

CHAPTER III

RESEACH METHODOLOGY

A. Research Type and Design

This study was conducted used a quantitative method. Quantitative research was a means of testing objective theories by examining the relationships among variables. These variables, in turn, were measure typically using instrument so that numerical data could be analyzed through statistical procedures. The final written report followed a structured format consisting of an introduction, literature and theory, methods, results, and discussion (Creswell, 2009).

In this design, two groups were involved: an experimental group and a control group, both of which were selected without random assignment. Only the experimental group received the treatment, which involved the application of the Jigsaw technique and movie clips as instructional media, while the control group was taught without jigsaw technique and movie clips as a media. After the treatment, both groups were given a post-test to assess differences in learning outcomes. This type of quasi-experimental design allowed the researcher to examine the effectiveness of the treatment while controlling for external variables as much as possible (Creswell, 2013).

B. Population and Sample of the Research

1. Population

This research was conducted at SMP Negeri 4 Bukittinggi, located at Jl. Panorama No. 36, Kayu Kubu, Kec. Guguk Panjang, Kota Bukittinggi, Prov Sumatra Barat. The researcher selected this location because, based on observations during the field practice program (PPL), there were issues identified that served as the basis for this study.

Futhurmore, the sample was understood as a subset of the population being studied, which was expected to be representative of the entire population. In this study, the population consisted of all eighth-grade there are VIII A, VIII B, VIII C, VIII D, VIII E, VIII F, VIII G, VIII H, and VIII I all of the total of nine classes.

In this research, the population can be seen through the following table. Junior High School in the academic year 2024/2025.

Table 3 1

The population of the research

No	Class	Student
1	VIII A	30
2	VIII B	30
3	VIII C	31
4	VIII D	32
5	VIII E	30
6	VIII F	31
7	VIII G	31
8	VIII H	30
9	VIII I	30
Total	9 classes	275

2. Sample

The sample in this research was taken using cluster random sampling. This technique was chosen because the population consisted of several intact classes that could not be randomly assigned individually. Cluster random sampling allows the researcher to randomly select groups (or clusters) as a whole rather than individuals, which is practical and appropriate in educational settings where students are already organized into classes.

In this study, the population included all eighth-grade students at SMPN 4 Bukittinggi. From the available classes, the researcher randomly selected two intact classes: VIII E and VIII I. After a lottery draw and consultation with the English teacher regarding class characteristics, Class VIII E was assigned as the experimental group and Class VIII I as the control group. These two classes were chosen because they had relatively similar academic abilities based on their previous English scores, and they had not previously received any treatment or special method related to listening skill. The details of the sample were shown in the table below:

Table 3 2

Sample of the research

No	Class	Number of students	Sample
1	VIII E	30	30
2	VIII I	30	30
Total sample			60

C. Technique of Collecting Data.

The technique of collecting data in this research was carried out through a listening post-test administered to both the experimental and control classes after the treatment. The purpose of this test was to measure students' listening skill after receiving different teaching techniques. The experimental class was taught using the Jigsaw technique integrated with movie clips, while the control class was taught using the Three-Phase Technique and speaker-based audio without the integration of Jigsaw or video media. The instrument used was a listening skill test in the form of multiple-choice questions based on recount texts. The post-test consisted of 20 questions designed to assess students' understanding of main ideas, specific information, and sequence of events presented in the listening materials.

The data collection was conducted in the following steps:

1. Preparation of Intrument

The researcher prepared the post-test questions, answer sheet, audio script, and selected movie clips for treatment in the experimental class. The intrsument was validated by two English teachers and tried out on a different class to test its validity and reliability.

2. Conducting the Treatment

The tratment was given five meetings over period of approximately one month, from May 8 to May 28, 2025, with different techniques for each group.

3. Administering the Post-Test

The post-test was conducted on May 30, 2025 in both classes (VIII E and VIII I) after the final treatment. Students listened to audio (with or without video, depending on the group) and answered the provided multiple-choice questions.

4. Collecting the Students' Answer Sheets

The students' answers were collected, checked, and scored to obtain the final data for analysis.

D. Research Instrument

1. Test

The instrument used in this research was a listening skill test in the form of 20 multiple-choice questions. The test was designed to measure students' ability to understand recount texts after being taught through different techniques in the experimental and control groups. The test items were based on the content of the movie clips (for the experimental group) and audio texts (for the control group), which had been selected and adjusted to the students' level. The questions focused on understanding specific information, sequence of events, and details in the text, which are essential elements of listening skill.

Before being administered to the sample, the test was first tried out to another class on May 3, 2025, to check its validity and reliability. After the test was revised, it was used as the post-test for both groups. During the post-test, students listened to the audio once and answered the

questions individually within a limited time. The test was administered on May 29, 2025, in each class using speaker media. For further information about the instruction of the text, the researcher presents the blueprint of the test according to Shocking Awful (n.d) in iRubric as follows:

Table 3 3
The Blue Print of Listening Skill Test

Variable	Indicators	Number Item
Listening Skill	Ability to focus - Student ability in understanding what is the main idea of the movie clips	1,6,11
	General Understanding - Students' ability in analyzing the meaning of words, action, verbs, and adverbs of time, place in comprehension text	2,5,7,14,16,18,
	Listening for details - Students' understanding the listening for specific information in the movie clip	3,8,10,13,19
	Accuracy of answer - Students' ability in accuracy of information based on movie clips	4,9,12,15,20

E. Technique of Data Collection

In this research, the data were collected through a listening skill test in the form of a multiple-choice post-test. The purpose of the test was to measure students' understanding of recount texts after being taught through different methods in both the experimental and control groups.

Before conducting the actual post-test, a try-out test was held on May 3, 2025, at SMPN 4 Bukittinggi, involving 30 students outside the research

sample to analyze the validity and reliability of the test items. Based on the results, valid and reliable items were selected for the final post-test.

The post-test was then administered to both the experimental and control groups. In the experimental group, students were taught using the Jigsaw technique combined with movie clips, while the control group was taught using the Three-Phase Technique and speaker media. The same test was given to both groups to ensure objectivity in data collection.

The post-test was conducted in the classroom using a speaker to play the audio/movie clips. Students answered the questions individually within a fixed time. The students' answers were then scored and analyzed to determine the effectiveness of the applied teaching methods.

1. Validity of test

Before administering the actual post-test to both the experimental and control groups, the researcher conducted a try-out of the listening skill test to measure its validity. The try-out was held on May 3, 2025, at SMPN 4 Bukittinggi, involving 30 students of Grade VIII who were not included in the research sample. The steps of the try-out process were as follows:

1. The researcher prepared 20 multiple-choice questions based on recount texts, using a format and difficulty level similar to the actual post-test.
2. The try-out was conducted in the classroom by playing the audio using a speaker.

3. The students listened carefully to the recording and answered the questions individually.
4. After collecting the students' answers, the researcher analyzed the data using Pearson Product Moment correlation formula with the help of SPSS 20 software to determine the validity of each test item.
5. Items with valid correlation scores were retained for the final test, while invalid items were revised or removed.

To determine the validity of the instrument, the researcher calculated it using SPSS version 20. The standard value for validity is $r_{\text{item}} > r_{\text{table}}$. Based on the results of the tryout, it was concluded that all items were valid. The results of the tryout are as follows:

Table 3.4

Table the item validity of try out

No. item	r table	Pearson correlation r (item)	Sig. (2-tailed)	Result	Interpretation
1	0,361	0,562	0,001	Valid	Sufficient
2	0,361	0,612	0,000	Valid	High
3	0,361	0,607	0,000	Valid	High
4	0,361	0,593	0,001	Valid	Sufficient
5	0,361	0,514	0,004	Valid	Sufficient
6	0,361	0,625	0,000	Valid	High
7	0,361	0,625	0,000	Valid	High
8	0,361	0,776	0,000	Valid	High
9	0,361	0,530	0,003	Valid	Sufficient
10	0,361	0,583	0,001	Valid	Sufficient
11	0,361	0,519	0,003	Valid	Sufficient
12	0,361	0,562	0,001	Valid	Sufficient
13	0,361	0,571	0,001	Valid	Sufficient

14	0,361	0,690	0,000	Valid	High
15	0,361	0,637	0,000	Valid	High
16	0,361	0,552	0,002	Valid	Sufficient
17	0,361	0,649	0,000	Valid	Sufficient
18	0,361	0,592	0,001	Valid	Sufficient
19	0,361	0,571	0,001	Valid	Sufficient
20	0,361	0,658	0,000	Valid	High

This try-out ensured that the test instrument used in this study was valid and appropriate for assessing students' listening skills.

6. Reliability of test

After conducting the validity test, the researcher also analyzed the reliability of the listening skill test to ensure that the instrument was consistent and dependable. The reliability of the test was calculated using the Cronbach's Alpha formula with the help of SPSS 20 software. The result showed that the reliability coefficient was 0.75. According to George and Mallery (2003), a reliability coefficient of 0.70 or higher is considered acceptable. Therefore, the test used in this research was considered reliable and appropriate for measuring students' listening skill of recount texts.

According to Creswell (2012), the guidelines for reliability were as follows:

Table 3. 5

Category of Reliability

Cronbach Alpha	Intrument Test
0.86-100	Very high
0.66-0.85	High

0.36-0.65	Low
0.20-0.35	Very low
0.00-0.10	Unreliable

In this research, the researcher used SPSS version 20 to calculate the reliability of the test. The result of the reliability test was as follows:

Table 3. 6

The result of reliability test

Reliability Statistics

Cronbach's Alpha	N of Items
.750	20

Based on the result of the reliability test using Cronbach's Alpha, the score obtained was 0.75. According to George and Mallery (2003), a reliability coefficient of 0.70 or higher is considered acceptable. Therefore, the listening comprehension test used in this study is considered reliable.

F. Data Analysis

The data collected from the post-test scores of both the experimental and control groups were analyzed quantitatively. The analysis aimed to determine whether there was a significant difference in listening comprehension achievement between the two groups after receiving different treatments. The analysis procedure included the following steps:

1. Descriptive Statistics

Descriptive statistics a statistical method used to summarize and describe the main characteristics of a dataset in a clear and informative way. It provides a general overview of the data collected from research subjects, enabling researchers to understand the distribution and trends within the dataset.

According to Creswell (2012), descriptive statistics help researcher to describe, organize, and summarize data through numerical values and visual representations, such as tables or graphs. The main components of descriptive analysis in this research included:

a. Mean (Average)

The mean is the arithmetic average of a set of scores and is commonly used to represent the central tendency of the data. It is calculated by dividing the sum of all scores by the number of observations.

b. Standard Deviation (SD)

The standard deviation indicates how spread out the scores are from the mean. A smaller standard deviation means that the scores are clustered closely around the mean, while a larger standard deviation indicates greater variability. It reflects the consistency of the data and how much each score deviates from the average score.

c. Minimum Score

The minimum score is the lowest score obtained in the dataset. It represents the lower boundary of the data and helps identify the weakest performance or outcome in the group.

d. Range

The range is the difference between the highest and the lowest scores in the dataset. It provides a basic measure of variability.

2. Pre-Requisite Data Analysis

a. Normality Data

The simplest form of normality testing involved graphing the frequency distribution of the available scores. However, such testing depended heavily on the researcher's ability to carefully observe the data plot. If the sample size was large and the distribution was not perfectly normal, relying solely on visual observation could lead to inaccurate conclusions. To address this limitation, experts developed more objective statistical methods for assessing normality. Two widely used tests were the Kolmogorov-Smirnov Test and the Lilliefors Test. Another commonly used method, particularly with small sample sizes, was the Shapiro-Wilk Test. In the Shapiro-Wilk Test, the proposed hypotheses were as follows:

- 1) If the significance value (p) > 0.05 , the data is considered to be normally distributed.

- 2) If the significance value (p) < 0.05 , the data is considered not to be normally distributed.

Once the data was confirmed to be normally distributed, a homogeneity of variances test was conducted using the Levene test, also with SPSS version 20 for windows.

b. Homogeneity Test

The homogeneity test was conducted to determine whether the data obtained from the two groups had equal variances. This test was essential to fulfill one of the assumptions for parametric statistical analysis. In this study, researcher was conducted to determine whether students in the class have homogeneous variance or not.

Homogeneity test of two variances on the result of post test experiment class and control class data using Levene's test with Statistical Package For Social Sciences (SPSS) 20 for windows software. The decision criteria in the homogeneity test in SPSS according to Arrifin (2017) are:

- 1) If the significance value (p) < 0.05 means that the data is declared non-homogeneous.
- 2) If the significance value (p) > 0.05 means that the data is declared homogeneous.

From the test result, the data of both group have the same variance, so the hypothesis test is carried out with the Levene test using the one sample test.

3. Hypothesis testing

Hypothesis testing is a statistical procedure used to determine whether the data collected in a study provide sufficient evidence to support or reject a proposed hypothesis. It allows researcher to make inferences about population parameters based on sample data.

According to Creswell (2012), hypothesis testing involves formulating two competing hypotheses the null hypothesis (H_0) and the alternative hypothesis (H_1) and using statistical analysis to decide which one is supported by the data. In this study, the hypotheses were as follows:

- a. Null Hypothesis (H_0): There is no significant difference in students' listening skills between those who were taught using the jigsaw technique with movie clips and those who were not.
- b. Alternative Hypothesis (H_1): There is a significant difference in students' listening skills between those who were taught using the jigsaw technique with movie clips and those who were not.

To test the hypothesis, an independent samples t-test was used, as the research involved comparing the post-test results of two independent groups (control and experimental). The SPSS version 20 software was used to perform the statistical calculation.

- a. If the significance value (Sig. 2-tailed) < 0.05 , then H_0 is rejected, and H_1 is accepted.

- b. If the significance value(Sig. 2-tailed) > 0.05 , then H_0 is accepted, and H_1 is rejected.

In the results, the significance value obtained from the independent samples t-test was 0.000, which is less than 0.05. Therefore, the null hypothesis was rejected and the alternative hypothesis was accepted. It can be concluded that the use of the jigsaw technique with movie clips had a significant effect on improving students' listening skills.

