III.   RESEARCH METHODS

This chapter deals with certain points; research design, population and samples, variables, data collecting technique, scoring system, research procedure, data treatment and hypothesis testing.

3.1.   Research Design

This research is a quantitative research, in which data tend to use statistic as measurement in deciding the conclusion (Hatch and Farhady: 1982). In conducting this research, the researcher used Static Group Comparison Design that deals with two groups, the first one as experimental group one that was given treatment through group work and another one as experimental group two that was given treatment through individual work. The research design is as follows:

$$G = T1\ X1\ T2$$
$$G2 = T1\ X2\ T2$$

In which:

G1   : Experimental group 1

G2   : Experimental group 2

T1   : Pre-test

T2   : Post-test

X1   : Treatment (applying group work technique)

X2   : Treatment (applying individual work technique)

(Hatch and Farhady, 1982:22)

In this research, both classes was received the same pretest and posttest. Pretest was conducted by using a reading test provided by the researcher and the posttest was given after the treatments have done. In which, the treatment was conducted for three times.

3.2.   Population and Samples

A population can be defined as the whole subjects of a research. And the population of this research was the first year students of SMA Negeri 15 Bandar Lampung in the 2012/2013 academic year. The reason why researcher choose that school was because the location near from Lampung University and besides that based on researcher's pre-observation, the school was needed this research to determine the significant different of student's reading comprehension achievement between the students who are taught by using group work technique and the students who are taught by using individual work technique.

In this research, Grade X was chosen considering the previous knowledge they have from previous semester in which grade X consists of six classes and each class consist about 32 until 39 students. Class X.1 was ignored since the class belongs to the superior; it means that the students in the class were selected from all students who get high achievement in all subjects. The research took one class

as try out class and two classes as the sample of the research, one as experimental class 1 and another one as experimental class 2. In choosing those three classes the researcher used simple random probability sampling to make sure that the students' abilities were homogeneous or not by seeing the data of the teacher in the school. The classes that supposed have homogeneous ability were Class X.3, Class X.4, and Class X.5, and Class X.6.

Therefore, this research took Class X.3 as try out class, Class X.4 as experimental class 1 and Class X.5 as experimental class 2. Considering the data of students' score from the teacher and also the data of students' pretest in experimental class 1 and experimental class 2 (see Appendix 20), it can be stated that the two experimental classes were equal in term of their ability before implementing the techniques that are used in this research.

3.3. Variables

A variable can be defined as an attribute of a person or an object which varies from object to object (Hatch and Farhady, 1982:12). Besides, in order to assess the influence of the treatment in research, variables can be defined as independent and dependent variables. According to Hatch and Farhady (1982:15), the independent variable is the major variable that a researcher hopes to investigate, and the dependent variable is the variable that the researcher observes and measures to determine the effect of the independent variable. In this research, the

independent variable is used as the treatment variable. And the writer proposed three variables in this research, as follows:

1. Reading comprehension as dependent variable (Y)
2. Group work technique as independent variable 1 (X_1)
3. Individual work technique as independent variable 2 (X_2)

3.4. Data Collecting Techniques

In this research, the researcher used reading test as the instrument in collecting the data. There were two kinds of test that is pretest and posttest. But, before that, to prove whether the research instrument has good quality or not, it must be tried out first to measure its validity and reliability. Then, the data was obtained from the student's pretest and posttest scores. Overall, the test can be said to have a good quality if it has good validity, reliability, level of difficulty, and discrimination power, as they were elaborated in the following section:

3.4.1. Try Out Test

This test was conducted to find out whether the test items that would be used for pretest and posttest. The tests can be said good or not by proving several criteria; validity, reliability, level of difficulty, and discrimination power, as they were elaborated in the following section:

1. Validity

Validity refers to the extent to which an instrument really measures the objective to be measured and suitable with the criteria (Hatch and Farhady, 1982:250). A test must aim to provide a true measure of the particular skill which it is intended to measure. There are four types of validity, that's: (1) face validity: concern with the layout of the test; (2) content validity: depends on a careful analysis of the language being tested; (3) construct validity: measures certain specific characteristic in accordance with a theory of language learning; (4) criterion-related validity: concerns with measuring the success in the future, as in replacement test. In this research, the researcher used content validity and construct validity since the others are considered to be less needed.

a) Content Validity

According to Hatch and Farhady (1982:251) content validity is the extent to which the test measures a representative sample of the subject matter content. Content validity concerns whether the test are good reflection of the materials that need to be tested. To get content validity of the test, the researcher adopted the materials based on the objective of teaching in syllabus (KTSP 2006) for first grade of senior high school students, and represented the materials taught in the class. In line with the syllabus for the first grade of senior high school at the second semester, the students are required to be able to comprehend the short functional texts and it features in the form of descriptive text, for further information see the syllabus at Appendix 28

b) Construct validity

A test is said to have construct validity if it can be demonstrated that it measure. The word 'construct' refers to any underlying ability or trait which is hypothesized in a theory of language ability (Hughes, 1991:26). In this case, what means by reading is to understanding the ideas and information explicitly states in the passage (suparman: 2012). Based on that theory, some of the reading comprehension aspects that should be mastered are; identifying main idea, identifying details, making inferences, making inferences, and understanding vocabulary. Therefore, to make sure that the items of the test already good in the term of construct validity, the researcher specify them into table of specification as follows:

Table 1. Specification of Try Out

No	Reading Skills	Item Number	Percentage of Items
1.	Identifying a Main Idea	1, 7, 10, 12, 14, 17, 26, 29, 31, 38.	25%
2.	Finding detail information	8, 9, 13, 20, 21, 24, 35.	18%
3.	Finding Inferences	3, 2, 4, 18, 22, 23, 27, 28, 32, 33, 39.	27,5%
4.	Finding References	6, 11, 15, 19, 37.	12%
5.	Understanding Vocabulary	5, 16, 25, 30, 34, 36, 40.	17,5%
Total		40 items	100%

(Suparman,2012)

2. Reliability

Reliability is a necessary characteristic of good test: for it to be valid at all, a test must first be reliable as measuring instrument. Reliability refers to the extent to which a test produce consistent result when administered under similar condition (Hatch and Farhady, 1982:244). To find out the reliability of the test the

researcher used split-half method. Split-half method means the test was divided into halves; the first half and the second half. The first half contained the question number 1 to 20 and the second half covered question 21 to 40. To measure the coefficiency of theirs reliability, the researcher has used Pearson Product Moment, and here was the formula:

$$r_1 = \frac{\sum x}{\sqrt{(\sum x^2)}\,(\sum y^2)}$$

Where :

r1   = coefficient reliability between 1st half and 2nd half groups

x   = total number of 1st half group

y   = total number of 2nd half group

x2   = square of x

y2   = square of y

Then to know about the coefficiency of their correlation as whole the researcher used Spearman Brown's formula and here is the formula:

$$r_k = \frac{2r_1}{1 + r_1}$$

Where :

r_k = full test reliability

r_l = half test reliability

The criteria of reability are:

0.80-1.00   = very high

0.60-0.79   = high

0.40-0.59   = average

0.20-0.39   = low

0.00-0.19   = very low

<div align="right">(Hatch and Farhady,1982:246)</div>

3. Level of Difficulty

Level of difficulty is determine as the proportion of correct responses; that the higher the difficulty, the lower proportion correct and the higher the proportion incorrect (Henning, 1987:49). The test items that is too difficult or too easy for examines often show low reliability. Here, the students of try out class were divided into two groups that are upper and lower groups. The students' scores of try out test are listed from the highest score to the lowest scores and to see the level of difficulty, the researcher uses the following formula:

$$LD = \frac{U+L}{N}$$

Where:

LD   : level of difficulty

U   : the number of upper group who answer correctly

L   : the number of lower group who answer correctly

N   : the total number of students in upper and lower groups

The criteria are:

< 0.30   : difficult

0.30 – 0.70   : average

> 0.70   : easy

<div align="right">(Shohamy, 2985:79)</div>

4. Discrimination Power

Discrimination power refers to "the extent to which the item differentiates between high and low level students on that test. A good item according to his criterion is one in which good students did well and bad students failed." (Shohamy, 1985:81). In this research, the items which has discrimination power than or equal to 0.20 and the items which fall smaller than 0.20 (poor items) were absolutely dropped. To determine the discrimination power, the researcher uses the following formula:

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

Where:

DP : Discrimination Power

U : the number of upper group who answer correctly

L : the number of lower group who answer correctly

N : the total number of students in upper and lower groups

The criteria are:

1. If the value is positive discrimination a large number of more knowledge students then poor students get the item correct. If the value is zero, no discrimination.

2. If the value is negative, it means that more low students than high level students get the item correct.

3. If the value is zero, it means that there was no discrimination.

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

(Shohamy, 2985:82)

3.4.2. Pretest

This test was administered to the experimental class 1 and experimental class 2 at the beginning of the research, in order to determine the student's background knowledge of reading comprehension. The pretest was aimed at gaining the data of the students' initial reading ability. Besides, this test was administrated to ensure that the students of both classes have the same English background and proficiency before they receive the treatments. In addition, this pretest was carried out to find out the equivalent of experimental class 1 and experimental class 2. The total items of the test were 25 in form of multiple choices and this test was conducted within 60 minutes.

3.4.3. Posttest

The posttest was administered to the students after treatments. The aim of this test was to find out whether there is any significant difference between students' achievement before and after the treatments which was conducted to both of the experimental classes. This test consisted of 25 items of multiple choices and it was conducted within 60 minutes.

3.5. Scoring System

In order to get the final score of reading test, the scoring is based on the students' correct number in answering the test items and the formula is as follows:

$$S = \frac{R}{N} \ x \ 100\%$$

Where :

S : the score of the test

R : the total of the right answer

N : the total items

<p align="right">(Heaton, 1973:183)</p>

3.6. Research Procedure

The researcher used the following procedure in order to collect the data:

3.6.1. Determining the research problems

Based on the researcher's background of problem in the first chapter, the researcher tried to find out whether there was any significant different of student's achievement in reading comprehension achievement between the two experimental classes, and it was assumed that group work and individual work were compared to determine which one of those two teaching techniques was better for the student's reading comprehension achievement.

3.6.2. Determining the populations and then selecting the samples

The population of this research was the first grade of SMA Negeri 15 Bandar Lampung and the samples was chosen randomly. The researcher took two classes, the first one was Class X.4 as the experimental class 1 and the second one was Class X.5 as the experimental class 2.

3.6.3. Administering the Try out Test

The researcher prepared the try out materials and gives the try out test to another class which has the same characteristic as the samples that was chosen, that is Class X.3. The try out was conducted before the pretest is administered. This was expected to measure the validity and reliability of pretest and posttest. This test was consists of 40 items of multiple choice tests with five options. One of them as the correct answer and the others as the distracters. 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 should be revised or dropped for the pretest and the posttest.

3.6.4. Administering the Pretest

The pretest was aimed for gaining the data of the students' initial reading ability. Besides, this test administrated to ensure that the students of both classes have the same English background and proficiency before two classes received the treatments. In addition, this pretest was carried out to find out the equivalent of experimental class 1 and experimental class 2. The total items of the test were 25 in form of multiple choices and this test was conducted within 60 minutes.

3.6.5. Conducting the Treatment

There are three treatments were focused for experimental class 1 and experimental class 2. The teaching materials were arranged by preparing 3 lesson plans for each experimental class and add the material from network. The material that was used

is in form of descriptive text because it relevant to KTSP 2006 for students at X class in the second semester.

3.6.6. Implementing the Techniques (group work and individual work)

In this term, the researcher applied two techniques to both classes. The experimental class 1 was taught by using group work and the experimental class taught 2 was taught by using individual work.

3.6.7. Administering the Posttest

The researcher prepares the posttest to both experimental classes. At the end of the research, the researcher administrates posttest in order to find out the result of the treatments. The aim of this test was to find out whether or not there was any significant difference between students' achievement after the treatments which was conducted to both of the groups. This test consisted of 25 items of multiple choices for 60 minutes.

3.6.8. Analyzing the Data

In this step, both of the pretest and posttest results of the two classes were analyzed by using normality test, homogeneity test, random test and hypothesis test. Then the researcher compared mean pretest to posttest from each experimental class to see the increase of students' reading comprehension achievement before and after treatment.

3.6.9. Testing Hypothesis

The hypothesis testing taken from the comparison between the students' increased from pretest to posttest score in both classes that computed through SPSS 17.0 and also by the statistical analysis t-test with the significant level of 0.05. Therefore, if the result of SPSS was less than α and the result of statistical analysis t-test was more than α, it can be stated that the hypothesis was accepted.

3.7. Data Treatment

The aim of data treatment was to determine whether the students' reading comprehension achievement was increases or not. The data of the research was examined by using independent group T-test, because the independent variable has more than one group, that is; group work technique and individual work technique, which means that two different groups (experimental class 1 and experimental class 2) were compared. And the data is statically computed through the Statistical Package for Social Science (SPSS) version 17.0. In doing so, the researcher was analyzed the data statistically by administering the normality test, homogeneity test, random test, and hypothesis test.

3.7.1. Normality Test

Normality Test is used to measure whether the data of the test have normal distribution or not. It is because the students' score of pre-tests and post-test both group are analyze to gain the normality test. The researcher used SPSS (One-

Sample Kolmogrov-Smirnov Test). And the hypotheses for the normality test are as follows:

H_0 : The data is not distribute normally

H_1 : The data is distribute normally

In this research, "H_0 is accepted if significant value does not exceed level of significance at 0.05, meanwhile, H_0 is rejected if significant value exceeds level of significance at 0.05". The result of the normality testing can be seen in table 11 below:

Table 2. Normality Testing

	Kolmogrov-Sminorv Z	
	N	Sig. (2-tailed)
Pretest X.4	38	0.628
Posttest X.4	38	0.461
Pretest X.5	38	0.644
Posttest X.5	38	0.769

Table 2 inferred that the result of normality of the pretest in the experimental class 1 (X.4) showed that the value of two tailed significance was 0.628. In this case the hypothesis was accepted if Sign>α, 0.628>0.05. It means that the distribution of the data of the test was normal. Meanwhile, the result of the normality test of the posttest in the experimental class 1 showed that the value of two tail significance was 0.461. Since Sign> α, 0.461>0.05, it can be concluded that the data of the posttest in the experimental class 1 was normally distributed.

In other side, the similar results were found from the experimental class 2 (X.5). The value normality test of the pretest from this group showed in the number of 0.644. This hypothesis was accepted if Sign> α, since 0.644>0.05. Meanwhile, the

result of the normality test of the posttest in the experimental class 2 showed that the value of two tail significance was 0.769. Since Sign> α, 0.769>0.05, it could be stated the hypothesis was accepted both in the experimental class 1 and experimental class 2, which mean that the distribution data in both classes were normal. Furthermore, the result of computation of normality can be seen completely can be seen completely in Appendices 16 and 17.

3.7.2. Homogeneity Test

The homogeneity testing is intended to test whether the variance of the data in the experimental class 1 and experimental class 2 is equal or not. The data is statically will be computed through SPSS (Independent Sample Test). And the hypotheses for the homogeneity test are as follows:

H_0 : The data is not homogenous

H_1 : The data is homogenous

In which, "H_0 is accepted if significant value does not exceed level of significance at 0.05, meanwhile, H_0 is   rejected if significant value exceeds level of significance at 0.05". The result of homogeneity testing is as follows:

Table 3. Homogeneity Testing of Pretest

Variables	Sig. (2-tailed)	Conclusion
Experimental Class 1 Experimental Class 2	.867	Homogeneous

Table 3 shows that the data are homogenous since the significance is 0.867. As the significance is more than 0.05, it illustrates that the data of both classes are homogeneous. The complete result of computation can be seen in Appendix 20.

3.7.3. Random Test

In this research, the researcher is used SPSS (Number of Runs Test) to see whether the data in the experimental group 1 and experimental group is random or not. The hypotheses for random test are as follow:

H_0 : The data is not random

H_1 : the data is random

In which, "H_0 is accepted if significant value does not exceed level of significance at 0.05, meanwhile, H_0 is   rejected if significant value exceeds level of significance at 0.05". The result of random test is stated in the table below.

Table 4. The Random Test of Pretest in the Experimental Class 1 and Experimental Class 2

Variables	Test Value (a)	Sig. (2-tailed)	Conclusion
Experimental Class 1	64	.307	Random
Experimental Class 2	64	.255	Random

Table 4 illustrates that the random test from the pretest in the experimental class 1 showed the two tails significance was 0.307. Seeing the result, it could be concluded that the data was random since Sign> α (0.307>0.05). Meanwhile, The analysis of the random test of the pretest in the experimental class 2 showed the value of two tails was 0.255. Since Sign> α, (0.255>0.05). It indicates that the

significance of the data is greater than 0.05 and could be summed up that the data

of pretest both classes are taken from the population at random.

Table 5. The Random Test of Posttest in the Experimental Class 1 and Experimental Class 2

Variables	Test Value (a)	Sig. (2-tailed)	Conclusion
Experimental Class 1	80	.919	Random
Experimental Class 2	69	.250	Random

Table 5 indicates that the result of the random test from the posttest in the

experimental class 1 showed in the number of 0.919. The value was Sign$>\alpha$, in

which 0.919>0.05. It could be stated that the data was random. Meanwhile, the

result of random test from the posttest in the experimental class 2 showed the

value of two tails significance was 0.250. Since the value was higher than α

(0.250>0.05), the data from this group was also determined random.

In summary, the data of pretest and posttest from the experimental class 1 and

experimental class 2 showed the value of two tail significance were higher than

alpha (Sign$>\alpha$), which meant that the data from two groups were random. The

complete result of computation can be seen in Appendix 18 and 19.

3.8. Hypothesis Testing

The hypothesis is used to prove whether the hypothesis proposes in this research

is accepted or not. In this case, the data of the test is analyzed by comparing the

means score from both classes by using SPSS 17.0. Then data from the reading

test is analyzed by using the statistical analysis t-test with the significant level of 0.05. The formulation is as follow:

$$t = \frac{X_1 - X_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

With :

$$S^2 = \frac{(n_1 - 1)\,S_1^2 + (n_1 - 1)S_2^2}{n_1 + n_2 - 2}$$

Where :

X_1 = the arithmetical mean of the experimental class 1 (gain)

X_2 = the arithmetical mean of the experimental class 2 (gain)

S   = standard deviation

n_1 = the number of students in experimental class 1

n_2 = the number of students in experimental class 2

(Hatch and Farhady, 1982)

The hypotheses for the research question are:

H_0 : There is no significant difference between experimental class 1 and experimental class 2 on the student's reading comprehension achievement.

H_1 : There is significant difference between experimental class 1 and experimental class 2 on the Student's reading comprehension achievement.

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