

III. RESEARCH METHOD

This part discusses the research design, the population and sample, data collecting technique, instrument, try out, result of try out, research procedures, data analysis, and hypothesis testing.

3.1 Research Design

In conducting this research, the writer applied one group pretest- posttest, a pre-experimental design. In this design, the pretest and posttest were administered to investigate whether task-based approach can be used to increase students' reading comprehension achievement. Then, the means (average scores) of both pretest and posttest were compared to find out the progress before and after three times treatment.

There was one class as a subject of the research. This class had both pretest and posttest and three times treatment. The design could be illustrated as follows:

T1 X T2

Which:

X : Treatment, three times (using task-based approach)

T1 : Pretest (before treatment)

T2 : Posttest (after treatment)

(Setiyadi, 2006: 131)

The pretest was administered to find out students' reading comprehension achievement before the treatments. Subsequently, the students got treatments by using task-based approach. Eventually, a posttest was administered to find out the students reading comprehension after three times treatment.

3.2 Population and Sample

Population of this research was the first year students of SMK Negeri 2 Bandar Lampung. The consideration of choosing this school was based on the writer's assumption that SMKN 2 Bandar Lampung as one of good and favourite school in Bandar Lampung, it should have good students with appropriate language skill to make the research run well. The researcher considered if the students were beginners, it would be hard to apply it on them because in their level they just do what is told (Sevda, 20011). They can not give an output separately. More over, the students should be accustomed to experiment or do the lessons that they learn (task).

Then, the writer chose the first year students in the second semester of academic year 2012/ 2013. There were fifteen classes which each class consisted of 25 to 35 students. Then, the writer took one of those classes as the experimental class, it was class X TKR 2. The experimental class was selected using simple random sampling, taken by using lottery. 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.

3.3 Data Collecting Technique

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

1. Administering the Pretest

Pretest was administered in order to find out the students' reading comprehension achievement before treatment. The test was in form of multiple choices that consist of 25 items.

2. Administering the Posttest

Posttest was given after the treatment in order to find out whether there is any improvement of students' reading comprehension achievement. The test was multiple choices consisted of 25 items and all the items were the same as the pretest.

3.4 Research Instrument

The instrument of this research was objective reading text of multiple choices test. This is supported by Henning (1987: 48), who stated that to measure reading comprehension, requesting students to write short-sentence answers to written questions is less valid procedure than multiple-choice selection. Objective text was used for pretest and posttest. Each test consisted of 25 items of multiple choices of comprehension questions and reading texts. The questions had five alternative answers (A,B,C, D and E) for each item, one was the correct answer and the rest were the distracters.

3.5 Try Out

Whenever a test or other measuring device is used as part of the data collection process, there are four criteria of a good test should be met: validity, reliability, reliability, level of difficulty, and discrimination power. The purpose of doing this try out is to determine the quality of the test as instrument of the research, because a good instrument only would produce quite consistent result when administered under similar condition to the same participant and in different time (Hatch and Farhady, 1982).

1. Validity of the Instrument

A test can be said valid if the test measures the object to be measured and suitable with the criteria (Hatch and Farhady, 1982: 250). According to Hatch and Farhady (1982: 251), there are four basic 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. Criterion-related validity concerns 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 refers to the extent to which a test measures a representative sample the subject matter contents, the focus of the content validity is adequate of the sample and simply on the appearance of the test (Hatch and Farhady, 1982:251). In order to know whether the test is good reflection of what would be taught and

of the knowledge which the teacher wants the students to know, the researcher compared this test with table of specification. A table of specification was an instrument that helps the test constructor plans the test.

Table 3.1 Table specification of the instrument

No	Objectives	Item Numbers	Total Items	Percentage
1	Identifying the main idea	1, 8, 14, 24,30, 36	6	15 %
2	Understanding Vocabulary	7, 10, 19, 29, 35, 37	6	15 %
3	Identifying Specific information	2, 4, 11, 12, 13, 15, 16, 18, 20, 22, 23, 25, 31, 38	14	35 %
4	Finding Inference	5, 6, 9, 17 ,26, 28, 32, 33, 40	9	22,5%
5	Finding Reference	3, 21, 27, 34, 39	5	12,5 %
TOTAL			40	100%

b. Construct Validity

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

2. Reliability of the Instrument

Reliability refers to the extent to which the text is consistent in its score, and gives us an indication of how accurate the test score are (Hatch and Farhady, 1982: 244). To test the reliability of the instruments, the writer 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 have been taken in twice. The first half contained odd numbered and the second half contained even numbered. The correlation between those two parts encountered the reliability of half test by using Pearson Product Moment (Henning, 1987:60), which is formulated as follows:

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

Where,

- n = number of students
- r = coefficient reliability between first and second half
- x = total number of first half
- y = total number of second half
- x² = square of x
- y² = square of y
- Σx = total score of first half items
- Σy = 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 is 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

3. Level of Difficulty

Henning (1987: 49) states that the most important characteristic of an item to be accurately determined is its difficulty. To see the level of difficulty, the writer uses 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)

4. 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:

$$DP = \frac{\text{correctUpper} - \text{correctLower}}{1/2 N}$$

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 Results of Try out Test

Before conducting the pretest and posttest, a try out test was carried out. This test was administered in order to determine the quality of the test as instrument of the research. The try out test was administered in the class which did not belong to the experimental class, it was class X TKJ 2. It was conducted on February 15th, 2013 in 60 minutes. The writer prepared multiple choices test that consisted of 40 items. After analyzing the data, the writer got that 26 items were good and 14 items were poor and should be dropped, but the researcher took 25 items only. To find out the reliability of the test, the writer used statistical formula namely *Spearman Brown's Prophecy Formula*. If the reliability tests reach 0.50 the researcher will consider that it has been reliable. The result of the reliability found through this research was 0,80 (see Appendix 3). By referring to the criteria of the reliability proposed by Hatch and Farhady (1982:268), the test had moderate reliability in the range 0,50-0,89 it indicated that this instrument would produce quite consistent result when administered under similar condition to the same participant and in different time (Hatch and Farhady, 1982).

Based on the computation of difficulty level (see Appendix 4), the writer found that there were five items which were more than 0.70 which meant that the items

were easy and ten item was below 0.30 which meant difficult. Meanwhile the rest were between the ranges 0.30-0.70 or belonged to average.

Meanwhile from the computation of discrimination power (see Appendix 4) the writer got fourteen poor items (having less than 2.00 index), 25 items were satisfactory (having higher than 2.00 index) and one good item (has higher than 4.00 index). In general, it can be stated that all items tested had good discrimination power and positive value. In this research, the writer omitted 15 items and administered 25 items that were satisfactory to be used in pretest and posttest.

3.7 Research Procedures

There were some procedures that were applied for taking the data:

1. Selecting the instrument materials

The instrument materials (reading test) were taken from English text book, internet and other sources. The selecting process was considered the materials that would be taught to the students and the students' interest.

2. Determining research instrument

The materials in this research were based on the school based curriculum 2006. The materials were taken from the students' text book, and internet.

3. Determining the population and the sample

The researcher took one class as the sample or experimental class, it was class X TKR (Teknik Kendaraan Ringan) 2.

4. Administering the pretest

The test aim was to know the input or the state of students' ability in reading comprehension before they were given the treatment. The test was in form of multiple choice questions with five alternative answers for each question. One was the key answer and the rest were distracters.

5. Giving the treatment

There were three times treatments in this research. The Descriptive and procedure text were used as the media in teaching reading to the students by using task based approach.

6. Administering the posttest

The next step was administered the posttest. The type of the test was similar to the pretest. The urgency of giving the test was to find out whether there is any increase of the students' reading comprehension achievement.

7. Analyzing the result of both pretest and posttest

The next step of the research was analyzing the data. In this step, the researcher drew conclusion from the tabulated results of the pretest and posttest administered.

3.8 Data Treatment

In order to determine whether the data were good or not, the researcher analyzed the data by:

1. Scoring the pretest and posttest.

2. Tabulating the scores and calculating the mean of the pretest and posttest. The researcher used a very simple statistic formula to compute the average score or mean of the pretest and posttest. The formula was as follows:

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

Where:

\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. Administering the Normality Test

This test was used to measure whether the distribution of the data normally distributed or not. The data were tested by One-sample Kolmogorov-Smirnov Formula (SPSS 15).

The criteria of normal distribution is:

the hypothesis is accepted if the result of the normality test is higher than 0.05 (sign > α). In this case, the researcher used level of significance of 0.05.

4. Calculating from the tabulated scores of the pretest and posttest administered.

Statistically, researcher analyzed the data using repeated measure T-test to test whether the difference between pretest and posttest is significant or not.

It was used as the data comes from the same sample (Setiyadi, 2006: 170).

The researcher used SPSS version 17.0 to calculate the data.

5. Drawing conclusion from the data.

3.9 Hypotheses Testing

The hypotheses testing was used to prove which one of the two hypotheses in this research was accepted or not. The hypothesis was statistically analyzed using repeated measures T-test that is used to draw conclusion in significant level of 0.05 ($P < 0.005$).

The formulation was:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_D} \quad \text{With: } S_D = \sqrt{\frac{\sum D^2 - (\sum D)^2 / n}{n-1}}$$

Where:

\bar{x}_1 = Mean from pretest

\bar{x}_2 = Mean from posttest

S_D = Standard error of differences between means

n = Subjects on sample

(Hatch and Farhady, 1982:114)

The criteria are:

1. If the t-ratio is higher than t-table: H_1 is accepted
2. If the t-ratio is lower than t-table: H_0 is accepted

In order to get good and valid result, researcher used SPSS version 17.0 to analyze the data.