

III. RESEARCH METHODS

A. Research Design

In this research, the researcher intends to find out whether there is any significant difference or reading comprehension achievement between students who are taught through reciprocal teaching technique and those taught through ordinary technique. The researcher selected two classes, one as the experimental group 1 and another as the experimental group 2. The researcher applied true experimental design. The criteria of increasing were determined by the differences of the posttest score of the two groups. The pretest was given to measure the initial ability and to make sure whether they had the same ability. Besides that, the subjects were chosen randomly to convince the validity. The design of the research was described as follows:

G1 (random) = T1 X1 T2

G2 (random) = T1 X2 T2

Where:

G1: experimental class 1 (reciprocal teaching technique)

G2: experimental class 2 (ordinary technique)

T1: pretest

T2: posttest

X1: treatment (using reciprocal teaching technique)

X2: treatment (using ordinary technique)

(Hatch and Farhady, 1982: 22)

B. Population and Sample

The population of this research was the eight grade students of SMP Negeri 6 Metro in 2010/2011 academic year in the third semester. The number was 150 students. The researcher took two classes as the sample of the research, class 8.1 as experimental class 1 and class 8.2 as experimental class 2. In addition, one class was assigned as try-out class that was 8.3. There were five classes of the eight grades and each class consists of 30 students. The classes were chosen randomly by lottery.

C. Data Collecting Technique

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

They were as follow:

1. Try Out Test

This test had aim 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 ability before they were given the treatments. The pretest consisted of 30 items of multiple choices. The pretest was conducted within 60 minutes.

3. Posttest

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

D. Variables

There are three variables in this research. They are:

1. Reading Comprehension as dependent variable (Y)
2. Reciprocal Teaching Technique as independent variable 1(X1)
3. Ordinary Technique as independent variable 2 (X2)

E. Research Procedures

The procedures in administering the research were as follow:

1. Determining the sample of the research

The samples were chosen by simple random probability sampling, using lottery since the students' ability were parallel and all students had the same chance. The

researcher took two classes of the second grade students of SMP Negeri 6 Metro. They were 8.1 as experimental class 1 and 8.2 experimental class 2. One class, 8.3 was also taken for try out test.

2. Determining the research instrument

The instrument of this research was objective reading test. This is supported by Henning (1975), who states that to measure reading comprehension, requesting students to write short-sentence answers to written questions was less valid procedure than multiple-choice selection (as cited in Henning, 1987:48). Objective test was used for pretest and posttest. Each test consisted of 30 items of multiple choices of comprehension question and some reading texts. The question had four alternative answers for each (A, B, C and D), one was correct answer and the rest were the disasters. The scoring system was that the load of each correct answer had five points. For both reading test, most of the materials were taken from students' English textbook and students' task sheet. In this research, the researcher used one type of reading text that was recount text and the theme was about Holiday.

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 was used, whether the items were 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 was 100 points; each correct answer had 2 points.

4. Administering the pretest

This test had aim to know the students' basic reading comprehension ability before they were given the treatments. It consisted of 30 items of multiple choices and was conducted within 60 minutes. At least, if a student could answer all items correctly, s/he got 100 points.

5. Conducting treatments

The researcher taught the students reading comprehension in reading text using reciprocal teaching technique for the experimental class 1 and ordinary technique for the experimental class 2. The researcher gave four times of treatments in four meetings, which took 2X40 minutes in every meeting. The texts were taken from the students' English textbook that were Let's Talk and English on Sky for Grade VIII of Junior High School and internet for second grades.

6. Administering the posttest

The researcher 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 (pretest and posttest)

In this step, the pretest and posttest results in experimental class 1 and experimental class 2 were analyzed by using independent group T-Test to compare the data of the two means scores (Hatch and Farhady. 1982: 108).

F. Instrument

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

G. Scoring System

The researcher in scoring the students work, the researcher used Arikunto's formula (1982:271). The ideal highest score would be 100. The score of pretest and posttest were 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)

H. Criteria of Good Test

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 had a good validity, the researcher used content and construct validity since the other two were 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 was 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, 1985).

In this research, the content of the test items is presented in the table of specification below.

Table1. 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.	Understanding Vocabulary	9, 17, 22, 29, 34, 39, 44, 47	16%
TOTAL			100%

b. Construct validity

Regarding the construct validity, it measures whether the construction had already referred to the theory, meaning that the test construction had already in line with the objective of the learning (Hatch & Farhady, 1982:251)

2. Reliability

Reliability refers to the extent to which a test produces consistent result when administered under similar condition (Hatch and Farhady, 1982: 244). In addition, Hatch and Farhady (1982: 246) also state that, there are three basic methods of estimating reliability: (1) test-retest, (2) parallel test, and (3) internal consistency methods.

The first, test-retest is administered in order to determine the stability of the test results. Reliability is obtained by administering a form test to the same students twice and computing the correlation between the two administrations. The second, parallel test is administered in order to determine the correlation between two alternate or parallel forms of tests, and called as a coefficient of equivalence. The tests has equivalent in length, difficulty, time limits, format and all other such aspects. The third, internal consistency method is administered in order to estimate reliability from a single administration of a single test. There are three basic methods for calculating reliability from an examination of internal consistency of the test: split-half method, Kuder-Richardson Formula 20, and Kuder-Richardson Formula 21 (Hatch and Farhady, 1982: 246).

Split-half method was used by the researcher to estimate the reliability of the test since this formula is simple to use. Besides that it avoids troublesome correlations and in addition to the number of item in the test, it involves only the test, mean and standard deviation, both of which are normally calculated (Heaton, 1991: 164). To use the split-half method, the researcher classified the test items into two similar parts, i.e. odd and even numbered. By splitting the test into two equal

parts, it is made as if the whole tests had been taken twice. The correlation between those two parts encountered the reliability of half test by using Pearson Product Moment (Henning, 1987: 60). After researcher had obtained the reliability of half test, the researcher then used Spearman Brown's Prophecy Formula (Hatch and Farhady, 1982: 246) to determine the reliability of the whole test. To measure the correlation coefficient of the reliability between odd and even number (reliability of half test), the researcher used Pearson Product Moment (Henning, 1987: 60) in the following formula:

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

Where:

r_{xy} : the correlation coefficient of reliability between odd and even

N : the number of students who take part in the test

x : the total numbers of odd number items

y : the total numbers of even number items

x^2 : the square of x

y^2 : the square of y

$\sum x$: the total score of odd number items

$\sum y$: the total score of even number items

(Henning, 1987: 60)

After getting the reliability of half test, the researcher then used Spearman Brown's Prophecy formula (Hatch and Farhady, 1982: 246) 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

(Hatch and Farhady, 1982: 247)

The criteria of reliability are:

0.90 – 1.00 : High

0.50 – 0.89 : moderate

0.0 – 0.49 : Low

3. Level of difficulty

In order to see the level of difficulty, the researcher 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)

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 were 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 researcher 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:

- a. If the value is positive, it means that more high level students get correct answer than low students.

- b. 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).
- c. If the value is zero, it means that there is no discrimination.

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)

I. Data Analysis

Analyzing data had aim to determine whether the students' reading comprehension achievement increase or not. The data of the research was examined by using independent-group T-Test. Independent t-test is used in this research. Two means of two different groups (experimental group 1 and experimental group 2) are compared. The data was statistically computed through the Statistical Package for Social Science (SPSS) version 17.0.

J. Data Treatment

According to Setiyadi (2006: 168-169), using t-test for the hypothesis testing has three underlying assumptions, namely:

1. The data is interval ratio.
2. The data is taken from random sample in a population.
3. The data is distributed normally.

Therefore, the researcher used the following procedures to treat the data treatment:

1. Random test

Random test or run test is used to make sure whether the data is random or not (Setiyadi, 2006: 168-169). The researcher used 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 used level of significance 0.05.

2. Normality test

Normality test is used to measure whether the data in experimental group 1 and experimental group 2 are normally distributed or not (Setiyadi, 2006: 168: 169). The students' scores of pretests and posttests both group are analyzed to gain the normality test. The hypotheses for the normality test are as follow:

H_0 : the data is not distributed normally

H_1 : the data is distributed normally

In this research, H_1 is accepted if $p > \alpha$, and the researcher used level of significance 0.05.

3. Homogeneity Test

This kind of test is used to know the data in experimental class 1 and 2 are homogenous or not. In this research, the researcher used Independent Samples Test in SPSS 17.0 to know the homogeneity of the test. The hypothesis for homogeneity of variance test was:

H0 : there is no significant difference (equal)

H1 : there is a significant difference (not equal)

In this case, H0 was accepted if $p > \alpha$ (p = the significant score of students, α = the significant level). Here, the researcher used the significant level (α) 0.05.

4. Hypothesis Test

The hypothesis of this research was as follows

“There is a there is a significant difference of students’ reading comprehension achievement between students who are taught through reciprocal teaching technique and those who are taught through ordinary (lecture) technique.”

The hypothesis was statistically analyzed using Independent Group T-Test. It was used to draw conclusionat the level of 0.05 ($p < 0.05$).

