

III. RESEARCH METHODS

A. Research Design

This research was a quantitative in which one-group pretest-posttest was applied (Hatch and Farady, 1982: 22). Pretest was conducted to find out the preliminary ability students and posttest was conducted to see whether *think-pair-share* technique is able to increase the students' reading achievement. The students were assigned randomly to one or the other group.

The design can be shown in the following way:

T1 X T2

Where :

T1 : pretest

T2 : pretest

X : treatment (using think-pair-share technique)

(Hatch and Farhady, 1982: 22)

B. Population and Sample

The research was conducted at SMAN 8 Bandar Lampung. The population of the research was the second year students of SMAN 8 Bandar Lampung. There were eight classes of the second year consisting of 32-36 students. Their age rate was 18 years old and come from different family background. The research used one sample class and the sample was taken randomly by using lottery since there was no special justification and priority class. So, the researcher took one class from the population as the sample. Class XI IPA 1 was assigned as experimental class.

C. Data Collecting Techniques

In collecting the data, some tests were administered. The type of the tests that were used as follows:

1. Pre Test

Pretest was administered to experimental group in the beginning of the treatment in order to find out the student's reading comprehension achievement

2. Post Test

Post Test was given to the experimental group at the end of the treatment in order to know the results of the student's reading comprehension of the whole treatment.

D. Research Procedures

1. Determining Research problem

The problem of this research was determined based on the pre-observation which was conducted by the researcher in SMAN 8 Bandar Lampung. The researcher asked 10 second year students 10 questions related to reading.

2. Selecting and Determining the Materials

The materials of this research were based on the School Based Curriculum (KTSP) 2006 for the second year students. The material were taken from and internet.

3. Determining the Research Instrument

The instrument of this research was objective reading test of multiple choices test. In measuring reading comprehension, multiple-choice selection is more valid than short-sentence answer (Henning, 1987: 48). Objective text was used for pretest and posttest. Each test consists of 30 items of multiple choices of comprehension questions and some reading texts. Each question had four alternative answers (A, B, C, and D), one correct answer and three distracters.

4. Administering Try-Out Test

The try-out test was conducted in 90 minutes and there were 40 reading multiple-choice questions with four options.

5. Administering Pre-Test

Pre-test is administered to identify the student's basic reading comprehension before the treatments. The test was administered in 60 minutes with 30 items of multiple choices reading test.

6. Conducting Treatment

The treatment that was applied to the experimental group in the classroom activity is *think-pair-share* technique. This treatment was conducted in three times in 2 x 45 minutes. The topics of the materials were about phenomena that should or should not be the case or be done.

7. Administering Post-Test

Post-test was given at the end of treatments in order to find out the development of the class. The test was administered in 60 minutes with 30 items of multiple-choice reading test.

8. Analyzing the Result of the Test

The result of pretest and posttest was analyzed using *T-test* to compare the data of two mean scores (Hatch and Farhady, 1982: 108).

9. Reporting the Results of the Research

The data that had been analyzed were written and reported in the script as the result of the research. It was presented after all the data had been collected and analyzed.

E. Scoring System

The researcher used Shohamy's formula in scoring the students' result of the test.

$$S = \frac{R}{N} \times 100$$

Where:

S = the score of the test

R = the right answer

N = the total of the items

(Shohamy, 1985:90)

F. Try Out

A research instrument will be said to have a good quality if it has good validity, reliability, level difficulty and discrimination power (Heaton, 1991: 5).

1. Validity

Validity is 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 can be considered to be valid if it can precisely measure the quality of the test.

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 will use content and construct validity since the other two were considered be less needed. Face validity only concerns with the appearance of the test. Criterion-related validity is concerned with measuring the success in the future, as in replacement test (Hatch and Farhady, 1982: 251). The two types will use in this research were:

a. Content Validity

Content validity is the extent to which the test measures a representative sample of the subject matter content. The focus of the content validity is adequacy of the sample and not simply on the appearance of the test (Hatch and Farhady, 1982: 251).

Content validity is intended to know whether the test items are good reflection of what will be covered. The test items are adapted from the materials that have been taught to the students should be constructed as to contain a representative sample

of the course. (Heaton, 1975: 160). In order to know whether the test have a good content validity, the items of the test will be discussed with the advisors to measure the degree of agreement. The composition of the test items was presented in table 1: table of specification below.

Table 1. Specification of the Validity Test.

No	Skills of Reading	Item Numbers	Percentage of Items
1	Determining main idea	4, 14, 24, 30, 36	12.5%
2	Finding specific information	2, 8, 9, 19, 20 21, 23, 38, 39	22.5%
3	Inference	10, 15, 18, 28, 32, 40	15 %
4	Reference	6, 16, 22, 33, 37	12.5 %
5	Understanding vocabulary	7, 17, 25, 29	10%
6	Determining features of text forms	1, 3, 5, 11, 12, 13, 26, 27, 31, 34, 35	27.5 %
Total		40	100%

(Suparman, 2012)

b. Construct Validity

Construct Validity is concerned with whether the test is actually in line with the theory of what it means to know the language (Shohamy, 1985: 74). Regarding the construct validity, it measures whether the construction had already in line with the objective of the learning (Hatch and Farhady, 1982: 251). Basically, the construct and content validity are overlap. It is a representative of the material from the subject. In line with Nuttal (1985) the relation validity of the instrument refers to construct validity in which the question represents five of sort reading skills, i.e. determining main idea, finding the detail information, reference, inference and vocabulary. Skills of reading in the test are a part of the construct validity and the item numbers are a part of the content validity.

1 Reliability

Reliability refers to whether the test is consistent in its score and gives us an indication of how accurate the test score are (Shohamy, 1985: 70).

A test is called reliable if the score gained by the examinees is constant whenever and by whomever the test is conducted. A test will not be a good parameter unless the test is suitable or constant. To measure the reliability of the test, the researcher used Spearman Brown formula. The formula is as follows:

$$R_k = \frac{2 \cdot r_l}{1 + r_l}$$

R_k = The reliability of the test

r_l = The reliability of half the tests

The criteria are:

0.00-0.19 Very low reliability

0.20-0.39 Low reliability

0.40-0.59 Average reliability

0.60-0.79 High reliability

0.80-1.00 Very high reliability

(Hatch and Farhady, 1982: 247)

2 Level of Difficulty

Level of difficulty relates to “how easy or difficult the item is from the point of view of the students who took the test. It is important since test items which are too easy (that all students get right) can tell us nothing about differences within the test population.” (Shohamy, 1985: 79).

Level of difficulty was calculated by using the following formula:

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

LD = level difficulty

R = number of students who answers it right

N = total number of students

The criteria are:

LD < 0.30 = difficult

LD = 0.31- 0.70 = average

LD > 0.71- 1.00 = easy

(Heaton, 1975: 182)

3 Discrimination Power of the Test

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 this criterion is one in which good students did well, and bad students failed.”

(Shohamy, 1985:81)

The formula is:

$$DP = \frac{\text{Upper} - \text{lower}}{\frac{1}{2(N)}}$$

DP = discrimination power

Upper = proportion of “high group” students getting the item correct

Lower = proportion of “low group” students getting the item correct

N = total number of students

The criteria are follows:

DP = 0.00-0.20 = poor

DP = 0.21-0.40 = satisfactory

DP = 0.41-0.70 = good

DP = 0.71-1.00 = excellent

(Heaton, 1975: 1985)

G. Data Analysis

The data were analyzed in order to determine whether there is any significant increase of students' reading comprehension when they are taught using *think-pair-share* technique. The researcher will analyze the data quantitatively after collecting the data. The researcher will examine the student's scores using the following step:

1. Scoring the pretest and posttest scores
2. Tabulating the results of the test and calculating the scores of the pretest and posttest
3. Drawing conclusion from the tabulated result of the pretest and posttest that will be statistically analyzed using *Independents T-Test* through SPSS version 20.0

According to Setiyadi (2006) 168-169), using t-test for the hypothesis testing has three underlying assumptions, there are:

1. The data is interval ratio
2. The data is taken from random sample in a population
3. The data is distributed normally

Therefore, the researcher used the following procedures to treat the data are as follow:

1. Normality Test

Normality test is used to measure whether the data in experimental classes are normally distributed or not (Setiyadi, 2006: 168-169). The students' scores of pretest and posttest were analyzed to gain normality test. The hypotheses for the normality test are as follow:

H₀ : the data is not distributed normally

H₁ : the data is distributed normally

In this research, H₁ is accepted if $p >$ and the researcher used level of significant 0.05.

H. Hypothesis Testing

The hypothesis is analyzed by using *Paired sample T-Test* in order to compare the different mean scores from different groups (Setiyadi, 2006: 169) with SPSS version 20.0. The hypotheses were as follow:

H₀ : There is no a significant increase of reading comprehension when they are taught by using *think-pair-share* technique.

($p > \alpha$)

H₁ : There is a significant increase of reading comprehension when they are taught by using *think-pair-share* technique.

($p < \alpha$).

The researcher used level of significance i.e. 0.05, and the probability error in the hypotheses was 5 %.

