# III. RESEARCH METHOD

This chapter discusses research design, population and sample, variables, data collecting technique, scoring system, research procedures, 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 using authentic material and another one as experimental group two that was given treatment using teacher-made material. The research design is as follows:

$$\frac{G_1 = T1X1T2}{G_2 = T1X2T2}$$

Notes:
$G_1$ : The experimental class I
$G_2$ : The experimental class II
$X_1$ : Treatment I (authentic material technique)

$X_2$ : Treatment II (teacher-made material technique)
T1 : Pre-test
T2 : Post-test
(Hatch and Farhady, 1982:20)

The research was conducted in six meetings with presentation as follows. The first meeting was for try-out test, the second meeting was for pre-test, the third, fourth, fifth meetings were for treatment and the sixth meeting is for post-test.

## 3.2  Population and Samples

### 3.2.1.  Population

The population of this research was the first grade students of SMA 5 Bandar Lampung of 20012/2013. There are six classes in the first year. The samples of this research were two classes of the first year students. Each class consists of 33 students and has the same level of ability.

### 3.2.2.  Sample

The sample classes were taken through lottery, because all the classes have the same opportunities to be chosen as the sample of this research and to make sure that the students' abilities were homogeneous or not by seeing the data of the teacher in the school. Therefore, the researcher took X.3 class as the experimental class 1, and the other one is X.4 class as the experimental class 2. In this case, the researcher asked the leader of the each class to take a small piece of paper in order to know the class will be as experimental class 1 or experimental class 2.

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

1.  Reading comprehension as dependent variable (Y)
2.  Authentic material as independent variable 1 ($X_1$)
3.  Teacher-made material as independent variable 2 ($X_2$)

## 3.4  Data Collecting Techniques

The instrument of this research was reading text. There were two kinds of test, pre-test and post-test. But, to prove whether the research instrument has good quality or not, it tried out first to measure its validity and reliability. Then, the data was obtained from the student's pre-test and post-test 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 will be elaborated in the following section:

## 1. Try Out

The try-out was administered to determine the quality of the test that was used in taking the data. In order to determine the quality of the test, the researcher analyzed four terms, that is: the validity, the reliability, the level of difficulty, and the discrimination power of the test.

## a. Validity

*Validity* is the extent to which a test does the job desired of it; the evidence may either empirical or logical (Lyman, 1971:196). A test can be said valid if it measures the object to be measured and suitable for the criteria (Hatch and Farhady, 1982: 251). There are four basic types of validity: content validity, criterion-related validity, face validity, and construct validity. In this research, the researcher used content validity and construct validity since the others are considered to be less needed.

## 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 of news item text.

**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 the Test**

| No | Reading Skills | Item Number | Percentage of Items |
|----|----------------|-------------|---------------------|
| 1. | Identifying a Main Idea | 1, 9, 15, 21, 26, 31, 36, 46 | 16% |
| 2. | Identifying details | 2, 3, 4, 10, 16, 17, 23, 24, 27, 28, 30, 32, 37, 41, 47, 48, 49 | 34% |
| 3. | Making/Confirming/Inferences | 11, 18, 22, 33, 38, 42 | 12% |
| 4. | Identifying References | 7, 19, 39, 43, 45 | 10% |
| 5. | Understanding Vocabulary | 5, 6, 8, 12, 13, 14, 20, 25, 29, 34, 35, 40, 44, 50 | 28% |
| Total | | 50 items | 100% |

(Suparman, 2012)

**b. 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 produces consistent result when administered under similar condition (Hatch and Farhady, 1982:244). To investigate the reliability of the reading test, the researcher use Pearson Product Moment which measured the correlation coefficient of the reliability between odd and even number (reliability of half test) in the following formula:

$$R_{xy} = \frac{N\sum xy - (\sum x)(\sum y)}{\sqrt{N\sum x^2 - (\sum X)^2}\sqrt{N\sum y^2 - (\sum y)^2}}$$

Where:

$r_{xy}$      : the correlation coefficient of reliability between odd and even

N      : the number of students who take part in the test

x      : the total numbers of odd number items

y      : the total numbers of even number items

(Hatch and Farhady, 1982:199)

After getting the reliability of half test, the researcher use Spearman Brown to determine the reliability of the whole test, as follows:

$$r_k = \frac{2r_{xy}}{1 + r_{xy}}$$

(Hatch and Farhady, 1982:247)

The criteria of reliability are:

0.90 – 1.00     : high

0.50 – 0.89     : moderate

0.0 – 0.49     : low

### c. Level of Difficulty

The difficulty level of an item shows how easy or difficult that particular item done by the participants, (Heaton,1975:182). Level of difficulty is generally expressed as the fraction (or percentage) of the students who answered the item correctly. It is calculated by the following formula:

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

Notes:
LD : the level of difficulty
R : the number of the students who answer correctly
N : the total of the students in the higher and lower group
(Heaton,1975:182)

The criteria of the difficulty level are
< 0.30 = difficult
0.30- 0.70 = average
> 0.70 = easy

### d. Discrimination Power

The discrimination power (D) 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. The discrimination power of an indicate item the extent, to which the item discriminates between the test taker from the less able. The formula of the discrimination power is:

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

Notes:
D : discrimination power

U  : the number of students from the upper who answer correctly
L  : the number of students from the lower who answer correctly
N  : the number of the students
(Shohamy, 1985:82)

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, 1985:82)

In this research, the instrument to be tried was administered to Class X.1 that consists of 33 students who has the same characteristic as the samples that was chosen. The material was based on the objective of teaching syllabus for the students at the first grade of senior high school. It means that, this test was completely proved the criteria of content validity. Then, to find the reliability of the test, the researcher used the data from try-out test (see Appendix 5). If the reliability tests reach 0.50, the researcher will consider that the test has been reliable. Meanwhile, in this research the reliability

was 0.84. So, it could be stated that the test had a very high reliability since 0.84 include to 0.80 – 100 range of very high reliability.

In this test, there were 50 items of multiple choices with five options (a,b,c,d or e), one of them as the correct answer and the rest were distracters. This was conducted in 100 minutes. After analyzing the data in level of difficulty, the writer got that 4 items were difficult, 18 items were average. Meanwhile, in the discrimination power, there were 33 good items and 17 poor items. In this case, there were 30 items that was dropped because it had easy or difficult in level of difficulty, and poor in the discrimination power. Therefore, there were 20 items that possible to use for pretest and posttest.

## 2. Pre Test

Pretest was administered before the treatment applied. It was done in order to know how far the competence of students in reading comprehension before the treatment. It was also needed to know whether both the two classes were equal or not in the terms of their reading comprehension achievement. The test was multiple choices.

**Table 2. Specification of Pre-test**

| No | Reading Sub-skills | Item Number | Percentage of Items |
|---|---|---|---|
| 1. | Identifying a Main Idea | 1, 4, 9, 16 | 20% |
| 2. | Identifying details | 2, 5, 10, 41, 17, 18, 19 | 35% |
| 3. | Making Inferences | 15, 12 | 10% |
| 4. | Identifying References | 6, 11 | 10% |
| 5. | Understanding Vocabulary | 3, 7, 8, 13, 20 | 25% |

**3. Post Test**

Posttest was administered after the treatment applied in order to find out whether is difference of students' reading comprehension between those who are taught using authentic material and those who are taught using teacher-made material. The posttest was done after 3 meetings of the treatments. The result of the posttest of two classes was compared in order to know whether teaching reading using authentic material and teacher-made material is effective or not.

**Table 3. Specification of Posttest**

| No | Reading Sub-skills | Item Number | Percentage of Items |
|----|-------------------|-------------|---------------------|
| 1. | Identifying a Main Idea | 1, 6, 11, 16 | 20% |
| 2. | Identifying details | 4, 7, 8, 10, 13, 17, 19 | 35% |
| 3. | Making Inferences | 20, 5 | 10% |
| 4. | Identifying References | 14, 3 | 10% |
| 5. | Understanding Vocabulary | 2, 9, 12, 15, 18 | 25% |

**3.5 Scoring System**

In scoring the result of students' test, the researcher will use Percentage Correct (Lyman, 1971:95). The percentage correct score is used in reporting the result of classroom achievement tests. The researcher will calculate the average of the pre-test and post-test by using this formula:

$$X_{\%c} = 100\frac{R}{T}$$

(Lyman, 1971: 95)

40

Where:

$X_{\%c}$ = percentage of correct score

R = number of right answers

T = total number of items on test.

## 3.6 Research Procedure

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

**1. Determining the research problem**

The first step of this research was determine the problem. The writer will determine what kind of problems appear in the classes.

**2. Determining the population and then selecting the sample**

The population of this research was the first grade of SMAN 5 Bandar Lampung and the samples were chosen randomly. The researcher took two classes, the first one was experimental class 1 (authentic material) and the second one was experimental class 2 (teacher-made material).

**3. Arranging the material that will be teach**

The researcher arranged the material that was taught to each class by prepared the lesson plan, searched and added the material from network. The research used types of text for treatments. The material based on KTSP 2006.

**4. Administering the try-out test**

The researcher prepared the try out materials and gave the try out test to another class which had the same characteristic as the samples that was chosen, it is X.1. The try-

out was conducted before the pre-test was administered. This was expected to measure the validity and reliability of pretest and posttest. This test was multiple choice tests and it was conducted in 90 minutes. There were 40 items of multiple choices with five options and one of them was as the correct answer, the test items were reduced or kept depends on its reliability and validity. The aim of try-out is to determine the quality of the test used as the instrument of the research, and to determine which item should be revised for the pre-test and the post-test.

## 5. Administering the pre-test

The pre-test was aimed for 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 two classes received the treatments. In addition, this pre-test was carried out to find out the equivalent of experimental class 1 and experimental class 2. The total items of the test were 20 in form of multiple choices and this test was conducted within 50 minutes.

## 6. Implementing the materials (authentic material and teacher-made material)

In this term, the researcher applied two materials to both classes. The experimental class 1 was taught by using authentic material and the experimental class 2 was taught by using teacher-made material.

## 7. Administering the post-test

The researcher prepared the post-test to both experimental classes. At the end of the research, the researcher administrated post-test in order to find out the result of the

treatments. The aim of this test is to finding whether or not there is a significant difference between students' achievement after the treatments which is conducted to both of the groups. This test consisted of 20 items of multiple choices for 50 minutes.

## 8. Analyzing the data

The data was analyzed by using normality test, homogeneity test, random test, and hypothesis test.

## 3.7  Data Treatment

The aim of data treatment was to determine whether the students' reading comprehension achievement is 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, those are; authentic material and teacher-made material, which means that two different groups (experimental class 1 and experimental class 2) were compared. And the data was statically computed through the Statistical Package for Social Science (SPSS) version 17.0. In doing so, the researcher analyzed the data statistically by administering the normality test, homogeneity test, random test, and hypothesis test.

## 1.    Normality Test

Normality Test was used to measure whether the data of the test have normal distribution or not. It caused the students' score of pre-test and post-test were analyzed to gain the normality test. The hypotheses for the normality are as follows:

$H_0$ : The data is distribute normally

$H_1$ : The data is not distribute normally

The data will be determined normal if it met the criterion, and the criterion for the hypothesis is:

The hypothesis will be accepted if sign$> \alpha$. In this case, the researcher will be used the level of significance of 0.05.

## 2. Homogeneity Test

The homogeneity testing is intended to test whether the variance of the data in 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, $H_1$ is rejected if significant value exceeds level of significance at 0.05".

## 3. Random Test

In this research, random test was used to see whether the data in experimental group 1 and experimental group was random or not. The hypothesis for random test is as follow:

$H_0$        : The data is random

$H_1$        : the data is not random

The data was determined random if it met the criterion, and the criterion for the hypothesis is:

$H_0$ will be accepted if sign> $\alpha$. In this case, the researcher will be used the level of significance of 0.05.

**4.  Hypothesis Test**

This test was to know the hypothesis which was proposed by the writer is accepted or not. The formulation of the hypotheses that (Independent group t-test) is as follows:

$$t = \frac{x_1 - x_2}{S \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$$

In which S = $\sqrt{S^2}$

$$S^2 = \frac{N_1 - 1 \; S_2^1 + N_1 - 1 \; S_2^2}{N_1 - 1 + N_1 - 1}$$

Notes  :

$x_1$        : the mean of the experimental class 1

$x_2$        : the mean of the experimental class 2

S        : standard deviation

$N_1$        : the number of the students in the experimental class 1

$N_2$        : the number of the students in the experimental class 2

The criteria are :

$H_0$      : t-ratio $\leq$ t $-$ table $=$ accepted

$H_1$      : t-ratio $\geq$ t $-$ table $=$ accepted

(Hatch and Farhady, 1983)

## 3.8 Hypothesis Testing

The hypothesis was used to prove whether the hypothesis proposes in this research is accepted or not. The researcher used SPSS (Independent group T-test). The hypothesis for the research questions are:

$H_0$  : There is no significant difference between authentic material and teacher-made material on the reading comprehension achievement.

$H_1$  : There is significant difference between authentic material and teacher-made material on the 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