ELTR Journal, e-ISSN 2579-8235, Vol. 8, No. 1, January 2024, pp. 18-32
English Language Teaching and Research Journal
http://apspbi.or.id/eltr
English Language Education Study Program Association, Indonesia

THE EFFECTIVENESS OF PEER FEEDBACK THROUGH INSTAGRAM ON TENTH-GRADE STUDENTS' SPEAKING SKILLS IN AN EFL HIGH SCHOOL

Renita[1] and Rifki Irawan[2*]
[1,2]Universitas PGRI Yogyakarta, Indonesia
nre94277@gmail.com[1] and rifkizam@upy.ac.id[2]
*correspondence: rifkizam@upy.ac.id
https://doi.org/10.37147/eltr.v8i1.181
received 29 November 2023; accepted 19 January 2024

Abstract
In the flow of learning objectives (ATP), students are required to master speaking skills, especially in the aspects of pronunciation, fluency, and accuracy. Students need help getting it, and one of the causes is the need for more effective learning media. This research aims to find out the effectiveness of peer feedback through Instagram on tenth-grade students' speaking skills at SMAN 1 Gamping. This research was carried out at SMAN 1 Gamping and used quantitative methods with a quasi-experimental design. Class X B is the control class that is treated with traditional peer feedback. In contrast, class X C is the experimental class that is treated with peer feedback through Instagram. The instrument used was a test (speaking skill) with the following research steps: pre-test – treatment – post-test. The hypothesis test used is the t-test with = 5% (0.05). The paired sample t-test proves that peer feedback through Instagram is effective in improving tenth-grade students' speaking skills at SMAN 1 Gamping. The independent sample proves that traditional peer feedback is superior and is more recommended for implementation.

Keywords: Instagram, peer feedback, speaking skill, traditional

Introduction
English is an international language. In Indonesia, English has various potentials if used in several fields, including business and economics, international relations, understanding international news, communication with foreigners, and education (Lauder, 2008). It means that English subjects must be taught in school to adapt to the increasingly advanced outside world (Alfarisy, 2021).

In the grade 10 sequence and learning objectives, students are expected to be able to master the three learning outcomes of phase E, namely listening and speaking, reading and understanding, and writing and presenting (Usmirawati, 2021). According to Brown (2004, p. 173), in speaking ability, there are five points of assessment, namely grammar, vocabulary, comprehension, fluency, and pronunciation. In fact, in learning speaking skills, several problems have been found that arise in class X students of SMAN 1 Gamping. The problem is that the student needs to take more time to remember or think about the following sentence,

often mispronounces words, and rarely practices because they only use English during lessons.

There are many questions about 'what makes a good teacher?' What makes a good teacher is that the teacher must make teaching and learning activities enjoyable for students. A teacher must create an active classroom atmosphere that students can enjoy (Harmer, 1998, p. 1). Teachers in Indonesia often need to be more precise in teaching methods or learning media, especially in speaking skills (Wahyuningsih & Afandi, 2020). The reality that the researcher encountered during teaching practice, according to students, was that the learning media provided had less interest in learning, and for practice learning to improve speaking skills still needed to be improved.

When teaching speaking skills, students are instructed to show their speaking skills. Giving speaking assignments is suitable to make it easier for teachers to monitor students. The things that underlie the idea of giving oral tasks are rehearsal, feedback, and engagement. Trial means the teacher provides students the opportunity to discuss outside the classroom. Feedback on assignment results can come from teachers and peers, and students can confidently find things they can improve in their speaking ability (Harmer, 1998, pp. 87–88). On the other hand, the researcher saw that the tenth-grade students in the class were rarely allowed to give feedback on their peers' work.

According to Berk (2012), teenagers have a better capacity than adults to see a person's strengths and weaknesses. From this, in terms of idealism and criticism, a teenager can provide significant benefits. They respond to create better conditions. On the other hand, researchers saw that tenth-grade students had anxiety or doubts when giving comments or suggestions to their peers. Students also have anxiety in speaking because they are afraid of making mistakes and are afraid of being laughed at by their friends. That condition is very contrary to the conditions that should be, as stated by Tanjung and Amelia (2017) self-confidence is beneficial for daily life. Confidence is needed to give meaning in doing everything including communicating so that the message can be better conveyed to the other person.

In this research, researchers chose to use Instagram as a learning medium for learning activities. Instagram can be an innovation and a valuable tool for improving students' speaking skills. That is in line with the results of previous research that examined the impact of Instagram on students' language skills. Handayani (2016) states that Instagram can be an innovative and valuable tool for improving students' language skills, such as speaking skills. It is said to be an innovative tool because Instagram can be used in many activities; for example, students can post videos of speaking practice, and peers or followers can give feedback in the comments column. The results of that research were reaffirmed by Ishak and Yaacob (2022), who stated that uploading videos on Instagram and peer feedback focused on speaking skills, resulted in the finding that Instagram is the right tool for practicing speaking skills and a tool for interacting with peers that is fun and flexible in its implementation.

Based on existing phenomena in the field, demands for student achievement in the learning objectives (ATP), and previous research concerning the impact of peer feedback through Instagram on improving students' language skills, the researcher is interested in finding out more to investigate the effectiveness of peer

feedback on speaking skills through video assignments uploaded on Instagram with a focus on the text used, which is descriptive text.

Literature Review

Speaking skill

Speaking is not the oral output of written language. In speaking, students are required to master various sub-skills, which, when combined, will result in overall competence in spoken language (McDonough et al., 2013, p. 156). Speaking refers to the ability on which someone is assessed "at face value." In other words, people may rate our English proficiency based on aspects of our speech skills more than any other language skill (McDonough et al., 2013, p. 156).

Feedback

Feedback is an activity where someone gives their response to the performance results of others according to their understanding, and this aims to improve the performance results of others to a better level by correct knowledge. Improved performance results are considered better when using corrective feedback types than those that only provide feedback in the form of praise, value, or punishment (Wisniewski et al., 2020). According to Hyland and Hyland (2006), there are two types of feedback: written and oral.

Traditional peer feedback

Many ways can be done to implement peer feedback, one of which is the interaction of peer feedback in the classroom without incorporating elements of technology. The goal of this peer feedback is to improve their understanding and the quality of the work. Traditional peer feedback involves students actively participating in conferences. Peer feedback can be done directly by providing comments or suggestions on paper (Lei, 2017).

Peer feedback through social media

According to Irawan and Wiyanah (2021), in teaching and learning one of the language skills, students feel interested in peer feedback through online platforms because of the ease of correcting peers' work and getting feedback from peers. This statement is in line with Pikhart and Botezat (2021), who state that Generation Z, or generations born after the 2000s, wants fast interaction on the e-learning platform; the interaction in question is communicating and giving each other input between friends/peer feedback with short waiting time. They suggest that teachers use several tools to make it easier for them to interact in learning, for example, Facebook, Instagram, YouTube, and Twitter. This type of learning is also called eLearning 3.0 because it deals with machines and is time efficient with the help of mobile phones.

Theory of teenager character

Teenagers in their teens are pretty successful in learning languages because of their remarkable ability, persistence, and strong determination to do something. These teenagers are earnest when they participate, and they are serious because they already have an awareness of the future and the burden they carry. However, beyond its advantages in learning languages, we also have to be aware that at the

age of their teens, the child is still developing mentally. The teacher is considered to have to understand the students first so that students feel comfortable and are not humiliated. In addition to the attention that the teacher must give, teenagers are also very concerned about the opinions of their peers (Harmer, 2007, p. 83).

The "Merdeka" curriculum

A "Merdeka" curriculum is an educational concept that aims to give more freedom to schools, teachers, and students. The freedom of teachers and schools lies in determining curriculum content that is relevant to students' needs. The freedom of students with the opportunity to study in a learning atmosphere that students are interested in, for example, fun and free from stress and pressure. Freedom in learning is expected so that students can show their natural talents and form a superior generation while still upholding national values (Ainia, 2020).

Method

Setting and participants

This research was carried out from July to August 2023 at SMAN 1 Gamping. Participants came from classes X B and X C and were selected using a simple random sampling method with the condition that the two classes had to be homogeneous first.

Research design

This research was conducted using quantitative methods. This method is closely related to the measurement and analysis of cause-and-effect relationships. Data collection can be in the form of surveys or experiments (Hardani et al., 2020, p. 8). Establish an objective relationship between the researcher and those being studied. Researchers want to have research results that can be generalized at the research location (Sugiyono, 2013, p. 242).

This research used Quasi-Experimental to determine the effect of a treatment on the object of research (Hardani et al., 2020, p. 362). In this research, researchers used a nonequivalent control group design so that the control and experimental classes will have a pretest-posttest (Sugiyono, 2013, p. 357). Class X B as a control class was treated like normal learning activities like giving feedback to friends who came forward to describe something, the feedback in the form of writing on paper. Class X C was an experimental class with treatment in the form of instructing students to upload a video containing their recordings describing something to Instagram reels and asking other friends to comment on the post.

Instrument of the research

The main component of the research was data. To obtain these data, the researcher should create research instruments that follow the type of data needed and the research problem. The instrument was measured based on the guidelines that had been made.

Table 1. Guidelines for pre-test and post-test speaking skills

Material	Learning Objectives	Assessment Indicator	Speaking Skills Indicator	Item
Description text	10.C.3 Presenting orally in the form of descriptive texts related to the topic of the physical and social environment of society by paying attention to social functions, text structure, and linguistic elements according to the context in a polite, critical, creative, and independent manner with an optimal level of fluency and accuracy (speaking skill).	Students present descriptive texts related to the topic of the physical and social environment of the community with an optimal level of pronunciation.	Pronunciation Fluency Accuracy	3 (oral test)

According to Brown (2004, pp. 172–173), state that pronunciation, fluency, and accuracy proficiency scoring categories are as follows:

Table 2. Categories of scoring pronunciation

Score	Indicators
5	Equivalent and recognized by trained native speakers.
4	Mispronunciation is quite rare.
3	Errors never get in the way of understanding and rarely bother native speakers. The accent is very foreign.
2	The accent is understandable, although there are occasional mistakes.
1	Mispronunciations are common but understandable for native speakers who are used to dealing with foreigners trying to speak their language.

(Brown, 2004, p. 173)

Table 3. Categories of scoring fluency

Score	Indicators
5	It has perfect fluency, so native speakers can fully receive it.
4	Able to use the language without long pauses and able to be in any conversation with high fluency.
3	Can discuss an issue with rarely pausing to think about the following sentence.
2	Have high confidence, but in everyday conversations, there are often long pauses in the middle of a conversation.
1	There is no specific fluency description.

(Brown, 2004, p. 173)

Table 4. Table categories of scoring accuracy of grammatical

Score	Indicators
5	Have the ability equivalent to native speakers.
4	Accurate use of language in all professional requirements. Grammar errors are rare.
3	The use of language that is accurate enough to follow conversations on specific topics and is simple. Grammar errors are standard.

Score	Indicators
2	They need to be more confident even though they can follow basic conversations with sufficient grammar.
1	Conversation content can be understood even using lousy grammar.

(Brown, 2004, p. 172)

According to Brown (2004, p. 174), subcategories of oral proficiency scores are as follows:

Table 5. Subcategories of oral proficiency scores

Level	Description
0	Does not work with spoken language.
0+	Able to meet urgent needs through speaking practice.
1	Able to maintain minimum standards of politeness and have effortless personal conversations on familiar topics.
1+	Able to initiate and maintain predictable personal conversations and respond to limited social demands.
2	Able to meet daily social needs and limited work needs.
2+	Often able to meet most job requirements, but only sometimes in acceptable and compelling language.
3	Able to speak the language with structural accuracy and sufficient vocabulary to participate effectively in most formal and informal conversations on practical, social, and professional topics.
3+	Often able to use language to meet broad professional needs and complex and demanding tasks.
4	Able to use the language fluently and accurately at all levels usually associated with job requirements.
4+	Language skills are regularly better in every way and are usually at the level of a trained and fluent native speaker.
5	Speech skills are functionally on par with experienced and highly educated native speakers, reflecting the cultural norms of the country where the language is spoken.

(Brown, 2004, p. 174)

The way of marking is as follows:

Final score= $\frac{GAINED\ SCORE}{MAX\ SCORE} \times WEIGore = \frac{15}{15} \times 100 = 100$

Validity and Reliability of the Instrument

1. Validity

   Content validity relates to the ability of an instrument to measure the content that must be measured. Content validity is also related to the process of logical analysis (Siregar, 2013, p. 46). Research instruments were consulted with expert lecturers or validator lecturers, Mrs. Wiyanah, M.Hum. Construct validity is validity related to the ability of an instrument to measure what must be measured (Siregar, 2013, p. 47). In this research, construct validity was done using SPSS.

Table 6. Instrument validity correlation coefficient criteria

Correlation Coefficient	Correlation	Validity Interpretation
0,90 ≤ r_{count} ≤ 1,00	Very high	Very valid
0,70 ≤ r_{count} < 0,90	High	Valid
0,40 ≤ r_{count} < 0,70	Moderate	Quite valid

Correlation Coefficient	Correlation	Validity Interpretation
0,20 ≤ r_{count} < 0,40	Low	Somewhat valid
r_{count} < 0,20	Very low	Less valid

(Lestari & Zakarsyi, 2017, p. 193)

2. Reliability
   Reliable here means that it can be used to measure many times at different times but has the same result (Sugiyono, 2013, p. 393). Researchers used internal reliability with Cronbach's Alpha formula.

Table 7. Instrument Reliability Correlation Coefficient Criteria

Correlation Coefficient	Correlation	Reliability Interpretation
0,90 ≤ r_{count} ≤ 1,00	Very high	Very reliable
0,70 ≤ r_{count} < 0,90	High	Reliable
0,40 ≤ r_{count} < 0,70	Moderate	Quite reliable
0,20 ≤ r_{count} < 0,40	Low	Somewhat reliable
r_{count} < 0,20	Very low	Less reliable

(Lestari & Zakarsyi, 2017, p. 206)

The technique of data analysis

1. Normality test
   The normality test that was used in this research is Kolmogorov- Semirnof. The SPSS 25 application assists with the calculation.
   The hypothesis used is:
   H_0 = data comes from a normally distributed population
   H_1 = data does not come from a normally distributed population
   If probability (Sig.) ≥ 0,05, then cap H_0 is accepted (Sig.) < 0,05, then H_0 is rejected.

2. Homogeneity Test
   For the homogeneity of the data, the Levene test was carried out. The SPSS application assists in the calculation.
   The hypothesis used is:
   H_0 : $\sigma_1^2 = \sigma_2^2$, both data are homogenous
   H_1 : $\sigma_1^2 \neq \sigma_2^2$, both data are not homogenous
   If probability (Sig.) ≥ 0,05, then H_0 is accepted.
   If probability (Sig.) < 0,05, then H_0 is rejected.

3. T-test
   According to Creswell (2017, p. 220), there are criteria for selecting a statistical test, one of which is if the question criteria are group comparisons, the number of independent and dependent variables is 1. The scores are normally distributed, then someone can use the t-test. These criteria match this research, so researchers will choose the t-test as a statistical data test. In the comparative data of the two sample groups with interval or ratio data types, there are two types of t-tests, namely paired t-test (for correlated data) and independent t-test (for independent data) (Siregar, 2013, p. 177).

a. Paired sample t-test

A paired sample t-test was used to answer the formulation of the problem, "Is peer feedback through Instagram effective in improving students' speaking skills in tenth grade at SMAN 1 Gamping?" Sample t-tests was carried out on pre-test and post-test data in both classes.

According to Djudin (2013, p. 17), the t-test formula for paired samples is:

Hipotesis:

H_0 = Peer feedback through Instagram is not effective in improving students' speaking skills in tenth grade at SMAN 1 Gamping.

H_a = Peer feedback through Instagram is effective in improving students' speaking skills in tenth grade at SMAN 1 Gamping.

Decision criteria:

If Sig. (2-tailed) > 0,05 or $T_{count} > T_{table}$, H_0 was accepted and H_a was rejected. It means that there was no effect of peer feedback through Instagram on tenth-grade students' speaking skills in the descriptive text at SMAN 1 Gamping.

If Sig. (2-tailed) < 0,05 or $T_{count} < T_{table}$, H_0 was rejected and H_a was accepted. It means that there is an effect of peer feedback through Instagram on tenth-grade students' speaking skills in the descriptive text at SMAN 1 Gamping.

b. Independent sample t-test

An Independent t-test was carried out to see if there were differences in student scores after being treated using peer feedback through Instagram and not being treated using this learning media.

According to Djudin (2013, p. 22), the t-test formula for independent samples is:

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Hypothesis:

H_0 = There is no difference in students' scores after treatment using peer feedback through Instagram and not treatment using this learning media on tenth-grade students at SMAN 1 Gamping.

H_a = There is a difference in students' scores after treatment using peer feedback through Instagram and not treatment using this learning media on tenth-grade students at SMAN 1 Gamping.

Decision criteria:

If Sig. (2-tailed) > 0,05 or $T_{count} < T_{table}$, H_0 was accepted and H_a was rejected. It means that students' scores were the same after treatment using peer feedback through Instagram and not treatment using this learning media on tenth-grade students at SMAN 1 Gamping.

If Sig. (2-tailed) < 0,05 or $T_{count} > T_{table}$, H_0 was rejected and H_a was accepted. It means that there was a difference in students' scores after treatment using peer feedback through Instagram and not treatment using this learning media on tenth-grade students at SMAN 1 Gamping.

Findings and Discussion

Description of Research Data

The control class used traditional peer feedback, and the experimental class used peer feedback through Instagram. Descriptive statistics data processing was assisted by SPSS 25; for more concise, it can be seen in the image below:

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Pre test control	36	63	83	73.69	4.827	23.304
Post test control	36	68	88	77.89	5.159	26.616
Pre test experiment	36	65	83	72.97	4.279	18.313
Post test experiment	36	65	83	74.81	5.291	27.990

Figure 1. The output of descriptive statistics test results

Based on the results of the pre-test control class, the lowest score was 63, and the highest score was 83, with an average of 73.69. In the class control post-test, there was a lowest score of 68 and a highest score of 88, with an average of 77.89. Based on the results of the experimental class pre-test, the lowest score was 65, and the highest score was 83, with an average of 72.97. In the experimental class post-test, there was a lowest score of 65 and a highest score of 83, with an average of 74.81.

Validity and Reliability of the Instrument

1. Validity

In the construct validity test, the test results can be seen in the image below

Correlations

		Pronounciation	Fluency	Accuracy	Total
Pronounciation	Pearson Correlation	1	.682**	.677**	.908**
	Sig. (2-tailed)		.000	.000	.000
	N	36	36	36	36
Fluency	Pearson Correlation	.682**	1	.665**	.883**
	Sig. (2-tailed)	.000		.000	.000
	N	36	36	36	36
Accuracy	Pearson Correlation	.677**	.665**	1	.860**
	Sig. (2-tailed)	.000	.000		.000
	N	36	36	36	36
Total	Pearson Correlation	.908**	.883**	.860**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	36	36	36	36

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 2. The output of validity test results for pre-test questions

Based on Table 6, the pronunciation assessment aspect correlates very highly with excellent validity interpretation. The aspects of fluency and accuracy are highly correlated with reasonable validity interpretation.

Correlations

		Pronounciation	Fluency	Accuracy	Total
Pronounciation	Pearson Correlation	1	.712**	.683**	.910**
	Sig. (2-tailed)		.000	.000	.000
	N	36	36	36	36
Fluency	Pearson Correlation	.712**	1	.580**	.889**
	Sig. (2-tailed)	.000		.000	.000
	N	36	36	36	36
Accuracy	Pearson Correlation	.683**	.580**	1	.834**
	Sig. (2-tailed)	.000	.000		.000
	N	36	36	36	36
Total	Pearson Correlation	.910**	.889**	.834**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	36	36	36	36

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 3. The output of validity test results for post-test questions

Based on Table 6, the pronunciation assessment aspect correlates very highly with excellent validity interpretation. The aspects of fluency and accuracy are highly correlated with reasonable validity interpretation.

2. Reliability
The formula used for reliability testing is Alpha Cronbach; show the results below:

Reliability Statistics

Cronbach's Alpha	N of Items
.851	3

Figure 4. The output of reliability test results for pre-test questions

Based on Table 7, the instrument is highly correlated with reasonable reliability interpretation.

Reliability Statistics

Cronbach's Alpha	N of Items
.847	3

Figure 5. The output of reliability test results for post-test questions

Based on Table 7, the instrument is highly correlated with reasonable reliability interpretation.

The technique of data analysis

1. Normality test
   In the normality test, the test results can be seen in the image below:

Tests of Normality

	Class	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
Speaking Test BC	Pre Test Control Class	.141	36	.070	.967	36	.341
	Post Test Control Class	.131	36	.122	.961	36	.226
	Pre Test Experimental Class	.141	36	.066	.955	36	.154
	Post Test Experimental Class	.126	36	.162	.946	36	.076

a. Lilliefors Significance Correction

Figure 6. The output of normality test results

In Image 6, it can be seen that all test data in both classes is declared normal. In the pre-test control class data, it has a Sig value of 0.07 > 0.05, so H_0 is accepted. In the post-test data, the control class has a Sig value of 0.112 > 0.05, so H_0 is accepted. The experimental class pre-test data has a Sig value of 0.066 > 0.05, so H_0 is accepted.

2. Homogeneity Test
   In the homogeneity test, the test results can be seen in the image below:

Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
speaking test	Based on Mean	.012	1	70	.912
	Based on Median	.013	1	70	.909
	Based on Median and with adjusted df	.013	1	69.997	.909
	Based on trimmed mean	.007	1	70	.934

Figure 7. The output of homogeneity test results

In Image 7, it has a Sig value. 0.912 > 0.05, so H_0 is declared accepted. The data significance value is more than 0.05, so it can be concluded that the data comes from a homogeneous population or that the data has the same variance.

3. T-test
   a. The effectiveness of peer feedback through Instagram on tenth-grade students' speaking skills at SMAN 1 Gamping.
      The results of the paired sample T-test can be seen in the image below:

Paired Samples Test

		Paired Differences Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Pretestcntrl - Posttestcntrl	-4.194	4.146	.691	-5.597	-2.792	-6.070	35	.000
Pair 2	Pretestexprmn - Posttestexprmn	-1.833	4.855	.809	-3.476	-.191	-2.266	35	.030

Figure 8. The output of paired sample t-test

The interpretation of Image 8 shows the difference in results between the pre-test and post-test in the experimental class. Please focus on the results in the experimental class; it shows Sig. (2-tailed) 0.030 < 0.05 and T_{count} $(-2{,}266)$ < T_{table} $(-2{,}030)$, so H_0 is rejected. That can be interpreted as the effect of peer feedback through Instagram on students' speaking abilities.

Before peer feedback through Instagram was applied to students, researchers observed students' weaknesses and found out the demands of students' speaking skills from the teaching module first. Researchers categorize assessments into three aspects, namely pronunciation, fluency, and grammatical accuracy. After the media was applied to students, researchers saw changes in students' speaking abilities based on the resulting scores. The interpretation of the results of the paired sample t-test has proved that.

From what has been explained, researchers can conclude that peer feedback through Instagram is effective in improving speaking skills in tenth-grade students at SMAN 1 Gamping. In addition, peer feedback through Instagram is work and can be applied as an alternative medium for learning in speaking classes if there is no other learning medium. These results are in line with the research Ishak and Yaacob (2022), which states that uploading videos on Instagram and peer feedback was focused on speaking skills, resulting in the finding that Instagram is the right tool for practicing speaking skills and a tool for interacting with peers that is fun and flexible in its implementation.

b. The difference in using learning media in improving students' speaking skills among tenth-grade students at SMAN 1 Gamping.
The results of the independent sample T-test can be seen in the image below:

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Result	Equal variances assumed	.012	.912	2.504	70	.015	3.083	1.232	.627	5.540
	Equal variances not assumed			2.504	69.956	.015	3.083	1.232	.627	5.540

Figure 9. The output of an independent sample t-test

In Image 9, the independent sample t-test result shows Sig. (2-tailed) 0.015 < 0.05 and dan T_{count} (2,504) > T_{table} (1,994), so H_0 is rejected, and H_a is accepted. That can be interpreted as indicating that there is a difference in the post-test scores of the control class and the experimental class in class 10 of SMAN 1 Gamping. The Image also shows the average difference. From these results, there are differences in learning outcomes between the post-tests of the two classes. The average difference is 3.083, but from this picture, researchers still need to show which learning medium is more effective.

Group Statistics

	Class	N	Mean	Std. Deviation	Std. Error Mean
Result	Post Test Control Class	36	77.89	5.159	.860
	Post Test Experimental Class	36	74.81	5.291	.882

Figure 10. The output of independent sample t-test statistics

To find out which learning medium is more effective in improving students' speaking skills, the reader has to look at Image 10. In this picture, it is clearly shown which post-test average is higher between the two classes. The results of the statistical group show an average of 77.89 for the class control post-test and 74.81 for the class experimental post-test average. The average in the control class is superior after being given traditional peer feedback treatment in the speaking class.

In this research, researchers found several essential points that might influence the post-test results. Firstly, in the experimental class, students felt embarrassed if their videos were seen by many people other than their classmates, even though they were required to upload assignment videos on Instagram. That causes most of the students in that class to upload videos on Instagram. However, soon, the post is deleted, leaving students with no documentation of the deficiencies they need to improve. Second, in the control class, students were more confident when practicing speaking in front of the class, and students kept papers containing feedback from friends. That has a positive impact because students have documentation to remember the parts that need to be improved.

From the discussion above, researchers can conclude that traditional peer feedback is more effective than peer feedback through Instagram in improving the speaking skills of class 10 students at SMAN 1 Gamping. In addition, traditional is more recommended for use as a learning medium in speaking classes. This result is in line with the results of the study by Tseng and Yeh (2019) which state that students prefer written feedback to video feedback. Students prefer written peer feedback because it is more efficient for finding the information they need. This type of peer feedback helps improve speaking skills, especially in the aspects of grammar rules and word usage, to achieve better linguistic accuracy.

Conclusion

Based on the research conducted, peer feedback through Instagram is effective in improving the speaking skills of tenth-grade students at SMAN 1 Gamping. Moreover, the independent t-test found that traditional peer feedback is more effective than peer feedback through Instagram. In addition, traditional peer feedback is recommended as a learning medium for speaking class.

References

Ainia, D. K. (2020). Merdeka belajar dalam pandangan Ki Hadjar Dewantara dan relevansinya bagi pengembangan pendidikan karakter. *Jurnal Filsafat Indonesia*, *3*(3), 95–101.

Alfarisy, F. (2021). Kebijakan pembelajaran Bahasa Inggris di Indonesia dalam perspektif pembentukan warga dunia dengan kompetensi antarbudaya. *Jurnal Ilmiah Profesi Pendidikan*, *6*(3), 303–313. https://doi.org/10.29303/jipp.v6i3.207

Berk, L. E. (2012). *Infant, children, and adolescents* (7th ed.). Boston, MA: Pearson Education.

Brown, H. D. (2004). *Language assessment: Principles and classroom practices* (Nachdr.). New York, NY: Longman.

Creswell, J. W. (2017). *Research design—Pendekatan metode kualitatif, kuantitatif, dan campuran* (4th ed.). Yogyakarta: Pustaka Pelajar.

Djudin, T. (2013). *Stastistika parametrik—Dasar pemikiran dan penerapannya dalam penelitian* (1st ed.). Yogyakarta: Tiara Wacana.

Handayani, F. (2016). Instagram as a teaching tool? really?. *Proceedings of the Fourth International Seminar on English Language and Teaching (ISELT-4)*, 320-327.

Hardani, H. A., Ustiawaty, J., Utami, E.F., Istiqomah, R.R., Fardani, R.A., Sukmana, D.J., & Auliya, N.H. (2020). *Metode penelitian kualitatif dan kuantitatif* (1st ed.). Yogyakarta: Pustaka Ilmu.

Harmer, J. (1998). *How to teach English* (7th ed.). Harlow: Pearson Education Limited.

Harmer, J. (2007). *The practice of English language teaching* (4th ed.). Harlow: Pearson Education Limited.

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, *39*(2), 83–101. https://doi.org/10.1017/S0261444806003399

Irawan, R., & Wiyanah, S. (2021). Asynchronous peer feedback in EFL writing. *Proceedings UPY International Conference on Applied Science and Education*, *2*(1).

Ishak, S. M. M., & Yaacob, A. (2022). Enhancing ESL students' speaking motivation through Instagram (IG). *Practitioner Research*, *4*, 53–69. https://doi.org/10.32890/pr2022.4.4

Lauder, A. (2008). The status and function of English in Indonesia: A review of key. *Makara Human Behavior Studies in Asia*, *12*(1), 9-20. https://doi.org/10.7454/mssh.v12i1.128

Lei, Z. (2017). Salience of student written feedback by peer-revision in EFL writing class. *English Language Teaching*, *10*(12), 151-157. https://doi.org/10.5539/elt.v10n12p151

Lestari, K. E., & Zakarsyi, M. W. (2017). *Penelitian pendidikan matematika*. Bandung: Refika Aditama.

McDonough, J., Shaw, C., & Masuhara, H. (2013). *Materials and methods in ELT : A teacher's guide* (3rd ed.). London: Blackwell Publishing Ltd.

Pikhart, M., & Botezat, O. (2021). The impact of the use of social media on second language acquisition. *Procedia Computer Science*, *192*, 1621–1628. https://doi.org/10.1016/j.procs.2021.08.166

Siregar, S. (2013). *Metode penelitian kuantitatif* (1st ed.). Jakarta: Kencana Prenada Media Group.

Sugiyono, S. (2013). *Metode penelitian kuantitatif kualitatif dan R&D* (19th ed.). Bandung: Alfabeta.

Tanjung, Z., & Amelia, S. (2017). Menumbuhkan kepercayaan diri siswa. *JRTI (Jurnal Riset Tindakan Indonesia)*, *2*(2), 1-4. https://doi.org/10.29210/3003205000

Tseng, S.-S., & Yeh, H.-C. (2019). The impact of video and written feedback on student preferences of English speaking practice. *Language Learning*, *23*(2), 145-158. https://doi.org/10125/44687

Usmirawati, E. (2021). *Alur dan tujuan pembelajaran Bahasa Inggis (SMA)*.

Wahyuningsih, S., & Afandi, M. (2020). Investigating English speaking problems: Implications for speaking curriculum development in Indonesia. *European Journal of Educational Research*, *9*(3), 967–977. https://doi.org/10.12973/eu-jer.9.3.967

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, *10*, 3087. https://doi.org/10.3389/fpsyg.2019.03087