

III. RESEARCH METHODS

This chapter describes the method that is used in conducting the data of the research such as research design, population and sample of the research, research instruments, validity and reliability of the instrument, research procedure, data analysis, and hypothesis testing.

3.1 Research Design

This research used a quantitative because it was very useful for providing factors connected with second language development. Setiyadi (2006:5) cites that quantitative design aims to investigate a theory has been existed and the data in order to support or reject it.

In conducting the research, the researcher applied *One Group Pretest-Posttest design*. One group pretest-posttest design is a research design where one group of participants is pretested on the dependent variable and then posttested after the treatment condition has been administered. Pre test was given to the students in order to measure the students' competence before they were given the treatment and post test was given to measure how far the students' achievement after they

were given the treatment. In this design, it showed that the difference between the pretest and posttest scores. The research design could be represented as follow:

$$T1 \quad X \quad T2$$

Where:

T1 : pre-test

T2 : post-test

X : treatment (using Self Questioning Strategy)

(Hatch and Farhady (1982: 24)

3.2 Population and Sample of the Research

A population could be defined as the whole subjects of the research. Setiyadi (2006:38) states research population is all individuals which are being target in research while research sample is individual who give the data. The population in this research was the first grade of SMAN 5 Bandar Lampung in the academic year of 2012/2013. There were seven classes of the first grade (X1-X7). Their age rate is 17 years old and come from different family background. This research employed two classes; the first class as try out class was X-4 and the second class as the experimental class was X-1. The researcher applied the classes by using random sampling.

3.3 Variables

In order to assess the influence of the treatment in research, variables can be defined as dependent and independent variables. Hatch and Farhady (1982:15) state that the independent variable is the major variable that a resesarcher hopes to

investigate and the dependent variable was the variable that the researcher observed and measured to determine the improvement of the independent variable. The research consists of the following variables:

1. Students' reading comprehension achievement as dependent variable (Y).
2. Self-questioning as independent variable (X).

3.4 Research Instrument

The instrument of this research is reading text. the researcher administered a pre test, treatments and post test. Then, the data analyzed from the result of those three activities which could be clarified as follows:

1. Pretest

Pretest was conducted in order to find out the students' reading comprehension achievement before the treatments. This test was multiple choice in which the students were asked to choose one correct answer from the options a, b, c, or d. In this test students were given 30 items of reading and it was conducted within 60 minutes.

2. Posttest

After conducting the treatments, the post test was administered. It was done in order to know the students' achievement after applying the treatments. This test consists of 30 items of multiple choice for 60 minutes. It can be stated that comparing pre-test and post-test scores, a much higher post-test score should indicate that a student has learned certain topics.

If the scores are about the same, or if the post-test score is lower than the pre-test score, may the indications the measures that topics were not learned in the teaching learning.

3. Questionnaire

The questionnaire was conducted in order to investigate the students' response in reading comprehension of descriptive text using self-questioning. The questionnaire consists of 8 statements that related to the aspects of reading (main idea, supporting details, inference, reference and vocabulary). Each item provided four options; completely agree (4), agree (3), disagree (2), totally disagree (1). The questionnaire can be seen in Appendix 22.

3.5 Research Procedures

The procedures of this research were:

1. Determining Research problem

The problem of this research was determined based on the pre observation which was conducted in SMAN 5 Bandar Lampung. The research problem can be seen in the background in Chapter 1.

2. Determining the Research Instrument

The researcher checked the students' reading achievement by giving two reading tests to the students. The reading tests were pre test and post test. In measuring reading comprehension, multiple-choice selection is more valid than short-sentence answer (Henning, 1987: 48). Each test consists of 30 items

and each item had four alternative answers a, b, c, or d. one correct answer and three distracters. In addition, the questionnaire was carried out in order to know the students' response toward self questioning of descriptive text. The questionnaire consists of 8 statements or questions about reading comprehension and self questioning.

3. Selecting and Determining the Materials

The materials of this research were based on the School Based Curriculum (KTSP) 2006 for the first grade students. Besides, the materials were added and searched from network. Based on the discussion in Chapter 1, descriptive text was chosen as as the focus. The result of the try out were analyzed in order to know which the items are good to be used in pretest.

4. Administering the try out

This test was administered to measure the level of difficulty (LD) and discrimination power (DP) in order to find out the reliability and validity of the test. The items of the test were 40 and it was allocated within 80 minutes. The test was administered to measure the quality of the test as the instrument of the research.

5. Administering pre-test

The pre test was administered in order to find out students' reading comprehension achievement before treatments. In this test, the researcher asked students to do multiple choice tests that consist of 30 items of descriptive text in 60 minutes.

6. **Conducting treatments**

In this research, the treatment was conducted in three meeting which took 2 x 45 minutes. The lessons plan consisted of three different topics. The students has given the different assignments for each session, they are structured assignment and unstructured assignment. The text or the material were taken from students' English book and internet for first grades. The materials was about descriptive text

7. **Administering post-test**

The aim of this test was to measure the students' reading comprehension achievement after giving treatments. The test was conducted in 60 minutes with 30 items of multiple-choice reading test.

8. **Administering the questionnaire**

The questionnaire was given for the students in order to investigate the students' response in learning reading comprehension of descriptive text in applying self-questioning technique.

9. **Analyzing test data and testing hypothesis**

After scoring student's work, the data were analyzed by using *T-test* to compare the data of two mean scores (Hatch and Farhady, 1982: 108).

3.6 The Results of Try Out of the Instruments

According to the procedures in determining the instrument, this research started by trying out the reading comprehension test which was administered to remove the bad items based on their level of difficulty, and discrimination power of the

test and also the reliability of the test, validity of the test. The test was conducted at the first grade students of SMAN 5 Bandar Lampung on April 29th until May 18th 2013. There were seven classes of the first grade of students and this research employed two classes; the first class as try out class was X-4 and the second class as the experimental class was X-1.

3.7 The Try Out of Instrument

There were four criteria of good test, namely, validity, reliability, level of difficulty, and discrimination power.

3.7.1 Validity

Validity refers to the extent to which the test measures what was intended to measure. There are two basic types of validity; content validity and construct validity (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 kinds of validity that are:

1. face validity, concerns with the lay out of the test;
2. content validity, depends on a careful analysis of the language being stated;
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.

According to the types of validity above, the writer also used content and construct validity. Both of them were explained as follows:

a. Content Validity

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,1988). To get the content validity of reading comprehension, the researcher tried to arrange the materials based on the standard competence in syllabus for second grade of senior high school students that is students are able to construct meaning of functional text. In order to establish the content validity of a measuring instrument, the researcher identify the overall content to be represented. By using this method the teacher should obtain a group of items which is representative of the content of the trait or property to be measured.

The validity of the instruments is referred to the content and constructs validity in which the question represents five sort reading skills, i.e. determining main idea, finding the detail information, reference, inference and vocabulary (Nuttal, 1985). All test items which has good validity were used to collect the data for this research and the bad one should be revised.

The content of the try out was presented in table of specification below:

Table 1. Reading Specification (Aspects of Reading)

No.	Reading Skills	Items Number of Pre Test	Items Number of Post Test
1.	Identifying main idea	1,5,8,16,21,28,30	1,5,9,16,23,27
2.	Identifying details	3,6,9,13,17,19,26	2,6,14,17,24
3.	Making inference	11,15,22,24,29	4,10,12,25,29
4.	Understanding vocabulary	2,7,10,12,15,23	3,13,14,18,22,28
5.	Reference	4,10,14,22,18,27	8,11,19,26,30
Total		30 items	30 items

See on p.13

b. Construct Validity

Construct validity concerns whether the tests are true reflection in line with the theory of what it means to know the language (Shohamy, 1985:74). If a test has construct validity, it is capable of measuring certain specific characteristics in accordance with a theory of language behaviour and learning. This type of validity assumes the existence of certain learning theories or constructs underlying the acquisition of abilities and skills (Heaton, 1988:161).

3.7.2 Reliability

Reliability refers to the extent 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). In other words, how far it can measure the subject at separated time, but it shows the same result relatively (Setiyadi.2006:113). Reliability can be defined as the

extent to which a test produce consistent results when administered under similar condition (Hatch and Farhady.1982:244).

The test was determined by using Pearson Product Moment which measures the correlation coefficient of the reliability between odd and even number (reliability of half test) in the following formula:

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

where:

- r_{xy} : coefficient of reliability between odd and even numbers items
- x : odd number
- y : even number
- $\sum x^2$: total score of odd number items
- $\sum y^2$: total score of even number items
- $\sum xy$: total score of odd and even number

After getting the reliability of half test, the researcher used “Spearman Bown’s Prophecy formula” (Hatch and Farhady,1982; 247) to determine the reliability of the whole tests, as follows:

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

where:

- r_k : the reliability of the whole tests
- r_{xy} : the reliability of half test

(Hatch and Farhady, 1982:247)

The criteria of reliability as follows:

- 0.90 – 1.00 = high
- 0.50 – 0.89 = moderate
- 0.0 – 0.49 = low

After the data were analyzed, the result showed that the reliability of half test (r_{xy}) was 0.952 and reliability of the whole test (r_k) is 0.975 (see Appendix 5).

Based on the criteria of reliability of the test, it can be stated that the tests have a high reliability since the range of the high criteria for the reliability test is 0.8 – 1.00 (Hatch and Farhady, 1982: 246). It can be interpreted that the test can be used as an instrument to collect the data of the research. In other word, the test is reliable.

3.7.3 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).

Moreover, the difficulty level of an item shows how easy or difficult that particular item done by the participants Heaton (1975:182). The students were divided into two group that were upper and lower groups. The students' scores of try out was listed from the highest score to the lowest score. It was calculated by 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 as follows:

< 0.03 : difficult
 0.03 – 0.70 : average
 > 0.70 : easy

(Shohamy, 1985: 79)

Based on the result of try out related to the criteria, these test consisted of 5 difficult items (12, 15, 33, 35, 39), 26 good items (1, 2, 3, 6, 7, 9, 10, 11, 13, 14, 16, 17, 19, 20, 21, 22, 24, 25, 26, 27, 28, 30, 31, 32, 36, 40) and 9 easy items (4,5, 7, 8, 18, 23, 29,37,38). The easy and difficult items were not used to collect the data and should be revised or dropped from the test. Then, the average items were administered in pretest and post test. For discrimination power, it was found that there were 7 poor items (1, 2, 3, 15, 20, 32, 39), 16 satisfactory items (7, 10, 13, 14, 18, 19, 21, 23, 24, 26, 29, 30, 34, 36, 37, 38) and 17 good items (4, 5, 6, 8, 9, 11, 12, 16, 17, 22, 25, 27, 28, 31, 33, 35, 40) (Appendix 3).

3.7.4 Discrimination Power

Discrimination power refers to the extent to which the items were able to differentiate between high and low level students on that test. Discrimination power used to differentiate between the students who have high ability and those who have low ability.

A good item according to his criterion was the one in which good students did well and bad students failed (Shohamy, 1985:81).

The formula are :

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

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 of discrimination power are:

- 0.00 – 0.20 : Poor
 0.21 – 0.40 : Satisfactory
 0.41 – 0.70 : Good
 0.70 – 1.00 : Excellent
 - (negative) : Bad items (should be omitted)

From the computation of discrimination power of try out, it was found that there were 10 items (5, 7, 15, 17, 22, 23, 29, 31, 33 and 35) belong to bad items (has less than 2.00 index) Then, there were 28 items were satisfactory (1, 3, 4, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 21, 24, 25, 27, 28, 29, 30, 34, 36, 37, 38, 39 and 40) and the rest items (items number 2, 6, 20, 26 and 32) included in good discrimination power (see Appendix 3).

From the result of the test, 30 which met the criteria of a good test were administered in pre test and post test.

3.8 Scoring System

To get the score of the students' result of the test, this research employ Lyman's formula. The ideal highest score is 100. The score of pretest and posttest will be calculated by using the following formula:

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

Where:

X%c = percentage of correct score

R = number of right answer

T = total number of items on test

(Lyman, 1971: 95)

3.9 Data Analysis

In order to know the students' progress in comprehending the text and the students' score were computed by doing three activities:

1. Scoring the pretest and posttest
2. Tabulating the result of the test and calculating the mean of pretest and the posttest.

The mean was calculated by applying the following formula:

$$M = \frac{\sum x}{N}$$

Notes:

M = mean (average score)

$\sum x$ = the total students' score

N = total number of students

(Hatch and Farhady:1982)

3. Drawing conclusion from the tabulated results of the test given, that was by statistically analyzing the data using statistical computerization i.e Paired T-Test of Statistical Package for Social Science (SPSS) to test whether the increase of students' gain was significant or not, in which the significance was determined by $p < 0.05$. It was used as the data from one sample. (Hatch and Farhady, 1982: 117). In order to know whether the students got any progress, the formula was as follows:

$$I = \overline{X}_2 - \overline{X}_1$$

Notes:

I = the increase of students' reading comprehension achievement

\overline{X}_2 = the average score of post test

\overline{X}_1 = the average score of pre test

3.10 Hypothesis Test

The pre test and post test were compared in order to know the gain. The researcher used Repeated Measure T-Test towards the average score of pre test and post test. Moreover, the result of t-test was used to investigate the significance difference on students' reading comprehension achievement before and after giving self questioning and to prove whether the proposed hypothesis was accepted or rejected. In this case, the researcher used significant level of 0.05 in which that the probability of error in the hypothesis was only about 5%.

The hypothesis are drawn as follows:

H_0 : There is no significance difference on students' reading achievement by using self-questioning strategy of descriptive text.

H_1 : There is significance difference on students' reading achievement by using self-questioning strategy of descriptive text.

(Hatch and Farhady, 1982:111)

The criteria for accepting the hypothesis are as follows:

1. H_0 is accepted if the t-value is lower than T-ratio.
2. H_1 is accepted if the t-value is higher than T-ratio.