

### III. RESEARCH METHODS

This chapter discusses about the methods of research used in this study, such as research design, population and sample, data collecting technique, research procedures, and criteria of good test, data analysis, data treatment, and hypothesis testing.

#### 3.1 Research Design

In this research, the researcher intends to find out the significant increase of students' reading comprehension achievement by comparing graphic organizer and literal translation in reading. To gain the objectives of the research, the researcher conducted quantitative design with *pretest posttest experimental group design*. The researcher selected two classes, one as the experimental group one and another as the experimental group two. According to Hatch and Farhady (1982:22) the design of the research is described as follows:

<b>G1</b>	<b>:</b>	<b>T1</b>	<b>X<sub>1</sub></b>	<b>T2</b>
<b>G2</b>	<b>:</b>	<b>T1</b>	<b>X<sub>2</sub></b>	<b>T2</b>

*Notes:*

G1	:	experimental group one
G2	:	experimental group two
T1	:	pre-test
T2	:	post-test

$X_1$  : treatment for Experimental Group one (Graphic Organizer Technique)

$X_2$  : treatment for Experimental Group Two (Literal Translation Technique)

In this research the pretest was conducted by implementing a reading narrative test provided by the researcher. The posttest was given by the end of the treatment. The treatment (X) was conducted three times. Before the test used to collect the data, the researcher firstly tried it out to the students. The try-outtest was held before all the tests and treatments would be given to the students. The aim of this try out test was to find out the quality of the items of tests which were used in this research.

## **3.2 Population and Sample**

### **3.2.1 Population**

The population of the research was the first year students of SMAN 1 Natar. The researcher chose the first year students in the second semester of academic year 2012/2013. There were nine classes of the first year students and each class consisted of 30 to 34 students. Their ages range from 15-16 years old. From those nine classes, there were around 288 students as the population of the research

### **3.2.2 Sample**

Based on the population above, three classes were taken as the sample of this research, the first class as try out class, the second class as the experimental group one, and the last class as the experimental group two. The two sample classes of this research were selected using simple random sampling. Those classes

were selected randomly by using lottery, since the first year students in SMAN 1 Natar was not stratified class. There was no priority class. It was applied based on consideration that every class in the population had the same chance to be chosen and in order to avoid the subjectivity in the research. Next, to determine which class was as the experimental class one and experimental class two, the researcher used coin by flipping it.

### **3.3 Data Collecting Technique**

In collecting the data, the researcher used the following steps:

#### **1. Administering the Pre-test**

The pre-test was given before the treatment, in order to find out how far the competence of the students in reading comprehension or their input before the treatments and to find out the experimental class' reading comprehension achievement, the test was multiple choices that consist of 20 items. The material was related to the curriculum used in the school and suitable with their level. The pretest was conducted within 60 minutes.

#### **2. Administering the Post-test**

The post-test was given after the treatment in order to find out whether there was any increase of students' reading comprehension achievement. The test was multiple choices consisted of 20 items. The materials tested, were related to the curriculum used in the school and suitable with their level. The post-test was done after four meetings of the treatments. The result of the post-test of the participant class was analyzed.

### 3.4 Research Procedures

There were some procedures applied for taking the data:

1. Determining the population and the sample.

The researcher takes three classes to determine as try out class, experimental class one, and experimental class two.

2. Determining the research instrument

Both the reading tests (pretest and posttest), the material was taken from students' textbook and from authentic materials, i.e. magazines and internet.

The researcher used four reading texts in her research. The topics of the reading text were about fairy tales and realistic fictions.

3. Administering try-out.

This was expected to measure the validity, reliability, level of difficulty, and discrimination power, to ensure the data used by the researcher is valid and reliable as the research instruments. This test was multiple choice tests and was conducted in 80 minutes. There were 35 items of multiple choices with four options and one of them was as the correct answer and the rest were the distracters, the test items could be reduced or kept depends on its reliability and validity. The aim of try-out was to determine the quality of the test used as the instrument of the research, and to determine which item revised for the pre-test and the post-test. This research used the result of the try-out test to measure the level of difficulty and discrimination power, to find out the validity and reliability of the test.

4. Administering the pretest

This test aimed to find out the students' basic reading comprehension. The researcher administered pretest before giving treatments by using reading

test and 25 items of multiple choices text. The scoring system was that the load of each correct answer of each correct answer was 5 points. So, if one participant answers all the items correctly, students would get 100 points.

The pretest takes 60 minutes.

#### 5. Conducting treatments

In this research the researcher taught the students reading comprehension in reading narrative text using graphic organizer for the experimental class one, and those who would be taught through literal translation technique for the experimental class two. The researcher gave three times of treatments in three meetings, which took 90 minutes in every meeting.

#### 6. Administering the posttest

The researcher conducted the posttest in order to find out the students achievements in reading comprehension after giving the treatments, which was using graphic organizer as their reading strategy. The researcher administered posttest after giving treatments by using reading text and 25 items of multiple choices test. The scoring system was that the load of each correct answer is 5 points. The posttest took 60 minutes.

#### 7. Analyzing the data (pretest and posttest)

This step was conducted in order to find out the students reading comprehension achievement. *Independent group T-Test* formula was used to compare the means of the pretest and posttest of both two groups. The data was computed through the Statistical Package for Social Sciences (SPSS) version 15.0.

### **3.5 Criteria of Good Test**

A test is said to have a good quality if it had a good validity, reliability, level of difficulty and discrimination power.

#### **1. Validity**

Validity referred to the extent to which the test measures what was intended to measure. This meant that it related directly to the purpose of the test (Shohamy, 1985: 74). A test could be said valid if the test measured the object to be measured and suitable with the criteria (Hatch and Farhady, 1982: 250).

According to Hatch and Farhady (1982: 251), there were four basic types of validity: face validity, content validity, construct validity and empirical or criterion-related validity. To measure whether the test had good validity, the researcher used content and construct validity since the other two are considered be less needed. Face validity only concerns with the layout of the test. Criterion-related validity concerned with measuring the success in the future, as in replacement test (Hatch and Farhady, 1982:251). The two types were used in this research are:

##### **a. Content validity**

Content validity referred to the extent to which a test measures a representative sample the subject matter contents, the focus of the content validity was adequate of the sample and simply on the appearance of the test (Hatch and Farhady, 1982:251). To know whether the test was good reflection of what would be taught and of the knowledge which the teacher wanted the students to know, the

researcher compared this test with table of specification. If the table represented the material that the researcher wanted to test, then it was valid from that point of view. A table of specification was an instrument that helped the test constructor plan the test.

Table 1. Table specification of try out

No	Objectives	Item Numbers	Total Items	Percentage
1	Identify the main idea	1, 9, 15, 19, 26,27	6	18 %
2	Vocabulary	6, 7, 17, 18, 24, 25, 33, 35	8	22 %
3	Specific information	4, 10, 12 ,13, 14, 21, 23, 28, 30, 32	10	28 %
4	Inference	2, 3, 11, 20, 22, 29, 31	7	20 %
5	Reference	5, 8, 16, 34	4	12 %
TOTAL			35	100%

## 2. Construct Validity

Construct validity was concerned with whether the test was actually in line with the theory of what reading comprehension means. To know the test was true reflection of the theory in reading comprehension, the researcher examined whether the test questions actually reflected the means of reading comprehension or not.

## 3. Reliability

Reliability referred to the extent to which the text was consistent in its score, and gave us an indication of how accurate the test score were (Hatch and Farhady, 1982: 244). To test the reliability of the instruments, the researcher used *split-half* method in which the reading tests were divided into halves (Hatch and Farhady, 1982: 246). By splitting the test into two equal parts (first half and second half); it was made as if the whole tests had been taken in twice.

To measure coefficient of the reliability of the first and second half group, the researcher used the following formula:

$$r_{xy} = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}}$$

*Notes:*

$n$  = number of students

$r$  = coefficient reliability between first and second half

$x$  = total number of first half

$y$  = total number of second half

$x^2$  = square of  $x$

$y^2$  = square of  $y$

$\sum x$  = total score of first half items

$\sum y$  = total score of second half items

(Hatch and Farhady, 1982: 222)

Then to know the coefficient correlation of the whole items, Spearman Brown's Paraphrase Formula was used. The formula was as follows:

$$rk = \frac{2rl}{1 + rl}$$

Where:

$rk$  = the reliability of full test

$rl$  = the reliability of half test

The criteria of reliability are:

0.90- 1.00 = high

0.50- 0.89 = moderate

0.0- 0.49 = low

#### **4. Level of Difficulty**

To see the index of difficulty, the researcher used the following formula:



$$LD = \frac{R}{N}$$

Where,

LD = level of difficulty

R = the number of the students who answer correctly

N = the total number of the students

The criteria are:

< 0.30 = Difficult

0.30- 0.70 = Average

> 0.70 = Easy

(Heaton, 1975: 182)

## 5. Discrimination Power

The discrimination power (DP) is the proportion of the high group students getting the items correct minus the proportion of the low-level students who getting the items correct. In calculating the discrimination power of each item, the following formula is used:

$$DP = \frac{U - L}{1/2 N}$$

Where,

DP = Discrimination Power

U = Number of upper group who answer correctly

L = Number of lower group who answer correctly

N = Total number of the students.

The criteria are:

DP: 0.00-0.19 = Poor

DP: 0.20-0.39 = Satisfactory

DP: 0.40-0.69 = Good

DP: 0.70-1.00 = Excellent

DP: - (negative)= Bad items, should be omitted

(Heaton, 1975: 182)

### 3.6 Data Analysis

The researcher computed the students' score in order to find out the students' achievement in reading narrative text through graphic organizer using the following steps:

- Scoring the pre-test and post-test.
- Tabulating the results of the test and calculating the score of the pre-test and post-test.
- Drawing conclusion from the tabulated results of the pre-test and post-test administered, that is by statistically analyzing the data using statistical computerization i.e. *Independent Groups T-Test of Statistical Package for Social Science (SPSS) version 15.0 for windows* to test whether the increase of students' gain is significant or not, in which the significance is determined by  $p < 0.05$ . It is used as the data come from the two samples (Hatch and Farhady, 1982:111).

### 3.7 Treatment of the Data

In order to determine whether the data are good or not, the researcher would analyze the data by:

1. Scoring the pre-test and post-test.

2. Tabulating the result of the thesis and calculating the mean of the pretest and posttest. To compute the average score or mean of the pretest and posttest, the researcher would use a very simple statistic formula as follows:

$$\bar{X} = \frac{\sum x}{N}$$

Notes:

$\bar{X}$  : mean (average score)

$\sum x$ : total number of the student's score

N : total number of the students

(Hatch and Farhady, 1982:5)

3. Calculating from the tabulated results of the pretest and posttest administered that is by statistically analyzing the data using *t-test* to test whether or not the difference between pretest and posttest is significant. It is used as the data comes from the same sample or known as paired data (Hatch and Farhady, 1982).

4. Administering Random Test

This test is used to make sure whether the data is random or not (Hatch and Farhady is quoted by Setiyadi, 2006: 168-169). The researcher uses SPSS version 15.0 to analyze the data. The hypotheses for the random test are as follow:

$H_0$ : the data is not random

$H_1$ : the data is random

In this research, the criteria for the hypotheses are  $H_1$  is accepted if  $p > \alpha$ , and the researcher uses level of significance 0.05.

#### 5. Administering the Normality Test

This test is used to measure whether the data in two classes are normally distributed or not. The data are tested by One-sample Kolmogorov-Smirnov Formula (SPSS 15). The hypothesis is accepted if the result of the normality test is higher than 0.05 ( $\text{sign} > \alpha$ ). In this case, the researcher uses level of significance of 0.05

#### 6. Administering the Homogeneity Test

This test is used to know whether the data of the posttest from the experimental class 1 and from the experimental class 2 are homogeneous or not. The data is tested by Independent Sample Test (SPSS 15). The criteria for the homogeneity of pre test are:

$H_0$ : There is no significant difference in the level of ability (equal)

$H_1$ : There is a significant difference in the level of ability (not equal)

The criteria for the hypothesis is  $H_1$  is accepted if the result of Homogeneity test of pretest is higher than 0.05 ( $\text{Sign} > \alpha$ ).

### 3.8 Hypothesis Testing

After collecting the data, the researcher recorded and analyzed them in order to find out whether there is an increasing in students' ability in reading

comprehension of narrative text or not after the treatment. The researcher used Independent Group T-test to know the level of significance of the treatment effect.

The formulation is: 
$$t_{obs} = \frac{\bar{X}_e - \bar{X}_c}{S_{(\bar{X}_e - \bar{X}_c)}}$$

With: 
$$S_{(\bar{X}_e - \bar{X}_c)} = \sqrt{\left(\frac{S_e}{\sqrt{n_1}}\right)^2 + \left(\frac{S_c}{\sqrt{n_2}}\right)^2}$$

$\bar{X}_e$  : Mean from the difference pre-test and post-test of experimental class one and experimental class two

$\bar{X}_c$  : Mean from the difference pre-test and post-test of experimental class One and experimental class two

$S_{(\bar{X}_e - \bar{X}_c)}$  : Standard error of differences between means

$n$  : Subjects on sample

(Hatch and Farhady, 1982:111)

The criteria are:

If the t-ratio is higher than t-table :  $H_1$  is accepted

If the t-ratio is lower than t-table :  $H_0$  is accepted