

## **CHAPTER III**

### **RESEARCH METHOD**

This chapter presents the research method. It focuses on research design, variables, research setting, research instruments, try out instrumen (validity and reliability), normality and homogeneity, data and data source, data collecting method, and data analysis.

#### **A. Research Design**

In this research, the researcher used quasi-experimental research. This study classified as quasi-experimental reasearch design because the researcher could not randomly put the subject, and classes that used by the researcher was already formed. Ary *et al* (2010:26) stated, that if the researcher could not randomly assign subjects which it had already assembled groups such as classes, it was called as quasi-experimental design.

The researcher used two different classes, namely experimental group and control group. Which is where the experimental group would be given pre-test, treatment, and posttest while the others are the control group that would be given pre-test and post-test. The researcher taught those classes for several meetings. In the experimental class, the researcher employed QAR strategy. On the other hand, in the control class, the researcher employed conventional methods to teach the class in the same way as acceptable. For more detail about the design of quasi-experimental, see the table below:

**Table. 3.1 The Design Pattern of the Study**

<b>Groups</b>	<b><i>Pretest</i></b>	<b><i>Treatment</i></b>	<b><i>Posttest</i></b>
E	Y1	X	Y2
C	Y3	-	Y4

(Ary, et.al 2010:316)

Note :

E : Experimental group

C : Control group

Y1 : Pretest for experimental group

Y2 : Posttest for experimental group

X : Treatment

Y3 : Pretest for control group

Y4 : Posttest for control group

Based on two groups on the table above the first was the experimental group that would receive a treatment (X) while the second group was the control group that did not receive treatment (-). Both of them would receive pre-test to obtain the first data about students' reading comprehension score before the treatment. The Experimental group (E) was given treatment of being taught by using QAR (X) while the control group (C) was being taught without QAR (-). For the conclusion is both of them would be given post-test to obtain the second data about students reading comprehension score. By using T-test, the scores were compared to find out if there was significant difference of reading comprehension ability before and after being taught by using QAR strategy.

## B. Population and Samples

### 1. Population

Population is the whole subject of the research which has certain quality and characteristics. Population is a set to which a researcher wishes to generalize. According to Ary et al (2010:148) population defines as all members of any well-defined class of people, events, or objects. While Arikunto (2013:173) states “Population is the whole subject of research”. Based on Lodico et al (2006:13) the population is the large group to which the researcher would like the result of a study to be generalizable. It means that the population is least one characteristic of differentiates it from other groups. Supported by Creswell (2008:151) population is a group of individuals who have the same characteristic. As a description above, the researcher take conclusion that the population is a whole research subject used by the researcher. So, population is very important part in a research.

Population in this study is all of the seventh graders of MTs. Ma’arif Bakung in the academic year 2018/2019.

**Table 3.2 List of population**

<b>Class</b>	<b>Total</b>
VII A	40
VII B	40
VII C	40
VII D	45
VII E	47
VII F	35
VII G	49
VII H	44
VII I	40
Total	380

## 2. Samples

According to Ary et al (2010:148) sample is a portion of a population. It means that the sample is a set of data consisting of only a part of the research. In other words, good sample must be represented of the entire as possible, so that the generalization of the sample as true as population. Supported by Creswell (2008:152) sample is a subgroup of the target population that the researcher plants to research for generalizing about the target population.

Sampling is the way that used by the researcher to select the number of individuals as a sample in study. In this research the researcher used purposive sampling technique. In this study, samples were taken from two classes from seventh grader of MTs. Ma'arif Bakung. They were VII D class as a control group taught without using Question Answer Relationship (QAR) strategy and VII E class as the experimental group taught by using Question Answer Relationship (QAR) strategy. The researcher decided to choose the class because the English teacher suggested using VII D and VII E class to conduct the research and the two classes have the same criteria and are possible to be used as research. The teacher suggested to using the class because the students in these class include into active students and good to give treatment for them.

### C. Variables

A variable is everything that will become that object of research or the influencing. According to Ary *et al* (2010) is a construct variable or

characteristic that can take on different values or scores. In other words, variable is anything that can change the result of a research. There are two kinds of variable, independent variable and dependent variable.

#### 1) Independent Variable (X)

Independent variable is a variable in which is observed the side effect. Independent variable can appear and exist by itself without any other supported. It influences and gives special effect. It was usually symbolized by "X". In this study, the used of Question Answer Relationship (QAR) was the independent variable.

#### 2) Dependent Variable (Y)

A dependent variable is an attribute or characteristic that is dependent on or influenced by the independent variable. This is the effect of independent variable. This variable was not manipulated by the researcher, but it was affected by the independent variable. It was symbolized by "Y". Dependent variable in this study was students' scores in reading comprehension.

### **D. Research Instrument**

In this study the researcher used a test as an instrument. Where a test is a set of stimulate presented to an individual in order to elicit response on the basis of which a numerical score can be assigned (Ary *et al* 2010:201). It means that the test conducted by the researcher would provide the data of students' score in reading comprehension in the form of numerical score. The test was given two times, that are pretest and post-test.

Test the validity of the test by viewing the material criterion of validity, the right to use words and language. Questions used in accordance with the basic competencies and indicators, the question is not confusing. This test instrument validated by the teacher of the class before the test is tested to class and grade control experiment.

### **E. Try Out the Instrument**

The purpose of try out the instrument was to find out the quality of the research instrument that used for collecting the data, the test was chosen as the instrument tried out before hand. The try out was conducted on March 1<sup>st</sup>, 2019 to different subject that is on VII F class before truly conducting the research to the sample of research. It was supposed to determine the quality of the test as the instrument of the research. The result of the try out the instrument was analyzed statistically to know the validity and reliability that will be discussed below:

#### **1. Validity**

Validity is measuring what is the designed to measure. The most complex criterion of an effective test and the most important principle of language testing is validity. It is the extent to which inferences made from assessment result are appropriate, meaningful, and useful in term of the purpose of the assesment (Brown 2004:22). It is measurement that shows the validity of instrument. According to Ary (2010:225) validity is the most important consideration in developing and evaluating measuring instruments. Historically, validity was defined as the extent to which an

instrument measured what it claimed to measure. The focus of recent views of validity is not on instrument itself but on the interpretation and meaning of indicates how deep the instrument can measure the target of the research. In this study, the researcher used content and construct validity to measure whether the test.

a) Content Validity

A test is said to have content validity if its contents constitutes a representative sample of language skills, structures, etc being tested. In order to judge whether or not the test has content validity, we need a specification of the skill or structure being stated. A comparison of test specification and test content is basis for judgment for content validity. It can be concluded that content validity which is used to assesses how well the instrument samples the content domain being measured.

In this study, the test which was given twice at pre-test and post-test was in the form of multiple choices. The test was made up based on course objective the syllabus of seventh grade of MTs. Ma'arif Bakung. The test spesifications of instrument in this research could be shown from the table :

**Table 3.3 The Test Specifications of Instrument**

Basic Competence	Indicators	Test Item	
		Pre-test	Post-test
3.7 Comparing social functions, text structures, and linguistic elements of several oral and	1. Understanding and identifying the characteristics or information about someone.	12,13,14, 15,16,17, 18,19,20	8,9,10,11, 16,17,18, 19,20

written descriptive texts by giving and asking for information related to the description of people, animals, and objects that are very short and simple, according to the context of their use.	2. Understanding and identifying the characteristics or information about animal.	5,6,7	1,2,3
	3. Understanding and identifying the characteristics or information about thing.	8,9,10,11	12,13,14,15
	4. Understanding and identifying the characteristics or information about place.	1,2,3,4	4,5,6,7
	Total	20	20

$$\text{SCORE} = \frac{\text{Number of correct items}}{\text{Total questions}} \times 100$$

The table above shows the specifications of instrument which represented the material. Total all of question are 20 items in the form of multiple choices test and the value score of every question is 5 points, thus a total of all is 100 points.

#### b) Construct Validity

According to Ary *et al* (2010:231), construct validity focuses on test scores as a measure of a psychological construct. The word construct refers to any underlying ability which is hypothesized in a theory of language ability. It means that the instrument was made up based on the theory which the instrument would measure. In this research, the instrument has been constructed based on reading comprehension theory. After the instrument

was constructed, the test was tried out and then the researcher used SPSS 18.0 of Pearson Correlation to count the validity test per items.

Basic decisions making in validity testing per items are as follows:

1. If the score of  $R_{hitung} > R_{table}$ , it means that the test items is valid.
2. If the score of  $R_{hitung} < R_{table}$ , it means that the test items is not valid.

The process calculation of validity testing by using SPSS 18.0 version for windows found that the 20 questions of multiple choices which had been tried out were valid. The result of validity can be seen as follows:

**Table 3.4 The Result of Construct Validity**

No Items	R. Hitung	R. Table	Keterangan
Item 1	0.470	0.334	Valid
Items 2	0.849	0.334	Valid
Items 3	0.821	0.334	Valid
Items 4	0.420	0.334	Valid
Items 5	0.821	0.334	Valid
Items 6	0.560	0.334	Valid
Items 7	0.470	0.334	Valid
Items 8	0.849	0.334	Valid
Items 9	0.570	0.334	Valid
Items 10	0.821	0.334	Valid
Items 11	0.470	0.334	Valid
Items 12	0.771	0.334	Valid
Items 13	0.508	0.334	Valid
Items 14	0.560	0.334	Valid
Items 15	0.553	0.334	Valid
Items 16	0.349	0.334	Valid
Items 17	0.627	0.334	Valid
Items 18	0.349	0.334	Valid
Items 19	0.660	0.334	Valid
Items 20	0.470	0.334	Valid

## 2. Reliability

According to Bolarinwa (2015), reliability refers to the degree to which the results obtained by a measurement and procedure can be replicated.

Reliability was necessary characteristic of any good test for it to be valid at all. Reliability was an indicator of consistency, that was an indicator of how stable a test score or data is across applications or time. A measure should produce similar or the same results consistently if it measures the same “thing.” A measure can be reliable without being valid. A measure cannot be valid without being reliable (Hale *et al* 2014:45). It meant the test could be valid if it was reliable as well.

Reliability test instrument can be done by using Cronbach’s Alpha. According to Riduwan (2004:118), the criteria of reliability can be divided into 5 classes as follow:

**Table 3.5 Cronbach’s Alpha Interpretation**

<b>Cronbach`s Alpha</b>	<b>Interpretation</b>
0,00 – 0,20	Less Reliable
0,21 – 0,40	Rather Reliable
0,41– 0,60	Quite Reliable
0,61 – 0,80	Reliable
0,81 – 1,00	Very Reliable

The result of reliability testing by using SPSS 18.0 can be seen from the table:

**Table 3.6 Reliability Result**

Reliability Statistics	
Cronbach's Alpha	N of Items
.905	20

Based on the table 3.6 above, the results of reliability testing can be seen from the reliable value in the cronbach alpha. If the significance value is  $> 0.6$ , the data can be said to be reliable. From the table above it can be seen that the cronbach alpha column shows a significance of 0.905 which means  $> 0.6$  so that it can be said to be reliable.

## F. Normality and Homogeneity

### 1. Normality Testing

Normality is conducted to know whether the data distribution is normal or not. Before conducting the real analysis, the data of the research should be tested normality of distribution. Good data is the data in the normal distribution.

This normality test is intended to find out that the distribution of research does not deviate significantly from the normal distribution. One way to find out the value of normality is by using the Kolmogorov Smirnov formula which is assisted by using the SPSS 18.0 for Windows application.

From the *One-Sample Kolmogorov-Smirnov Test table*, the probability number or *Asymp. Sig (2-tailed)* is obtained. This value is

compared with 0.05 (in this case using a significance level or  $\alpha = 5\%$ ) for decision making with guidelines:

1. If the value of significance or probability  $< 0.05$  it means that the distribution data is not normal.
2. If the value of significance or probability  $> 0.05$  it means that the distribution data is normal.

**Table 3.7 Normality Test Pre-Post Test Experimental Group**

One-Sample Kolmogorov-Smirnov Test			
		pretest_eksp	posttest
N		47	47
Normal Parameters <sup>a,b</sup>	Mean	60.96	80.21
	Std. Deviation	8.382	8.782
Most Extreme Differences	Absolute	.175	.192
	Positive	.166	.192
	Negative	-.175	-.127
Kolmogorov-Smirnov Z		1.198	1.314
Asymp. Sig. (2-tailed)		.113	.063

a. Test distribution is Normal.

b. Calculated from data.

Based on the table 3.7 above, output One Sample Kolmogrov-Smirnov Test shows that sample of experiment class are 47 students. The Asymp. Sig (2-tailed) test pretest in Experiment class was 0.113 and test posttest was 0.063.

**Table 3.8 Data Normality Test Decision**

Nama Variabel	Nilai Asymp. Sig.(2-tailed)	Taraf Signifikansi	Keputusan
pretest	0,113	0,05	Normal
posttest	0,063	0,05	Normal

**Table 3.9 Normality Test Pre-Post Test Control Group**

		One-Sample Kolmogorov-Smirnov Test	
		pretest	posttest
N		45	45
Normal Parameters <sup>a,b</sup>	Mean	53.56	60.56
	Std. Deviation	6.625	7.247
Most Extreme Differences	Absolute	.191	.178
	Positive	.191	.178
	Negative	-.185	-.130
Kolmogorov-Smirnov Z		1.284	1.196
Asymp. Sig. (2-tailed)		.074	.114

a. Test distribution is Normal.

b. Calculated from data.

Based on the table 3.9 above, output One Sample Kolmogorov-Smirnov Test shows that sample of control class are 45 students. The Asymp. Sig (2-tailed) test pretest in control class was 0.074 and test posttest was 0.114.

**Table 3.10 Data Normality Test Decision**

Nama Variabel	Nilai Asymp. Sig.(2-tailed)	Taraf Signifikansi	Keputusan
pretest	0,074	0,05	Normal
posttest	0,114	0,05	Normal

## 2. Homogeneity Testing

Homogeneity test intended to show two or more group of data sample come from population having the same variance. Homogeneity testing is conducted to know whether the gotten data has a homogeneous variance or not. The basis of decision making in this homogeneity test is:

1. If the value of significance  $< 0.05$  it means that the variant of two or more groups is not the same.
2. If the value of significance  $> 0.05$  it means that the variant of two or more groups is the same.

**Table 3.11 Test of Homogeneity of Post-test Variances of Experiment Group**

**Test of Homogeneity of Variances**

posttest

Levene Statistic	df1	df2	Sig.
1.043	6	39	.413

Based on the output table above it is known that the Sig. student learning outcomes in the experiment class based on the value of the pretest and posttest is 0.413, meaning that the data on student learning outcomes in the experiment class based on the pretest and posttest values have the same or homogeneous variant.

**Table 3.12 Test of Homogeneity of Post-test Variances of control Group**

**Test of Homogeneity of Variances**

posttest

Levene Statistic	df1	df2	Sig.
.877	5	38	.506

Based on the output table above it is known that the Sig. student learning outcomes in the control class based on the value of the pretest and posttest is 0.506, meaning that the data on student learning outcomes in the

control class based on the pretest and posttest values have the same or homogeneous variant.

### **G. Data and Data Source**

The data were information or facts used in discussing or deciding the answer of research question. The source of data in the study was the subjects from which the data can be collected for the purpose of research, Arikunto (2010:129).

In this research, the data was obtained from students' reading comprehension score using and not using Question Answer Relationship (QAR) Strategy. The data source was the seventh grade of MTs. Ma'arif Bakung in the academic year 2018/2019.

### **H. Data Collecting Method**

To know the effectiveness of Question-Answer Relationship strategy in this quantitative research, the researcher used test as the instrument to collecting the data. Test is a set of question and exercises used to measured the achievement or capacity of the individual of group. In this research, there are two kinds of test, pretest and post-test that were given to the students.

#### **1. Pre-test**

The researcher gave pre-test for experimental group conducted on April 1<sup>st</sup>, 2019 and for control group conducted on March 30<sup>th</sup>, 2019. The pre-test was administered before treatment process. Pre-test is needed to know how far the students' reading comprehension ability. The pre-test

was in the form of multiple choices consisting of 20 items. This kind of tests was chosen to avoid subjectivity that may affect unreliability of the tests. Multiple choice is the most obvious advantage is that scoring can perfectly reliable. Scoring should also be rapid and economical. A further considerable advantage was that it was possible to include more items than other forms of other tests since the test-takers have only to make a mark on the paper. The time was allocated for this test was 30 minutes. There are 47 students in experimental group and 45 students in control group.

## 2. Treatment

This treatment procedure was adapted from Sejnost (2009:45) by applying the steps of using QAR strategy. Treatment here meant that the researcher applied Question Answer Relationship (QAR) Strategy in teaching process. The treatment was given to experimental class two times. The treatment was given twice due to the limited time. The process of this strategy described as follow:

- 1) Introducing the concept of QAR strategy by explaining each types of question (In the Book and In My Head), providing a clear example of each, and discussing the difference in each.
- 2) Giving a short piece of text for students to read, in this case was a descriptive text.
- 3) Leading the students through the process of answering each type of question and making sure that they go back into text to verify their answer and they understand the whole of questions.

- 4) Increasing the number of questions of each type, until students are clearly able to understand the differences among types and can identify them with ease.
- 5) Asking students to read a longer passage and develop a set of questions for their classmates to identify and answer.

This teaching strategy would be taught for experimental group on April 1<sup>st</sup> and April 5<sup>th</sup>, 2019, while the control group would be taught without using Question Answer Relationship (QAR) Strategy.

### 3. Post-test

After the treatment, the post-test was given to the students for experimental group on April 5<sup>th</sup>, 2019 and for control group on April 1<sup>st</sup>, 2019. The test items in the post test the same as the pre-test, but it has the difference number from the pre-pest. The purpose of the post-test was to measure students' reading comprehension ability after being taught by using QAR strategy. The post-test was in the form of multiple choices which is consisting of 20 items. The time that was allocated for this post-test was 30 minutes.

## **I. Data Analysis**

Data analysis is a process of analyzing the acquired from the result of the research. Ary et al (2010:95) explains that data analysis indicate how the researcher will analyze the data to the test the hypothesis and/or answer the

research question. While Khotari (2004:18) explains after the data have been collected, the researcher turns to the task of analyzing the data.

The data was analyzed in order to see whether the students' descriptive text reading comprehension before and after being taught by using Question-Answer Relationship strategy is effective or not. Quantitative analysis was done using statistic which is called statistical analysis or inferential statistic. The quantitative data of this research in analyzed by using statistical computation. The data collected was processed by comparing the score between control group and treatment group to see whether there is significant different between students who given by treatment and not. In this research the researcher was used independent sample test through SPSS 18.0.