III. RESEARCH METHODS

This part discusses the design of this research and how to collect the data from the samples. The writer encloses the data collecting technique and the procedures of this research. The writer also gives the scoring system and how the data will be analyzed.

3.1 Research Design

To conduct this research, the researcher used *Control Group Pretest Posttest Design*. This design belonged to true experimental designs. True experimental designs have three basic characteristics: (1) a comparison group is present, (2) The sample are randomly selected and assigned to the groups, and (3) a pretest is administered to capture the initial differences between the groups (Hatch and Farhady 1982:22).

The researcher used this design because he wanted to give special treatment to the experimental class one using collaborative strategic reading in teaching reading comprehension. There were two classes of this experimental study; experimental class one which get treatment through collaborative strategic reading and another as a experimental class two which get treatment through self-questioning strategy. The pretest was administered first before the treatment. It was intended to measure the students' basic ability of both in order to ensure their entry point.

Experimental class two was needed for comparison purposes because it lets the writer interpret his findings more confidently. Both of them got the same materials.

Based on Hatch and Farhady (1982: 22), the researcher used the following design:

G1 : T1 X1 T2

G2 : T1 X2 T2

Notes:

G_1 = experimental Group 1

G_2 = experimental Group 2

T_1 = the pretest

T_2 = the posttest

X_1 = treatment by the researcher (Teaching through collaborative strategic reading)

X_2 = treatment by the teacher (Teaching through self-questioning strategy)

3.2 Population and Sample

3.2.1 Population

The population of the research was the first year students of SMAN 8 Bandar Lampung. The researcher chose the first year students in the second semester of academic year 2011/2012. There were eight classes of the first year students and each class consisted of 30 to 34 students. Their ages range from 15-16 years old.

3.2.2 Sample

Based on the population above, two classes were taken as the sample of this research, as experimental class one and experimental class two. The two sample classes of this research were selected using simple random sampling. Those classes were selected randomly by using lottery, since the first year students in SMAN 8 Bandar Lampung was not stratified class. 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. Next, to determine which class is as the experimental class one and experimental class two, the researcher used coin by flipping it.

3.3 Data Collecting Technique

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

1. Administering the Pre-test

   The pre-test was given before the treatment, in order to find out how far the competence of the students in reading comprehension or their input before the treatment and to find out the experimental class' reading comprehension achievement, the test was multiple choices that consist of 25 items. The materials tested was related to the curriculum used in the school and suitable with their level.

2. Administering the Post-test

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

pre-test. The materials tested, were related to the curriculum used in the school and suitable with their level. The post-test was done after three meetings of the treatments. The result of the post-test of the participant class was analyzed.

3.4 Research Procedures

There are some procedures that will be applied for taking the data:

1. Determining the population and the sample.

   The researcher took two classes to determine the experimental class one and experimental class two.

2. Administering try-out.

   The try-out test had been conducted before the pre-test was administered. This was expected to measure the validity and reliability of pretest and posttest, to ensure the data used by the researcher was valid and reliable to use as a research instruments. This test was multiple choice tests and was conducted in 80 minutes. There were 35 items of multiple choices with four options and one of them was as the correct answer, the test items could be reduced or kept depends on its reliability and validity. 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 for the pre-test and the post-test. This research used the result of the try-out test to measure the level of difficulty and discrimination power, to find out the validity and reliability of the test.

Criteria of Good Test

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.

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. Face validity only concerns with the layout of the test. 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). To know whether the test is good reflection of what will be taught and of the knowledge which the teacher wants the students to know, the researcher compares this test with table of specification. If the table represents the material that the researcher wants to test, then it is valid

from that point of view. A table of specification is an instrument that helps the test constructor plans the test.

Table 3.1 Table specification of try out

No	Objectives	Item Numbers	Total Items	Percentage
1	Identify the main idea	1, 9, 15, 19, 26,27	6	18 %
2	Vocabulary	6, 7, 17, 18, 24, 25, 33, 35	8	22 %
3	Specific information	4, 10, 12 ,13, 14, 21, 23, 28, 30, 32	10	28 %
4	Inference	2, 3, 11, 20, 22, 29, 31	7	20 %
5	Reference	5, 8, 16, 34	4	12 %
TOTAL			35	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 was true reflection of the theory in reading comprehension, the researcher examined whether the test questions actually reflected 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 is made as if the whole tests have been taken in twice. The first half contained passage 1 and 2 and the items

were number 1. until 18. The second half contained passage 3 and 5 involving question number 19. until 35. Moreover, by arranging the tests into first half and second half allowed the writer to measure the test reliability by having *split half method*.

To measure the coefficient of the reliability between the first and the second half, Pearson Product Moment was used, which was formulated as follows:

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

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
x^2 = square of x
y^2 = square of y
$\sum x$ = total score of first half items
$\sum y$ = total score of second half items

(Hatch and Farhady, 1982: 222)

Then to know the coefficient correlation of the whole items, Spearman Brown's Pharophecy Formula was used. The formula was 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

To see the index of difficulty, the writer used 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) was 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 was used:

$$DP = \frac{correctUpper - correctLower}{{}^{1}/_{2}\,N}$$

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)

5. 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 used by the researcher was multiple choice questions with four alternative answers for each question. One was the key answer and the last three were distracters.

6. Giving the treatment

   There were three times treatments in this research. The narrative text was used as the media in teaching reading to the students by using collaborative strategic reading in experimental class one and self-questioning strategy in experimental class two.

7. Administering the post test

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

8. Analyzing the result of both pretest and post test

   The next step of the research analyzed the data. Drawing conclusion from the tabulated results of the pre-test and post-test administered.

3.5 Scoring System

In scoring the result of students' test, the researcher used Percentage Correct (Lyman, 1971:95). The percentage correct score was 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)

Where:

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

R = number of right answers

T = total number of items on test

3.7 Data Analysis

The writer computed the students' score in order to find out the students' achievement in reading narrative text through Contextual Teaching and Learning using the following steps:

- Scoring the pre-test and post-test.
- Tabulating the results of the test and calculating the score of the pre-test and post-test.
- Drawing conclusion from the tabulated results of the pre-test and post-test administered, that was by statistically analyzing the data using statistical computerization i.e. *Independent Groups t-Test of Statistical Package for Social Science (SPSS) version 15.0 for windows* to test whether the increase of students' gain is significant or not, in which the significance was determined by $p < 0.05$. It is used as the data come from the two samples (Hatch and Farhady, 1982:111).

3.8 Treatment of the Data

In order to determine whether the data are good or not, the researcher will analyze the data by:

1. Scoring the pre-test and post-test.

2. Tabulating the result of the thesis and calculating the mean of the pretest and posttest. To compute the average score or mean of the pretest and posttest, the researcher will use a very simple statistic formula as follows:

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

Notes:

$\bar{X}$: mean (average score)
$\sum x$: total number of the student's score
N : total number of the students
(Hatch and Farhady, 1982:5)

3. Calculating from the tabulated results of the pretest and posttest administered, that was by statistically analyzing the data using ***t-test*** to test whether or not the difference between pretest and posttest is significant. It was used as the data comes from the same sample or known as paired data (Hatch and Farhady, 1982).

4. Administering Random Test

   This test was used to make sure whether the data is random or not (Hatch and Farhady is quoted by Setiyadi, 2006: 168-169). The researcher uses SPSS

version 17.0 to analyze the data. The hypotheses for the random test are as follow:

H_0 : the data is not random

H_1 : the data is random

In this research, the criteria for the hypotheses were H_1 is accepted if $p > \alpha$, and the researcher uses level of significance 0.05.

5. Administering the Normality Test

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

   The criteria of normal distribution are:

   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

6. Administering the Homogeneity Test

   This test was used to know whether the data of the posttest from the experimental class 1 and from the experimental class 2 are homogeneous or not. The data was tested by Independent Sample Test (SPSS 15). The criteria for the homogeneity of pre test were:

   H_0: There is no significant difference in the level of ability (equal)

   H_1: There is a significant difference in the level of ability (not equal)

   The criteria for the hypothesis is: H_1 is accepted if the result of Homogeneity test of pre test is higher than 0.05 (Sign > α).

3.9 Hypothesis Testing

After collecting the data, the writer recorded and analyzed them in order to find out whether there is an increasing in students' ability in reading comprehension of narrative text or not after the treatment. The writer used Independent Group T-test to know the level of significance of the treatment effect.

The formulation is:

$$t_{obs} = \frac{\overline{X}_e - \overline{X}_c}{S_{(\overline{X}_e - \overline{X}_c)}}$$

With:

$$S_{(\overline{X}_e - \overline{X}_c)} = \sqrt{\left(\frac{S_e}{\sqrt{n_1}}\right)^2 + \left(\frac{S_c}{\sqrt{n_2}}\right)^2}$$

$\overline{X}_e$: Mean from the difference pre-test and post-test of experimental class and control class

$\overline{X}_c$: Mean from the difference pre-test and post-test of experimental class and control class

S $(\overline{X}_e - \overline{X}_c)$: Standard error of differences between means

n : Subjects on sample

(Hatch and Farhady, 1982:111)

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