

**CHAPTER III**  
**RESEARCH METHOD**

This chapter covers research design, Data and data resources, Variable of the study, population, sample and sampling, Methods of collecting data and instrument, Method of data analysis, validity and reliability testing, Normality, homogeneity testing and T-test, Hypothesis testing.

**A. Research Design**

In this study, the researcher decides a quasi-experimental design to explore the cause and effect relationship between dependent and independent variables. This research involved two groups of subject from seventh grade of MTS Subulussalam Sriwangi with different between experimental and control class. The researcher apply the treatment using Dulingo application for experimental class. Another class called control class is taught without using Duolingo.

The design of the study was captured from Ary et al(2010:303) and presented in the diagram below:

**Table 3.1 The diagram of the Quasi-Experimental design:**

<b>Control Class</b>	<b>Pre-test</b>	<b>No Treatment</b>	<b>Post-test</b>
<b>Experimental Class</b>	<b>Pre-test</b>	<b>Experimental Treatment</b>	<i>Post-test</i>

Based on the table before, the process of quasi- experimental research used two classes pretest-posttest design were:

1. Before treatment, the researcher distributes a preliminary test to measure the level of vocabulary proficiency of both students in the control and experimental.
2. Applying the experimental treatment to the experimental class using the Duolingo application strategy, and applying the traditional educational strategy to the control class as usual.
3. Distributing a post-test to measure vocabulary mastery between the control class and the experimental class after receiving and not receiving treatment.

### **B. Data and Data source**

The data of this study will be got from the students' scores between control and experimental class. Those data will be used to know the student' vocabulary mastery between the students who are taught by Duolingo application Strategy and those who are not taught by Duolingo application Strategy. The data source in this study is the students of seventh grade of MTS Subulussalam Sriwangi.

### **C. Variable of the study**

Independent variables is the variable that the experimenter expects to influence the other (Nunan;1992). One independent variable takes place as the treatment variable. One or more group acts as the experimental manipulation or treatment. On the other hand , Duolingo application Strategy becomes the dependent variable. Dependent variable is reaction or reference variable that is supposed to be affected by independent treatment condition and other independent variables. The dependent variable in this study was the students' vocabulary mastery.

## D. Population Sampling and Sample

### 1. Population

Population is the number of large elements that become the research areas. It consists of a whole group of objects, notes, or points that have a common denominator (Sugiyono 2007). The research population will include the whole seventh grade of the second year students of MTS Subulussalam Sriwangi in academic year 2020/2021. The population taken is 100 students that divides into 4 classes show by the table below:

**Table 3.2 The Total Number of Population**

No	Class	Students
1	VII –A	26
2	VII-B	25
3	VII-C	26
4	VII-D	23

### 2. Sampling

After deciding population which consists of large elements, the researcher need to gain the sample. Sampling is the process of technique in choosing an appropriate sample. As (Sugiyono;2007) said “Sampling is a sample technique”. Purposive sampling used as the sampling in this research. The researcher choose the sample from VII A and VII B classes because those classes had low achievement in vocabulary than others.

### 3. Sample

According to Sugiyono states that sample is a Sample is a part of the population that has the quantity and characteristics. Sample is a piece of the

amount population. Sample becomes one of the important requirement in achieving the successful research. It can be defined that sample is one portion of population which has quality and characteristics. Based on the understanding, there are VII A and VII B classes as the sample of the study, VII-A class consists 26 students of as the experimental class and VII-B class consists of 25 students as the control class. As result, the total sample is 51 students.

#### **E. Methods of Collecting Data and Instrument**

The way of researcher collects data is called data collection method. There are some steps which are provided in processing the data collection. Two types of tests will be used. They are:

##### 1) Pre-test

Pre-test or preliminary test function as the measurement of students' vocabulary mastery that will be performed prior to experimental treatment. It will be administered for both VII-A class as the experimental class and VII-B class as control class before giving the treatment.

##### 2) Post-test

Post-test is a measurement given after treatment. After both experimental and control class are getting the treatment, the post-test is administered. Before conducting post-test The researcher will teach as treatment section during four meetings in both experimental and control classes. The researcher teaches using Duolingo application Strategy for experimental class and using conventional method in control class.

To collect the data in this study, the researcher used a test as the instrument. The instrument is a tool model used in the study. The test on this study focuses on Vocabulary testing. The kind of tests are in the form of multiple choice. Another test is administered through Duolingo Application which there is an automatic correction when students answer question correctly during treatment. There are 20 question with 60 minutes for the time allocation and this test is provided for Post test. The instrument for pre test is a form of multiple choice with 20 questions.

**Table 3.3 The Timeline in Collecting Data**

NO	ACTIVITY	TIMELINE	
		Experimental Class	Control Class
1.	Pre – Test	8 <sup>th</sup> of June 2021	8 <sup>th</sup> of June 2021
2.	Treatment 1	9 <sup>th</sup> of June 2021	14 <sup>th</sup> of June 2021
3.	Treatment 2	10 <sup>th</sup> of June 2021	14 <sup>th</sup> of June 2021
4.	Treatment 3	11 <sup>th</sup> of June 2021	15 <sup>th</sup> of June 2021
5.	Treatment 4	12 <sup>th</sup> of June 2021	16 <sup>th</sup> of June 2021
6.	Post – Test	13 <sup>th</sup> of June 2021	17 <sup>th</sup> of June 2021

#### **F. Method of Data Analysis**

To understand the results of the study needs a method which is called as data analysis process. In this part, the researcher uses the statistical program SPSS 26 version for windows to analyze the quantitative data. A quantitative analysis of the data is used to identify significant differences in the command of vocabulary mastery both students who are taught using Duolingo application strategy and those who learn without the Duolingo application strategy. The post-test data for both the experimental class group and the control class will be statistically analyzed using an independent sample T-test via SPSS 26 for

windows. The researcher will use a t-tester to find out that the value is greater than or less than 0.05.

## **G. Validity and Reliability**

### **1. Validity**

Validity is the most difficult criterion for the effectiveness of the main principle of language testing (Brown;2007). According to Gronlund (1998) as quoted in Brown (2003:22) said that Validity is the degree to infer drawn from an appropriate, meaningful and useful assessment results. Therefore, the measurement that the researcher wants to measure is a valid test. In this study, the researchers used the content, structure, and face validity of the test as research tools.

#### **a. Content Validity**

Content validity is the test that it measures knowledge about the content area whose knowledge it is designed to measure. Hughes (2003:26) stated that the representative sample of language skills, structures, others is the criterion of valid content. Brown (2003: 22) explained that the valid test requires learners to achieve a measurable behavior. It means that the good content of the test is matched with the lesson given or based on the standard and basic competence which is appointed on the 2013 curriculum. To see the basic competence and the test specification of this instrument can be seen on appendix.

The great validity content influence to the trusted data, in gaining the good instrument needs to construct the best test specification. The researcher took

try out the test items to 26 students of VII C students of MTS Subulussalam Sriwangi who are not included as sample group. To make sure that the instrument is valid, The validity of instrument and index difficulty analysis is presented clearly as follows.

- Index Difficulty

According to Sudjiono(2019) state that the criteria of index difficulty, there are three categories such difficult, middle and easy.

It shows in the table 3.4:

**Table 3.4 Index Difficulty Of Test Category**

No	Result of P	Interpretation
1.	<0,30	Difficult
2.	0,30-0,70	Middle
3.	>0,70	Easy

As we see on the table 3.4, It means that for the middle category, these items can be used to be as the next instrument test. For the difficult item and easy category means that these items can not be used as an instrument or need to be revised. To analyze the difficulty of the index, the researcher used the formula:

$$P = \frac{N \text{ correct}}{N \text{ total}}$$

Where:

P : Index of difficulty

N correct : The total number of students who right answer

N total : The total number of students

**Table 3.5 The result of Index Difficulty of test**

No	Question Items	Number of Students	Correct Answer	Difficulty Index	Criteria
1	Item 1	26	21	0,8	Easy
2	Item 2	26	19	0,73	Easy
3	Item 3	26	17	0,65	Middle
4	Item 4	26	15	0,57	Middle
5	Item 5	26	19	0,73	Easy
6	Item 6	26	11	0,42	Middle
7	Item 7	26	13	0,50	Middle
8	Item 8	26	13	0,50	Middle
9	Item 9	26	15	0,57	Middle
10	Item 10	26	12	0,46	Middle
11	Item 11	26	16	0,61	Middle
12	Item 12	26	14	0,53	Middle
13	Item 13	26	14	0,53	Middle
14	Item 14	26	15	0,57	Middle
15	Item 15	26	18	0,69	Middle
16	Item 16	26	12	0,46	Middle
17	Item 17	26	11	0,42	Middle
18	Item 18	26	15	0,57	Middle
19	Item 19	26	13	0,50	Middle
20	Item 20	26	21	0,80	Easy

Based on the table above, it can be seen for the instrument of try out test. The researcher found some some level of middle questions which there were 17 middle questions and 3 easy questions and there was no difficult item. It means that the 3 easy items should be revised or removed as good test items for doing post-test.



- Validity Of Instrument Analysis

The researcher will use statistical product and service solution (SPSS) 26 to calculate the validity of the test items. The researcher determines the significance level ( $\alpha$ ) of the test is 0.05 or 5% and the value  $r$  table of this test is 0.388.

- If the criteria value ( $r_{\text{obtained}}$ ) is positive and more than  $r$  table, it means that item is valid.
- If the criteria value ( $r_{\text{obtained}}$ ) is negative and less than  $r$  table, it means that item is invalid.

The result of validity of the instrument is shown on the table 3.3 below:

**Table 3.3 Validity Of Instrument**

No	Question Items	$r$ obtained	$r$ table	Conclusion
1	Item 1	0.150	0,388	Invalid
2	Item 2	-0.052	0,388	Invalid
3	Item 3	0.664	0,388	Valid
4	Item 4	0.560	0,388	Valid
5	Item 5	0.435	0,388	Valid
6	Item 6	0.468	0,388	Valid
7	Item 7	0.389	0,388	Valid
8	Item 8	0.432	0,388	Valid
9	Item 9	0.538	0,388	Valid
10	Item 10	0.383	0,388	Valid
11	Item 11	0.220	0,388	Invalid
12	Item 12	0.072	0,388	Invalid
13	Item 13	0.440	0,388	Valid
14	Item 14	0.429	0,388	Valid
15	Item 15	0.434	0,388	Valid
16	Item 16	0.015	0,388	Invalid
17	Item 17	0.446	0,388	Valid

18	Item 18	0.516	0,388	Valid
19	Item 19	0.411	0,388	Valid
20	Item 20	0.451	0,388	Valid

Based on the table above, it indicated that there were 5 questions were categorized invalid which was for item number 1, 2, 11, 12, and 16. There are some reasons why the item number is invalid. According to Tinambuan, (1988), Unclear direction, inappropriate level of difficulty of the test, ambiguity, improper arrangement of the test item, pattern of answers which is identified easily, some factors in administering and scoring the test may influence the invalid item of the test. One of the benefit of doing try out makes our instrument stronger which results the trusted data. To make all items are valid, the researcher need to look at the instrument and revise them.

#### b. Construct validity

From the point of view of construct validity, a test constructively valid if it shows that it measures exactly the ability that is assumed in the theory of language ability. As Brown (2003:25) explained, "The construction of any theory, hypothesis, or model that attempts to explain observable phenomena in perceptual universe." The construct validity in this study is measured by consulted to the expert.

#### c. Face Validity

To have the valid instrument, a test has to appropriate to what is meant to be measured. According to Ary et sl (2010); "Face validity is the degree

to which subjects believe that the instrument is measuring what it is supposed to measure". The tests in this study were designed to measure students' vocabulary mastery. The researcher provides test guidance to train students' vocabulary mastery on the multiple-choice test. In order to check the validity of the face on my device, an expert has properly verified it.

## 2. Reliability

A test that is considered reliable if the result is consistent and reliable. Mousavi (2002), quoted by Brown (2003: 20) stated , "When a student takes the same test, or if there are students who match different opportunities, the test may give same results.". The term "same result" here requires the score needs to be consistent. The researcher administers try out instrument for students twice in different time and it usually mentions as Try out with Test-Retest method aim to get two consistence scores. To get the correlation between two sets of scores, the researcher calculates using pearson product-moment formula. The try out test consists of 20 questions in multiple choice form for 60 minutes time allocation. After doing this try out, the Researcher will get two consistence scores. The try out score result can be seen below:

**Table 3.6 Try Out Score**

No	Name	Try Out 1	Try Out 2
1	LU	25	40
2	RDS	50	60
3	IPR	65	65
4	FM	60	55
5	MW	75	80

6	MSA	55	60
7	NR	70	65
8	MM	50	60
9	NO	85	90
10	KNH	70	90
11	IA	40	55
12	JA	35	55
13	JNL	40	40
14	AM	65	80
15	RK	40	45
16	RI	95	90
17	FZS	75	80
18	JNV	65	70
19	REP	40	45
20	PS	50	65
21	AZR	60	35
22	CAZ	40	45
23	EN	60	35
24	MM	55	40
25	EFW	35	40
26	ARD	50	35

Table 3.3 describes two sets of try out scores which obtained from test-retest method. The researcher conducted the try out twice in different time. From this try out, the researcher got two different scores of the students are same in the difficulties level test. The result of reliability test is shown below:

**Table 3.7 Pearson Product Moment from IBM SPSS Statistical 26**

Correlations			
		tryou1	tryout2
tryou1	Pearson Correlation	1	.759**
	Sig. (2-tailed)		.000
	N	26	26
tryout2	Pearson Correlation	.759**	1
	Sig. (2-tailed)	.000	
	N	26	26

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

According to the Table 3.7, the result of reliability test is counted by Pearson Product Moment formula through IBM SPSS Statistical 26 describes 0,759 which is higher than 0,50 or it can be said the result is near to 1. It concludes that the instrument in this research is reliable. The instrument can be unreliable if the result of Pearson product moment result is lower than 0,50.

## **H. Normality, Homogeneity Testing and T-test**

### 1. Normality Testing

The function of normality testing is to know whether the data collection from experimental and control classes is normally distributed or not. In this research, The normality testing is done by using IBM SPSS Statistic Version 26 which has requirements as follows: if the result of normality test is more than  $> 0.05$ , it can be categorized that the distribution of the data is normal, but if result scores is less than  $< 0.05$ , the distribution of the data is not normal.

### 2. Homogeneity Testing

The next step, after having the result from normality test is calculating the homogeneity of data. This test is used to know whether the data in both of classes is homogeneous or not. The researcher calculates a homogeneity test using SPSS 26 with a significance value ( $\alpha$ ) = 0.050. The data is called Homogeneous if the significance value  $> 0.050$  while the Data is not homogeneous if the significance value  $< 0.050$ .

### 3. t-Test

The data analysis process in this study uses T-test in order to know the significant difference between students' vocabulary knowledge using Duolingo

application in experimental class and students' vocabulary knowledge without Duolingo application in control class. The t-test used in this study is Independent Samples T-test with two-tailed test of significance by using IBM SPSS Statistic Version 26. If the result shows Sig. (2-tailed)  $>$  sig  $\alpha = 0.05$  (5%), then the null hypothesis is accepted. But, if Sig. (2-tailed)  $<$  sig  $\alpha = 0.05$  (5%), then alternative hypothesis is accepted.

### **I. Hypothesis Testing**

The hypotheses are formulated with the assumption and will be done on the result of statistical computation, T-test, as follows:

1. Alternative Hypothesis (Ha): There is an effect of using Duolingo application to students' vocabulary mastery or if  $<$  sig  $\alpha = 0.05$  (5%). It means that the alternative hypothesis (Ha) is accepted and the null hypothesis (Ho) is rejected.
2. Null Hypothesis (Ho): There is no effect of using Duolingo application to students' vocabulary mastery or if  $>$  sig  $\alpha = 0.05$  (5%). It means that the alternative hypothesis (Ha) is rejected and the null hypothesis (Ho) is accepted.