally distributed. The results are presented as follows:

Table 9. Tests of Normality

Group	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Learning outcomes Pre-Test Experimental	.354	28	.000	.698	28	.000
Post-Test Eksperimental	.236	28	.000	.840	28	.001
Pre-Test Control	.272	28	.000	.795	28	.000
Post-Test Control	.314	28	.000	.845	28	.001

a. Lilliefors Significance Correction

Based on the results of the Shapiro-Wilk normality test, both groups obtained significance values of .000 for the pre-test and .001 for the post-test. Therefore, the data in this research were not normally distributed, as the significance values were less than 0.05. Consequently, the researcher used a non-parametric test, namely the Mann-Whitney U test, to determine whether there was a significant difference between the two independent groups. The results of the Mann-Whitney U test, including the Ranks and Test Statistics, are presented as follows:

Table 10. Ranks

Groups	N	Mean Rank	Sum of Ranks
Learning outcomes Experimental group	28	25.14	704.00
Control group	28	31.86	892.00
Total	56		

Based on the "ranks" output, it was found that the experimental group consisting of 28 students had a mean rank of 25.14 with a sum of ranks of 704.00. Meanwhile, the control group, which also consisted of 28 students, had a mean rank of 31.86 and a sum of ranks of 892.00.

Table 11. Test Statistics^a

Learning outcomes	
Mann-Whitney U	298.000
Wilcoxon W	704.000
Z	-1.586
Asymp. Sig. (2-tailed)	.113

a. Grouping Variable: Groups

Based on the "Test Statistics" output, the Mann-Whitney U value was found to be 298.000, the Z value was -1.586, and the Asymp. Sig. (2-tailed) value was 0.113. According to the decision criteria, if the Asymp. Sig. value is less than 0.05, the alternative hypothesis (H_a) is accepted; whereas if the Asymp. Sig. value is greater than 0.05, the null hypothesis (H_0) is accepted. Since the Asymp. Sig. value of 0.113 is greater than 0.05, the null hypothesis (H_0) is accepted. Therefore, it can be concluded that there is no statistically significant difference in learning outcomes between the experimental group and the control group.

Discussion

This research aims to investigate whether the use of group work significantly improves the English-speaking skills of seventh-grade students at SMP Negeri 18 Palu. The experimental group received group work treatment for six meetings, while the control group was taught using conventional methods. The research focused on students' fluency and comprehensibility. To measure these aspects, the researcher administered a pre-test consisting of three oral questions to both groups before the treatment and a post-test with three oral questions after the treatment to assess students' improvement in speaking skills.

Moreover, after the researcher calculated the pre-test and post-test scores of both groups, it was found that there was an increase in the average scores from the pre-test to the post-test. In the experimental group, the mean score increased from 39.14 to 58.43, with a gain of 19.29 points. Meanwhile, in the control group, the mean score increased from 45.39 to 69.89, showing an improvement of 24.50 points. Therefore, it can be concluded that both groups experienced improvement; however, the control group demonstrated a higher mean increase and a more stable score distribution compared to the experimental group.

In the experimental group, although the average score increased, the wider spread of scores indicates that the improvement in students' speaking skills was uneven. Some students showed significant progress, while others remained at a lower level. This condition may be influenced by several factors, such as differences in students' speaking abilities, lack of participation from some students during treatment sessions, and students being more focused on interacting socially rather than academically within their groups.

Therefore, the differences in speaking skills and student activity in each group resulted in active and passive groups, causing uneven improvement in students' speaking skills despite the increase in average scores from the pre-test to the post-test. Nevertheless, the level of students' engagement during dialogue activities increased, indicating an improvement in their speaking comprehension. Most of the students became more active and confident in expressing themselves, leading to greater fluency and reduced anxiety while speaking. These findings suggest a decrease in speaking anxiety and an increase in students' self-confidence.

As a result, the students' speaking skills improved. Similarly, (Fahrurrozi et al., 2019) found that the use of group work techniques was effective in improving students' speaking skills based on their data analysis. Finally, to determine whether the difference in improvement between the two groups was statistically significant, the researcher conducted a Mann-Whitney U test. The result showed that the Asymp. Sig. (2-tailed) value was 0.113, which is greater than 0.05. Therefore, the null hypothesis (H_0) was accepted, indicating that the use of group work did not significantly improve the English-speaking skills of seventh-grade students at SMP Negeri 18 Palu.

CONCLUSIONS

After calculating and analyzing the data, it can be concluded that although there was an increase in the mean scores from the pre-test to the post-test, the use of group work did not significantly improve the English-speaking skills of seventh-grade students at SMP Negeri 18 Palu. Based on the results of the Mann-Whitney U test, there was no statistically significant difference in learning outcomes between the experimental and control groups. Based on the research findings, the researchers provided several suggestions and implications. First, teachers may use group work as an alternative strategy to reduce students' anxiety in speaking English, as learners tend to feel more comfortable discussing ideas with their peers. Teachers are also advised to form groups randomly without involving students in the selection process. This approach helps prevent the formation of groups based on personal preferences or similar ability levels, thereby reducing the potential gap between active and passive groups. Second, this research may serve as a reference for future researchers who wish to explore speaking skills or investigate the use of group work techniques in English language teaching. Future researchers are also encouraged to conduct studies with longer treatment durations to gain deeper insights into the effectiveness of group work in improving students' speaking performance.

REFERENCES

- Fahrurrozi, F., Waluyo, U., & Soepriyanti, H. (2019). The Effectiveness of Group Work Technique in Improving the Students' Speaking Skill a Survey at Second Semester Students of English in Academic Years 2017-2018. *International Journal of Multicultural and Multireligious Understanding*, 6(2), 280-288. <https://doi.org/10.18415/ijmmu.v6i2.647>
- Harmer, J. (2007). *The Practice of English Language Teaching* (Fourth). Longman.
- Heaton, J. B. (1988). *Writing English Language Test*. Longman group limited.

- Johnson, D. W., & Johnson, R. T. (1999). *Making cooperative learning work. Theory into Practice*. 38(2), 67–73. <https://doi.org/https://doi.org/10.1080/00405849909543834>
- Mcleod, S. (2025, March 18). *Vygotsky's Theory of Cognitive Development*. Simply Psychology. Retrieved December 14, 2025, from <https://www.simplypsychology.org/vygotsky.html>
- Muhammad, A., Navriansyah, A., & Aeni, N. (2024). The Use of Group Work Based on Interest to Improve Students' English Speaking Skill at SMA Negeri 3 Makassar. *Performance: Journal of English Education and Literature*, 3(1), 35–50.
- Rospinah, R., Ampa, A. T., & Nappu, S. (2021). The Effect of Group Work Activities to Improve Students' Speaking Skill. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 9, 120–129. <https://doi.org/10.24256/ideas.v8i2.1660>
- Sugiyono. (2013). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Alfabeta.
- Ur, P. (1996). *A Course in Language Teaching Practice and Theory* (M. Williams & T. Wright, Eds.). Cambridge University Press.