

### III.METHOD

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

#### 3.1. 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 was a research design where one group of participants was pretested on the dependent variable and then post-tested after the treatment condition had 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, there should be a 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 Collaborative Strategic Reading)

(Hatch and Farhady 1982: 24)

### **3.2. Population and Sample**

The population of this research was the second year students of SMPN 29 Bandar Lampung. There were ten classes of the second year students in 2014/2015 academic year. Each class consisted of about 24-25 students.

This research employed two classes; the first class was the class to implement the try out or called as try out class and the second class stood as the experimental class. The researcher applied the classes by using random sampling (lottery).

### **3.3. Variables**

In order to assess the influence of the treatment in research, variables could be defined as dependent and independent variables. Hatch and Farhady (1982:15) stated that the independent variable was the major variable that a researcher hoped 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. Collaborative strategic reading as independent variable (X).

### **3.4. Research Instrument**

The instrument of this research was pretest and posttest. The data was 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 treatment. 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, they were given 30 items of reading and it was conducted within 60 minutes.

#### **2. Posttest**

After conducting the treatment, the post test was administered. It was done in order to know the students' achievement after the treatment. This test consisted of 30 items of multiple choice and was done within 60 minutes. It could be stated when a much higher post-test score should indicate that a student had learned certain topics well.

### **3.5. Research Procedure**

Below are the procedures in administering the research:

#### **1. Determining Research problem**

The problem of this research was determined based on the researcher's experience in SMP N 1 Sumberjaya and as seen in chapter 1, it was intended to find out whether CSR or Collaborative Strategic Reading could be used to increased the students' reading comprehension or not.

## **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 selections were more valid than short-sentence answer (Henning, 1987: 48). Each test consisted of 30 items and each item had four alternative answers a, b, c, or d. There were one correct answer and three distracters.

## **3. Selecting and Determining the Materials**

The materials of this research were based on the current English curriculum for the second grade students. Those were also added and searched from internet. There were three lesson plans in the process of teaching reading which involved recount text inside.

The first lesson plan was used in the first meeting and purposed to teach the students about the strategies of applying Collaborating Strategic Reading (CSR) in the class. While the second lesson plan was used in the second meeting to make the students practice CSR in the real class discussion. The researcher acted as a teacher and divided the students into the groups included different role for each student. Last, the third lesson plan was done in the third meeting and purposed to strengthen the students' understanding and performance to do CSR.

## **4. Administering the try out**

The try out 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. Simply, it was administered to measure the quality of the test as the

instrument of the research. The items were 40 and allocated within 80 minutes.

### **5. Administering pre-test**

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

### **6. Conducting treatment**

In this research, the treatment was conducted in three meeting which takes 2 x 40 minutes. It was done by the lessons plan which consists of three different topics. The students were given the different assignments for each session and the materials which were about recount text were taken from the students' English book for the second grades and internet.

### **7. Administering post-test**

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

### **8. Testing hypothesis**

After scoring student's work, the data were analyzed by using *T-test* to compare the data of two mean scores.

## **3.6. Try Out Result**

The instrument of this research was objective reading test in form of pre-test and post-test. The researcher chose multiple choice items form since its marking was

rapid, simple, and most importantly reliable, not subjective or influence by the markers' judgment (Heaton, 1975:135).

Try out test purpose was for knowing the quality of the research instrument that was used in pre-test and post-test. In order to get a good test, the test item should fulfill some criterias such as: validity, reliability, level of difficulty, and discrimination power that will be discussed below.

### **3.6.1. Validity**

Validity referred to the extent to which the test measures and to what is intended to measure (Hatch and Farhady, 1982:250). Validity indicated how deep the instrument can measure the target of the research. There were four types of validity namely face validity, content validity, construct validity, and empirical validity or criterion-related validity.

To measure the test had a good validity, the researcher used content validity and construct validity. Face validity concerned with the layout of the test while the criterion-related validity was concerned with measuring the success in the future, as in replacement test (Hatch and Farhady, 1982:251). So, these two validities were considered to be less needed. An instrument was valid when it can provide the output accord with the researchers' analysis, its contents and constructs validity. The two types of validity that was used in this research as followed:

#### **a. Content Validity**

Content validity is intended to know whether the test items are good reflection of what will be covered or not. The test items which 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 the second grade of junior high school students. In order to establish the content validity of a measuring instrument, the researcher identified the overall content to be represented.

The validity of instruments were 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). The distribution of the items number was based on the current English curriculum, and the syllabus of second grade SMP students and represent of the materials that has been taught by the teacher. The content of the try out is presented in table of specification below:

**Table 3.1. Reading Specification (Aspects of Reading)**

No.	Reading Skills	Items Number	Percentage
1.	Identifying main idea	1,7,11,16,24,26,31,36	20%
2.	Identifying details	6,19,8,2,12,13,21,28,37,38	25%
3.	Making inference	15,9,20,5,25,29,35	17.5%
4.	Understanding vocabulary	4,9,14,17,22,30,34,39	20%
5.	Reference	3,10,18,23,27,33,40	17.5%
<b>Total</b>		<b>40 items</b>	

Based on the theory, the test was created by considering five reading skills and all of them had distributed well in all items so it could be concluded that the test had a good grade in content validity's context. The table above shows that identifying main idea was 20% (1,7,11,16,24,26,31,36) of the total items; identifying details

was 25% (6,19,8,2,12,13,21,28,37,38); making inference was 17.5% (15,9,20,5,25,29,35); understanding vocabulary was 20% (4,9,14,17,22,30,34,39); and reference was 17.5% (3,10,18,23,27,33,40) of the total items.

### **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).

According to Nuttal's theory in which the construct validity in this instrument represented by five sort reading skills: determining main idea, finding the detail information, reference, inference, and vocabulary; so the table 3.1 has explained that the test is in a good construct validity in which the five reading skills are distributed well percentage in the items of the test.

### **3.6.2. 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 produces consistent results when administered under similar condition (Hatch and Farhady.1982:244). The test was determined by using Pearson Product Moment which measured the



correlation coefficient of the reliability between odd and even number (reliability of half test) in the following formula:

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

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 0.975 and reliability of the whole test ( $r_k$ ) is 0.987 (see Appendix 6). 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 and in other word, the test is reliable.

### **3.6.3. Level of Difficulty**

Level of difficulty relates to how easy or difficult the item taken from the point of view of the students who take 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 groups that were upper and lower groups. The students' scores of try out were listed from the highest score to the lowest score. It is 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 tests consisted of 2 difficult items (23, 27), 32 good items (2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 25, 26, 28, 29, 30, 31, 32, 33, 34, 35, 37, 39) and 5 easy items (1, 6, 13, 24, 36, 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 (see appendix 4).

#### **3.6.4. Discrimination Power**

Discrimination power refers to the extent to which the items are able to differentiate between high and low level students on that test. It is used to differentiate between the students who have high ability and those who have low ability. A good item according to this criterion is the one in which good students do well and bad students fail (Shohamy, 1985:81). The discrimination power was calculated by this following formula:

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

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, For discrimination power, it was found that there were 9 poor items (1, 10, 15, 16, 23, 32, 34, 35, 38), 19 satisfactory items (2, 4, 6, 7, 8, 12, 14, 17, 20, 22, 24, 26, 27, 28, 29, 36, 37, 39, 40), 11 good items (3, 5, 9, 11, 13, 18, 19, 25, 30, 31, 33), and an excellent item (21) (Appendix 4). From the result of the test, 30 which met the criteria of a good test were administered in pre test and post test.

### **3.7. Scoring System**

To get the score of the students' result of the test, this research employed Lyman's formula. The score of pretest and posttest were 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)

### 3.8. 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 difference of students' gain was significant or not, in which the significance was determined by  $p < 0.05$ . It is used as the data from one sample. (Hatch and Farhady, 1982: 117). To know whether the students got any progress, the formula was as follow:

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

Notes:

I = the difference of students' reading comprehension achievement

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

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

### 3.9. Hypothesis Test

There are two hypotheses in this research. First, there is an increase between pre test and post test result after the students are being taught by using CSR. Second, identifying main idea is the most affected aspect of reading comprehension. To answer those hypotheses, the result of pre test and post test were compared so the researcher could identify further.

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 difference on students' reading comprehension achievement before and after the treatment and to prove whether the first proposed hypothesis was accepted or rejected. In this case, the researcher is significant level of 0.05 in which that the probability of error in the hypothesis is only about 5%. The hypotheses are drawn as follows:

$H_0$  : There is no increase on students' reading achievement after giving treatment by using collaborative strategic reading.

$H_1$  : There is increase on students' reading achievement after giving treatment by using collaborative strategic reading.

(Hatch and Farhady, 1982:111)

The criteria for accepting the hypothesis were as follows:

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

Furthermore, the researcher used the score of each reading skills in the pre test and the post test result. Those were compared to know whether the second hypothesis was accepted or rejected. It was accepted if identifying main idea was the most affected aspect but it was rejected if identifying main idea was not the most affected aspect.

A research was conducted by a method. After the background which explained why the research was needed to do and what the theories which were going to be the base, this chapter described how to conduct the data of the research. It was explained one by one; started from design, population and sample, research instruments, validity and reliability of the instrument, research procedure, data analysis, until hypothesis testing.