

III. RESEARCH METHODS

This chapter discusses about the research design, population and sample, data collecting technique, research procedures, instrument of the research, scoring system, criteria of good test, data analysis.

3.1 Research Design

In this research, the writer intends to find out whether there is any difference in students' reading comprehension score before and after being taught through reciprocal teaching technique. The writer applied one group pretest posttest design. The criteria of improving would be determined by pretest and posttest score. The pretest would be given to measure the initial ability. The design of the research is described as follows:

T1 X T2

T1 : pretest

T2 : posttest

X : treatment (using reciprocal teaching technique)

(Hatch and Farhady, 1982: 22)

3.2 Population and Sample

The population of this research was the second grade students of SMP Negeri 8 Metro in 2011/2012 academic year. There are five classes of the second year students and each class consists of 26 students. The writer took two classes, class VIII.A as experimental class which would be taught through reciprocal teaching technique and VIII.C as try out class. The class was chosen through lottery drawing.

3.3 Data Collecting Technique

The instrument of this research was reading test. There were three kinds of test that are try out test, pretest and posttest. They are as follow:

1. Try-out Test

This test aimed to know the validity and reliability of the test. The test was administered before the pretest. The total items are 50 and it is allocated within 80 minutes.

2. Pretest

The pretest was administered to measure the students' basic reading comprehension before they were given the treatments. The pretest consists of 30 items of multiple choices. The pretest was conducted within 60 minutes.

3. Posttest

The posttest is administered to the students after treatments. The aim is to know the result of students' reading comprehension after applying reciprocal teaching technique and in their reading. There are 30 items of multiple choices and it took 60 minutes.

3.4 Research Procedures

The procedures in administering the research are as follow:

1. Determining the Sample of the Research

The samples were chosen by simple random probability sampling, using lottery since the students' ability is parallel and all students have the same chance. The writer took one class of the second grade students of SMP Negeri 8 Metro.

2. Determining the Research Instruments

The instrument of this research is objective reading test. This is supported by Henning (1987:48), who states that to measure reading comprehension, requesting students to write short-sentence answers to written questions is less valid procedure than multiple-choice selection. Objective test is used for pretest and posttest. Each test consists of 30 items of multiple choices of comprehension question and some reading texts. The question has four alternative answers for each (A, B, C and D), one is correct answer and the rest are the disasters. The scoring system is that the load of each correct answer had five points. For both reading test, most of the materials are taken from students' English textbook and students' task sheet. In this research, the

writer used one type of reading text that is recount text and the topics are about Holiday, Experience, and Activity in the past.

3. Administering the Try-out Test

It was conducted to measure the reliability of pretest and posttest and to make sure whether the test is good or bad for students. The test was tried out to the students whose level is equal to the sample of the research. It was administered to find out the quality of the test before it is used, whether the items are good or not in validity, reliability, level of difficulty, and the discrimination power. This exam used reading text with 50 items of multiple choices in 80 minutes. The maximum score is 100.

4. Administering the Pretest

This test had aim to know the students' basic reading comprehension ability before they are given the treatments. It consisted of 30 items of multiple choices and was conducted within 60 minutes.

5. Conducting Treatments

The writer taught the students reading comprehension in reading text by using reciprocal teaching technique. The writer gave three times of treatments in different topics, which took 2 x 40 minutes in every meeting. The texts are taken from the students' English textbook.

6. Administering the Posttest

The writer conducted the posttest to measure the students' ability in reading comprehension after giving treatment. It consisted 30 items of multiple choices and reading texts which took 60 minutes.

7. Analyzing the Data

In his step, the pretest and posttest results in experimental class is analyzed by using *Repeated Measures T-Test* to measure the data (Hatch and Farhady. 1982: 114). And for analyzing all available data were selected into observation to investigate the process in teaching learning reading by using reciprocal teaching technique.

3.5 Instrument

The two reading tests were given to students to check reading comprehension ability. They are pretest and posttest. The purpose of the pretest is to know the students' basic reading comprehension ability before treatments. Then, posttest had purpose to know the students' improve after treatments.

3.6 Scoring System

The writer in scoring the students work, the writer used Arikunto's formula. The ideal highest score would be 100. The score of pretest and posttest are calculated by using the following formula:

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

Where:

S : the score of the test

R : the total of the right answer

N : the total items (Arikunto, 2005: 236)

3.7 Criteria of Good Test

A good test should meet four criteria: a good validity, reliability, level of difficulty and discrimination power.

3.7.1 Validity

Validity refers to 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: (1) face validity, (2) content validity, (3) construct validity, and (4) criterion-related validity.

To measure whether the test has a good validity, the writer uses content and construct validity since the other two are considered to be less needed. Face validity only concerns with the lay out of the test. Criterion-related validity is concerned with measuring the success in the future, as in replacement test (Hatch and Farhady, 1982: 251).

a. Content Validity

According to Hatch and Farhady (1982: 251), content validity is the extent to which the test measures a representative sample of the subject matter content.

Good test is the test which is appropriate with the material has been taught and the material is developed from the educational goal. The test instrument is designed to measure reading comprehension ability in line with: (1) educational goal stated on syllabus for second grade of junior high school students and (2) represent the materials taught in the class.

The validity of the instrument is referred to the content and construct validity in which the question represents five sort reading skills, i.e. determining main idea, finding the detail information, reference, inference, and vocabulary (Nuttal, 1982). In this research, the content of the test items is presented in the table of specification below.

Table 3.1 Specification of Data Collecting Instrument (Try-out Test)

No.	Reading Skills	Item Number	Percentage of Items
1.	Determining Main Idea	2, 11, 14, 23, 28, 37, 43, 50	16%
2.	Finding Specific Information	3, 6, 13, 18, 21, 24, 31, 40	16%
3.	Determining concept of text (generic structure / language features)	4, 8, 15, 20, 27, 33, 35, 41, 45	18 %
4.	Finding Reference	5, 7, 16, 25, 32, 36, 49, 46	16%
5.	Finding Inference	1, 10, 12, 19, 26, 30, 38, 42, 48	18%
6.	Vocabulary	9, 17, 22, 29, 34, 39, 44, 47	16%

b. Construct Validity

Construct validity concerns with whether the text is actually in line with the theory of what it means to know the language (Shohamy, 1985: 74).

3.7.2 Reliability

Reliability is a measure of accuracy, consistency, dependability or fairness of score resulting from administration of particular examination. Reliability of test can be defined as the extent to which a test produces consistent result when administrated under similar conditions (Hatch and Farhady, 1982:243). To estimate the reliability of the test this research used split-half technique. To measure the coefficient of the reliability between odd and even group, this research used the person product moment formula as follows:

$$r_1 = \frac{\sum XY}{\sqrt{(\sum X^2)(\sum Y^2)}}$$

Where:

r_1 : reliability between odd and even

x : the total numbers of odd number items

y : the total numbers of even number items

(Lado in Hughes, 1989: 32)

After getting the reliability of half test, the writer then used Spearman Brown's Prophecy formula (Hatch and Farhady, 1982: 256) to determine the reliability of the whole test as follows:

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

Where:

r_k : the reliability of the whole test

r_{xy} : the reliability of half test

The criteria of reliability are:

0.90 – 1.00 : High

0.50 – 0.89 : moderate

– 0.49 : Low

3.7.3 Level of Difficulty

In order to see the level of difficulty, the writer used the following formula:

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

Where:

LD : level of difficulty

R : the number of students who answer correctly

N : the total number of students following the test

The criteria are:

<0.30 : difficult

0.30 – 0.70 : average

>0.70 : easy

(Shohamy, 1985: 79)

3.7.4 Discrimination Power

The discrimination power is used to discriminate between weak and strong examinees in the ability being tested. The students of try out class are divided into two groups, upper and lower students. The upper students meant the students who answer the questions correctly are more than the lower students who answer the questions correctly (upper students' score > lower students' score). To determine the discrimination power, the writer used the following formula:

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

Where:

DP : discrimination power

U : the proportion of the upper group who answer correctly

L : the proportion of the lower group who answer correctly

N : the total number of the students

The criteria:

1. If the value is positive, it means that more high level students get correct answer than low students.
2. If the value is negative, it means that more low level students get correct answer than the high level students (it can be said that the test item is bad item, should be omitted).
3. If the value is zero, it means that there is no discrimination.

4. In general, the higher discrimination index is going to be the better. In classroom situation most items should be higher than 0.20 indexes.

(Shohamy, 1985:82)

3.8 Data Analysis

Analyzing data had aim to determine whether the students' reading comprehension achievement increase or not. The data of the research is examined by using *Repeated Measures T-Test* is used in this research. The data is statistically computed through the Statistical Package for Social Science (SPSS). It is used because the data come from the same sample (Hatch and Farhady, 1982:114).

3.9 Hypothesis Test

This test is used to test whether the hypothesis proposed by the writer accepted or not. The writer's hypothesis are:

H₀: There is no difference in students' reading comprehension score before and after being taught through reciprocal teaching technique.

H₁: There is any difference in students' reading comprehension score before and after being taught through reciprocal teaching technique.

The hypothesis was analyzed by using *Repeated Measures T-Test* which statistically computed through the Statistical Package for Social Science (SPSS) version 15.0 for windows in which the level of significance is $p < 0.05$.