

CHAPTER III

RESEARCH METHODOLOGY

A. Research Design

In conducting research, it need research design. Research design refers to the strategy to integrate the different components of research projects in cohesive and coherence way. Research design is plans and procedures for research to detailed methods of data collection and analysis (Creswell, 2009). The function of a research design is to ensure that the evidence obtain enables us to answer the initial question as unambiguously as possible.

This study is conducted based on a quantitative method. Quantitative research is a means for testing objective theories by examining the relationship among variables. These variables, in turn, can be measured, typically on instruments, so that numbered data can be analyzed using statistical procedures. The final written report has a set structure consisting of introduction, literature and theory, methods, results, and discussion (Creswell, 2009)

The type of this research is a Quasi-experimental design. This study includes two randomly selected groups as samples. The first group, known as the experimental group, receives treatment through an Animation video. The second group serves as the control and does not receive treatment with an Animation video. Prior to treatment, both groups undergo a pretest. Following the treatment, a post-test is administered by the researcher to compare the outcomes between the two groups.

Table III.1 shows the research design as follows:

Tabel III.1
Research Design

Group	Pre- test	Treatment	Post- test
Experimental class	T1	X	T2
Control class	T1	Θ	T2

Note:

T1 : Pre-test

T2 : Post-test

X : Receiving treatment by using Animation video

Θ : No treatment

B. Population and Sample of the Research

1. Population of the research

A Population is a group of individuals who have the same characteristics (Creswell, 2012). It means the population is the whole subject of the study. Understanding of the sample is a part of the subject in the population studied, which is certainly capable of representative can represent the population. This research took place at Islamic Junior High School Kota Padang Panjang, located in Panyalaian, Kec. Sepuluh Koto, Padang Panjang. The researcher selected this location because it was suitable for the researcher's pre-service teacher training.

In this research, the population consists of all eighth-grade students of Islamic Junior High School in the academic year 2024/2025.

Table III.2
The population of the research

NO	CLASS	STUDENT
1	VIII A	32
2	VIII B	32
3	VIII C	32
4	VIII D	32
5	VIII E	24
6	VIII F	24
7	VIII G	30
8	VIII H	30
9	VIII I	30
10	VIII J	30
11	VIII K	29
Total	11 classes	325

2. Sample of the research

According to Keppel as cited in Creswell (2009) Each individual has an equal probability of being selected from the population, ensuring that the sample will be representative of the population. In many experiments, however, only a convenience sample is possible because the investigator must use naturally formed groups (e.g., a classroom, an organization, a family unit) or volunteers.

The sampling technique used in this research is cluster random sampling. Cluster Random Sampling is used when the population is large, requiring it to be divided into several classes or groups. Thus, in this sampling method, the unit of analysis is not the individual but rather the group or class, which consists of several individuals (Machali, I., 2021). To obtain the sample for this study, the researcher used a lottery method to obtain the sample classes. The researcher wrote letters A through K corresponding to the class numbers on several small sheets of paper, put all the lotteries in a box, shook it, and took out two tickets. The lot taken was selected as the sample for this study.

The researcher takes 64 students as a sample using the cluster random sampling technique, where the researcher selects a sample randomly from two classes; VIII A, consisting of 32 students as the experimental group, and VIII B, consisting of 32 students as the control group.

The details of the sample are shown in the table below:

Table III.3
Sample of the research

No	Class	Number of students	Sampel
1	VIII A	32	32
2	VIII B	32	32
Total sample			64

C. Technique of collecting data

In this research, the researcher used a listening test as an instrument in collecting the data. To find out the data both pre-test and post-test, the researcher used listening test.

1. Test

a. Pre Test

The pre-test is the test given before some treatments to assess students' listening comprehension. In this pre-test, students answer multiple-choice questions consisting of 20 questions. They are asked to choose the correct answer from choices a, b, c, and d.

Table III.4
The Blue Print Of Listening Test

Variable	Indicators	Numbers Item
		Pre and Post – test
Listening Skill	Students' comprehension in understanding what is the main idea of the clip	2,10,15,13,19
	Students' comprehension in analyzing the meaning of words, action verbs, and adverbs of time, place in comprehension test	6,8,12,16,18
	Students' understanding the listening for specific information	1,3,4,7,11,17
	Students' comprehension in understanding inferences	5,9,14,20

b. Treatment

The treatment is given to the samples. This research was conducted three times treatment. In each meeting, the students are taught by animation video with different topics

1) First meeting, the researcher explained the material with the flag ceremony topic

2) Second meeting, the researcher explained the material with the library topic

3) Third meeting, the researcher explained the material with mandatory traffic signs topic

c. Post Test

The post-test is given after teaching several times. The post-test is used to collect the final data about students' listening comprehension after giving treatment to the experimental class.

After assessing students' writing results using the scoring rubric, the researcher provided a score classification of students' listening abilities.

The classification of the student's score is shown in the table as follows:

TABLE III.5
The classification of students' score

Score	Categories
Excellent	80-100
Very good	60-79
Good	40-59
Enough	20-39

Bad	Less than 20
-----	--------------

(Haris,1969)

2. Validity

In quantitative research, to obtain valid, reliable and objective data, researchers use valid and reliable instruments, select representative samples from the entire population, and ensure that data collection and analysis is carried out correctly. According to Cresswell (2012), validity involves gathering strong evidence to show that the interpretation of test results (scores that reflect the construct measured by the test) is appropriate for the intended purpose. In other words, validity refers to the extent to which the conclusions drawn from the data are appropriate, meaningful and relevant to the assessment objectives. In this study, to assess the validity of the test, the researcher applied construct validity.

According to Irwan Sulistyanto in Surapranata (2005), construct validity refers to the extent to which the instrument used in the research aligns with the theoretical framework on which the instrument is based. Construct validity is briefly explained through the standard competence, basic competence, and indicators, which are used to evaluate the test's construction. The researcher aims to measure students' listening comprehension.

To determine the validity of the instrument, the researcher calculated it using SPSS version 26. The standard value for validity is $r_{item} > r_{table}$. Based on the results of the tryout, it was concluded that all items were valid.

The results of the tryout are as follows:

Table. III.6

Table the item validity of try out

No Item	Rtable	Ritem	Result	No item	Rtable	Ritem	Result
1	0.246	0.373	Valid	11	0.246	0.523	Valid
2	0.246	0.477	Valid	12	0.246	0.482	Valid
3	0.246	0.516	Valid	13	0.246	0.540	Valid
4	0.246	0.377	Valid	14	0.246	0.670	Valid
5	0.246	0.369	Valid	15	0.246	0.418	Valid
6	0.246	0.551	Valid	16	0.246	0.524	Valid
7	0.246	0.516	Valid	17	0.246	0.569	Valid
8	0.246	0.477	Valid	18	0.246	0.463	Valid
9	0.246	0.426	Valid	19	0.246	0.676	Valid
10	0.246	0.399	Valid	20	0.246	0.506	Valid

Based on the table above, it was clear that there were 20 valid items of 20 items.

3. Reliability of test

According to Irwan Sulistyanto in Surapranata (2005) states that reliability means normal score which is gained by the same subject when did the repeatedly test in the difference situation. The function of test reliability is to determine how much variability which happens because of the fault of measurement and how many the real test score variability. Reliability has two constant, internal and external. Internal means that the items of questions are homogeny from the

difficulty until the form of the test while the external means level of the items to produce the score which constant every time.

According to Creswell (2012), the guidelines for reliability is as follow:

Table III.7

Category of reliability

Cronbach Alpha	Instrument Test
0,86-100	Very high
0,66-0,85	High
0,36-0,65	Low
0,20-0,35	Very low
0,00-0,19	Unreliable

In this research, the researcher used SPSS 26 version to calculate the reliability of the text. The result of test reliability is as follows:

Table III.8

Reliability Statistics

Cronbach's Alpha	N of Items
.834	20

It is said to be reliable if it has a Cronbach's alpha value of more than or equal to 0.6. The condition for a measuring instrument to show higher reliability is if the reliability coefficient (α) is close to one. If the alpha coefficient (α) is greater than 0.6 then the measuring instrument is considered reliable or has internal consistency reliability and conversely if alpha is less than 0.2 then it is

considered less reliable or does not have internal consistency reliability. The of test was 0.834. It is categorized into highly reliable level

D. Technique for data Analysis

In order to find out whether there is or not a significant effect of using animation video to improve students' listening comprehension, data of this research were analyzed statistically. To analyze the students' score in the pre-test and post-test, the researcher used the following formula:

$$S = \frac{B}{N} \times 100$$

Where:

S : students' score N : maximum correct item
B : total correct item

The data was analyzed by using descriptive analysis, used pretest and post-test. Specifically, to analyze the data, the researcher compared the result of the pre-test and post-test using Mann Whitney -test by using SPSS 26.0 Version. In this research, the researcher used these formula as follow:

1. Normality Test

Knowing the data was obtained from the students' score had normal distribution or not. Before choosing a statistical test technique for data with a numerical scale (ratio and interval), it is important to first know whether the data is normally distributed or not. This is crucial because parametric statistical tests, which are often used for ratio and interval scale data, require the data to be

normally distributed. If the data is not normally distributed, then the alternative used is a non-parametric statistical test. The researcher used SPSS 26.0 version.

2. Homogeneity test

The homogeneity test is a test to find out whether the variances of two or more data distributions are the same. This test is carried out to ensure whether the data on variables X and Y are homogeneous or not. Usually, a homogeneity test is carried out as one of the requirements before carrying out further data analysis or statistical analysis. (Setyawan,2021)

After doing normality test, the homogeneity test was conducted to find out, The homogeneity test is used to determine the similarities between two conditions or populations. The homogeneity test used in this research is the Levene test. The Levene test is a method of testing homogeneity of variance which is almost the same as the Bartlett test, but the data tested with the Levene test does not have to be normally distributed. The test criteria is that if the sig value is greater than $\alpha = 0.05$, it can be concluded that the data is homogeneous.

Table III.9
Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
hasil	Based on Mean	2.514	3	124	.062
	Based on Median	2.490	3	124	.063
	Based on Median and with adjusted df	2.490	3	120.076	.064
	Based on trimmed mean	2.494	3	124	.063

Based on table III.9, it can be seen that the results based on the mean significance value (sig) are $0.062 > 0.05$, so it can be concluded that the pretest and posttest data for the experimental class and control class are homogeneous.

3. Hypothesis Testing

After carrying out the normality test and homogeneity test, data analysis was carried out to test the hypothesis that had been proposed. This non-parametric statistical test is used when the data does not meet statistical assumptions, such as when the data is normally distributed but has unequal variances, not normally distributed but homogeneous, or not normally distributed and not homogeneous. If parametric assumptions are not met, the non-parametric statistical tests that will be used are the Wilcoxon signed-rank test and Mann-Whitney test.

a) Wilcoxon signed-rank test

The Wilcoxon Signed Ranks Test is a non-parametric statistical test that is used to compare two paired samples and see whether there are differences between them. This test is used to analyze whether there are differences between two paired data sets. The Wilcoxon test is an option to replace the T-Test if the data is not normally distributed. Proof of this method was carried out by statistical analysis using the SPSS version 26 program.

b) Mann-Whitney test

The Mann-Whitney test is part of non-parametric statistics when the data being analyzed is not normally distributed. This test was carried out to determine

whether there was a significant difference between students' listening comprehension using video animation (experiment class) and without using video animation (control class).

The Mann-Whitney test was carried out using SPSS version 26 software with decision-making:

1). If the Sigificance or Asymp.sig (2-tailed) value is smaller than the probability of 0.05 then H1 is accepted and H0 is rejected then the hypothesis is accepted.

2). If the Sigificance or Asymp.sig (2-tailed) value is greater than the probability of 0.05 then H0 is accepted and H1 is rejected then the hypothesis is rejected.

