# 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, data analysis, and hypothesis testing.

## 3.1 Research Design

To gain the objective of the research, the researcher conducted quantitative design with *pretest posttest control group design* which belongs to the true experimental designs because a control group was present for comparison purposes. The design allowed the researcher to interpret the findings with more confidence. Furthermore, Hatch and Farhady (1982:22) mention the characteristics of true experimental designs:

a. A control group (or groups) is present,

b. The class are randomly selected and assigned to the groups, and

c. Pretest is administered to capture the initial differences between the groups they argue further that those three basic characteristics will allow the researcher to avoid almost all the problems associated with internal and external validity.

Since, the researcher used true experimental design; there were two classes of research participants: one as experimental group which got the treatment through graphic organizer technique and the other as a control group which was taught with literal translation technique as usual by the classroom teacher. Having a control group allowed the researcher to justify the effect of his treatments. Both classes received the same pretest and posttest. The pretest was administered to see the students' basic ability within both groups in order to ensure that their ability was equal before treatments. Then, after giving the treatments the researcher administered the posttest to the both groups. The posttest was administered in order to prove that graphic organizer technique can be used to improve students' reading comprehension in experimental group, and also to see whether there was any significant difference of students' reading comprehension between both groups.

According to Hatch and Farhady (1982: 22), the research design is presented as follows:

$$G1 : T1 \quad X_1 \quad T2$$

$$G2 : T1 \quad X_2 \quad T2$$

*Notes:*

| | |
|---|---|
| G1 | : experimental group |
| G2 | : control group |
| T1 | : pre-test |
| T2 | : post-test |
| $X_1$ | : treatment for Experimental Group (Graphic Organizer Technique) |
| $X_2$ | : treatment for Control Group (Literal Translation Technique) |

**3.2 Population and Sample**

The population and sample that was used in this research were explained as the followings:

**3.2.1 Population**

The population of the research was the first grade students of SMA Negeri 1 Kalianda. There were six classes of the first year students and most classes consist of 35 students. From those six classes, there were around 210 students as the population of the research.

**3.2.2 Sample**

From 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, and the last class as the control group. The classes were chosen randomly by cluster sampling through lottery. The researcher prepared six papers cut that represented the all first grade classes, and asked the English teacher to choose three papers cut. The first paper cut was the try out class, the second paper cut was the experimental group, and the last paper cut was the control group. Class X.5 was chosen as try out class, X.3 as experimental group, X.1 as control group. The try out class consist of 34 students, but only 30 students who followed the try out test because the other 4 students were absent. Whilst, only 30 students of X.3 and X.1 who followed the pre-test, treatment of the research and post-test.

**3.3 Data Collecting Technique**

The data were collected by administering the pretest and posttest for both classes investigated.

**1. Pre-test**

*Pre-test* is the way to measure students' ability in the beginning before giving treatment. The test consisted of 20 items of multiple choice forms with five options. The materials were taken from some of try-out test items. The test conducted within 45 minutes. The researcher used pre-test because it was very important to know students' ability at the beginning in order to compare it with students' ability after treatment. Pre-test was given in order to know how far the competence of the students in reading comprehension before the treatment and to know whether both experimental group and control group were equal or not in the term of their reading comprehension achievement.

**2. Post-test**

*Post-test* is the test that is given to the students after giving treatment. The test consisted of reading text with 20 items of multiple choice tests. The post-test was conducted within 45 minutes. Post-test was very important in this research because it represented the result of the students' reading comprehension achievement after the treatment. The researcher wanted to find out whether there was significant difference in students' reading comprehension achievement between students who got the treatment through graphic organizer and those who was taught through literal translation technique, in order to investigate whether there was positive effect of graphic organizer technique on the improvement of students' reading comprehension.

**3.4   Research Procedures**

In conducting this study, the researcher conducted the following procedures:

**1.   Determining the research problem**

Based on the researcher's background of problem in Chapter I, it was assumed that the graphic organizer technique could be used to improve the students reading comprehension achievement. And he tried to find out whether there was significant improvement of learner's achievement in reading comprehension skill before and after being taught through graphic organizer technique.

**2.   Selecting materials**

The materials (English recount text) were taken from the authentic materials and also from students' text book based on the KTSP. The selecting process considers materials that had been taught to the students and the students' interest.

**3.   Determining the samples of the research**

Based on the population and sample explained previously, three classes of first grade students of SMA N 1 Kalianda were taken as the sample of this research, the first class as try out class, the second class as the experimental group, and the last class as the control group. The three classes were chosen randomly by cluster sampling through lottery.  And the result, Class X.5 was chosen as try out class, X.3 as experimental group, X.1 as control group.

**4. Determining research instrument**

For both reading test (the pre-test and the post-test), the researcher made the test by himself based on to the students' hand book English for first grade of senior high school based on the *Kurikulum Tingkat Satuan Pendidikan*(KTSP).

**5. Conducting try-out**

The try-out test conducted before the pre-test was organized. This was intended to measure the validity and reliability of pretest and posttest, to ensure that the data used by the researcher was valid and reliable to be used as a research instrument. This test was multiple choice tests and conducted in 90 minutes. There were 50 items of multiple choices with five options and one of them as the correct answer. The total score was 100 point, so if the students answered the whole questions correctly they would get 100 point. The aimed of try -out was to determine the quality of the test which would be used as the instrument of the research, and to determine which item should be 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, discrimination power, and also to find out the validity and reliability of the test.

Based on the result of the difficulty level and discrimination power, some items were dropped and revised to administer in pretest and posttest. The try out test consisted of 10 easy items, 36 average items and 4 difficult items. Meanwhile, for discrimination indexes, 6 items were bad, 14 items were poor, 9 items were good,

20 items were satisfactory, and 1 item was excellent. For detail information, see appendix 4.

Items that have average remark in the level of difficulty and excellent or satisfactory or good for the discrimination power were selected for the test (3., 4., 5., 7., 8., 9., 10., 11., 13., 15., 16., 17., 19., 20., 21., 22., 25., 26., 28., 30., 34., 35., 36., 37., 38., 41., 43., 44., 46., 48.). Moreover, the items that were poor but had average criterion in difficulty level were revised. The revised items are : 12., 14., 18,. 27., 31., 39., 40., 49. The items with negative and zero discrimination power were dropped.

## 6. Conducting pretest

Pretest was administered to find out the students' initial scores on reading comprehension before the treatments. The test was conducted before the treatment, and 45 minutes was distributed in this reading test. This pretest was in form of multiple choices with 20 questions taken from the try out test.

## 7. Implementing the treatment by using graphic organizer technique.

The graphic organizer technique was trained in two weeks. The treatment conducted in 2 meetings and 2 x 45 minutes was distributed for each meeting consist of the use graphic organizer technique for the experimental group and literal translation technique for the control group

## 8. Administering posttest

Both of groups received the same posttest as well as pretest. Since graphic organizer technique on reading text was supposed to be able to improve the

students' reading comprehension, administering the some posttest for both groups, therefore, could be very much useful to justify that the improvement of students' reading comprehension within experimental group was really because of the effect of treatments through graphic organizer technique. The posttest was in the form of multiple choices with 20 items and five options for each test item. It was administered in 45 minutes.

## 9. Analyzing the result of the test

The researcher analyzed the data by using *Independents Group T-Test* because the data are collected from two different groups (experimental group and control group) and also taken from different situation. This formula used to compare the average score (mean) of both groups and to investigate whether there was improvement of students' reading comprehension through Graphic organizer technique. Before determining whether there was improvement, the raw data put in tables by putting the highest score on the top. The data was computed through the SPSS version 17.0.

## 3.5 Data Treatment

There were several steps in doing the data treatment. First, the try out test result were computed. Here the reliability, level of difficulty and discrimination power of the test was computed. Second, the result of pretest and posttest were searched simultaneously with the normality, homogeneity and random. The complete procedure can be seen as the following.

**3.5.1 The Treatment of Try out Test Result**

A good test should meet four criteria: a good validity, reliability, and level of difficulty and discrimination power.

**1) Validity of the test**

A test can be said to be valid if it measures the object to be measured and suitable with the criteria (Hatch and Farhady, 1982:250). According to authors (1982: 251), there are four types of validity: face validity, content validity, construct validity and empirical or criterion-related validity. To measure whether the test has good validity, the researcher used content and construct validity since the other two were considered be less needed. Face validity only concerns with the layout of the test. While, Criterion-related validity is concerned with measuring the success in the future, as in replacement test (Hatch and Farhady, 1982:251). The two types used in this research were:

**a)  Content Validity**

*Content validity* is the extent to which a test measures a representative sample of the subject matter content, the focus of content validity is adequacy of the sample and simply on the appearance of the test. (Hatch and Farhady, 1982:251). To know whether the test is good reflection of what has been taught and of the knowledge which the teacher wants the students to know, the researcher compared this try out test with the table of specification. If the table represents the material that the researcher wants to test, then it is valid from the point of view. A table of classification is an instrument that helps the test constructor plans the test.

**Table 2. Table of Specification**

| No | Objectives | Items Number | Percentage |
|----|-----------|-------------|-----------|
| 1 | Identify the main idea | 1., 6., 7., 18., 22., 32., 40., 41., 49. | 20% |
| 2 | Vocabulary | 2., 9., 20., 24., 30., 34., 39., 44., 46., 47. | 20% |
| 3 | Finding detail information | 11., 13., 14., 16., 17., 21., 23., 26., 27., 29., 31., 35., 36., 37., 42. | 30% |
| 4 | Inference | 3., 8., 15., 19., 25., 33., 38., 43., 48. | 14% |
| 5 | Reference | 4., 5., 19., 12., 28., 45., 50 | 16% |

This table of specification is also based on the English syllabus especially for reading comprehension that was made by the English teacher of SMA 1 Kalianda.

**b) Construct Validity**

*Construct validity* is concerned with whether the test is actually in line with the theory of what reading comprehension means. (Hatch and Farhady, 1982:252)

To determine the test is true reflection of the theory in reading comprehension, the researcher examined whether the test questions actually reflected the means of reading comprehension or not.

**2) Reliability of the Test**

Reliability refers to the extent of test to which the test is consistent in its score and gives us an indication of how accurate the test score are (Hatch and Farhady, 1982:244).

Reliability of the test can be determined by using the Spilt half method in order to estimate the reliability of the test. To measure coefficient of the reliability the first and second half group, the researcher used the following formula:

$$r_l = \frac{\sum XY}{\sqrt{\left[\sum X^2 \sum Y^2\right]}}$$

*Notes:*

rl          : The coefficient of reliability between odd and even number

X          : The total score of odd number

Y          : The total score of even number

$X^2$      : The square of X

$Y^2$      : The square of Y

 (Lado in Hughes, 1991:3)

After getting the coefficient of reliability between odd and even number, then the researcher used the result in  "Spearman Brown's Prophecy Formula" (Hatch and Farhady, 1982: 256) to determine the reliability of  whole test items as follow:

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

rk          : the reliability of the test

r1          : the reliability of half test

Then, the criteria of reliability are:

0.90 – 1.00     =high

0.50 – 0.89     =moderate

0.00 – 0.49     =low

The result of the reliability found through this research is 0.9887 (see appendix 6). By referring to the criteria of the reliability proposed by Hatch and Farhady (1982:247), the test has high reliability that is in the range of 0.90-1.00.

It is indicated that the instrument produced consistent result when administered under similar condition, to the same participant and in different time (Hatch and Farhady,1982:244).

**3) Level of difficulty**

The level of difficulty is used to classify the test items into difficult items and easy ones. The items should not be too difficult or too easy for the students. To see the level 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 |

(Shohamy, 1985:79)

Based on the try out test related to those criteria there are 10 easy items (3., 14., 15., 19., 20., 21., 32., 40., 43., 49.), 36 average items (1., 2., 4., 5., 7., 8., 9., 10., 11., 12., 13., 15., 17., 18., 22., 23., 24., 26., 27., 28., 30., 31., 32., 33., 34., 35., 36., 37., 38., 41., 42., 44., 45., 46., 47., 48., 50.), and 4 difficult items (6., 25., 29., 39.), see appendix 4.

**4) Discrimination Power**

Sometimes an important feature of a test is the capacity to discriminate among the different candidates and to reflect the differences in the performance of the candidates in the group (Heaton, 1991:165). Discrimination power benefits to indicate the discrimination of the fail and the success of the students. To see the discrimination power, the researcher used the following formula:

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

*Notes:*

DP : the discrimination power

U : the number of upper group who answers correctly

L : the number of lower group who answers correctly

N : the total number of the students

In determine where is the upper and lower group, the researcher uses the Mean score of the total correct items answered correctly by the students in the try out class. The Mean score of the test, is the total correct items of try out test divided with the total number of the students who followed the try out test (840/30=28,04), see appendix 3 and 4 for more detail. If the student's correct answer is higher than the Mean score , it belongs to the upper group, and if it is lower than the Mean score, it belongs to the lower group.

And the criteria of the discrimination power are:

1. If the value is positive, it has discrimination because a large number or more knowledgeable students than poor students get the item correct. If the value is zero, it means no discrimination.

2. If the value is negative, it has negative discrimination because more low-level students than high level students get the item correct.

3. In general, the higher discrimination index, the better, in the classroom situation most items should be higher than 0.20 index.

    (Shohamy, 1985:82)

In accordance with Shohamy (1985: 81), there are some criteria of discrimination power of an item. An item is excellent if the discrimination index ranges from 0.75 to 1.00. A good item ranges from 0.41 to 0.70. A satisfactory item ranges from 0.21 to 0.40. An item is poor if the discrimination index ranges from 0.00 to 0.20, and an item is bad if the discrimination index is negative.

Based on the try out test related to those criteria there are 6 bad items ( 6., 24., 29., 42., 48., 50.) , 14 items are poor (1., 2., 12., 14., 18., 23., 27., 31., 32., 33., 39., 40., 45., 49.), 9 items are good (4., 5., 10., 11., 15., 22., 28., 38., 48.), 20 items are satisfactory (3., 7., 8., 9., 13., 16., 17., 19., 20., 21., 26., 27., 30., 34,, 35., 36., 41., 43., 44., 46.), and 1 item is excellent (37.).

### 3.5.2   Treatment of Pretest and Posttest Result

In order to determine whether the data are good or not, the data should meet the criteria, which should be normally distributed, homogenous, and random.

### 1.   Normality test

Normality test is used to test whether the input data is normally distributed or not. It is required by some complicated statistical tests, such as student's T-Test, one-way and two-way ANOVA, because they make assumption that data comes from a normally distributed population, and if such assumptions are not valid, the

results of the test will be unreliable. The data can be tested by Chi-quadrate

Formula or One-sample Kolmogorov-Smirnov Formula (SPSS 17.0).The criteria

for Random test are:

$H_1$: the data is random

$H_0$: the data is not random

The hypothesis is accepted if the result of random test is higher than 0.05 (sign >

α). In this case, the researcher used 0.05, level of significance.

### 2. Homogeneity test

This test is used to know whether the data in two classes are homogeneous or not.

It is often important to determine the homogeneity of the test if a set of data is

homogeneous before any statistical technique is applied to it. Homogeneous data

are drawn from a single population. In other words, all outside processes that

could potentially affect the data must remain constant for the complete time

period of the sample. If the data are not homogeneous, the results of the test will

be unreliable, because they are collected from two samples of the population that

does not have same level of proficiency in English language. The data can be

tested by Barleth Test or Independent Sample Test (SPSS 17.0) .The criteria for

the homogeneity of pre test are:

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

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

The criterion for the hypothesis is: $H_1$ is accepted if the result of Homogeneity test

of pre test is higher than 0.05 (Sign > α).

3. **Random test**

This test is used to determine whether the data are randomly selected or not. The data that come from random sample in a population is one of the basic assumptions that must be fulfilled in order to test it in parametric analysis *(T-Test)*. In the analysis comparison of *T-Test,* in accordance with Setiyadi (2006:168-169) there are three basic assumptions that must be fulfilled – they are:

a. The data is interval or ratio data.
b. The data are taken from random sample in a population (if the researcher wants to generalize the findings).
c. The data has a normal distribution.


If the data are not randomly selected, the researcher can not test it by using parametric analysis (*T-Test*), because the data do not meet the three basic assumptions for this test. Furthermore, the researcher also can not generalize the findings for all of the population. To test the data whether it is randomly selected or not, the researcher used Jacob Cohen Formula or Statistic Formula in Descriptive Formula (SPSS 17.0). The criteria for the homogeneity of pre test are:

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

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

The criterion for the hypothesis is: $H_1$ is accepted if the result of Homogeneity test of pre test is higher than 0.05 (Sign > $\alpha$).

**3.6   Hypothesis Test**

After collecting the data, the researcher recorded and analyzed the data in order to test the hypothesis. The hypothesis was analyzed by using Independent Group T-Test to compare the mean of posttest result of both classes. The formulation is:

$$t_{obs} = \frac{\overline{X}_e - \overline{X}_c}{S_{(\overline{X}_e - \overline{X}_c)}}$$

with

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

$\overline{X}_e$     : mean from the difference pre-test and post-test of experimental group

$\overline{X}_c$     : mean from the difference pre-test and post-test of control group

S $^{(\overline{X}_e - \overline{X}_c)}$ : standard error of differences between means

N        : subjects on sample

(Hatch and Farhady, 1982:111)

The criteria of this calculation are:

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

2.  If the t-value is lower than t-table: $H_0$ is accepted.