III. RESEARCH METHOD

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

3.1 Research Design

The researcher conducted quantitative research with pretest posttest control group design in her research that belongs to the true experimental design. The researcher used the design because she wanted to compare students' reading comprehension achievement between those taught trough Reciprocal Teaching Technique and those taught through Contextual Teaching and Learning and which one of them is more effective in increasing it. According to Hatch and Farhady (1982: 22) define the basic the characteristics of true experimental designs into three:

a. a control group is present,

b. the students are randomly select and assigned to the groups, and

c. a pretest is administered to capture the initial differences between the groups.

Those three basic characteristics allowed the researcher to avoid almost all the problems associated with internal and external validity.

The researcher used two classes – that was an experimental class and a control class. The researcher gave three treatments to the experimental class and control class. Each treatment was teaching reading comprehension (related to aspects – that is: main ideas, details, inferences, references, and vocabularies) through RTT. The control class received the treatment but they studied through CTL as regularly. Both classes received the some pretest and posttest. Pretest was administered to see the students' basic ability within both groups in order to ensure that their ability was equal before treatments. Then, after giving the treatments the researcher administered the posttest to the both groups. The posttest was administered in order to prove that RTT can increase students' reading comprehension between both classes.

According to Hatch and Farhady (1982: 22), the research design is presented as follows:

G1   : T1   X1   T2

G2   : T1   X2   T2

Where:

G1   : Experimental Class

G2   : Control Class

T1   : Pre Test

T2   : Post Test

X1   : Treatment 1 (Teaching reading through RTT)

X2   : Treatment 2 (Teaching reading through CTL)

3.2 Population and Sample

3.2.1 Population

The population of the research was the second grades students of SMPN 5 Metro. There were 21 classes in this school and each grade has 7 classes. The total number of the population was 630 students, consisting of 304 female students and 326 male students. In this research, the researcher chose the second grade in the first semester of academic year 2010/2011 were investigated. There were seven classes of the second grade students, they are 8.A, 8.B, 8.C, 8.D, 8.E, 8.F, 8.G and each class consists of 30 students. Their ages range from 13-14 years old.

3.2.2 Sample

The researcher took two classes as the sample of the research, class 8.F as experimental class that was given the treatment by the researcher (teaching reading of recount text using Reciprocal Teaching Technique) and class 8.G as control class that was given the treatment by the researcher (teaching reading of recount text using CTL). In addition, class 8.B was assigned as try-out class. The classes were chosen randomly by lottery.

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 had aim to know the validity and reliability of the test. The test was administered before the pretest. The total items were 50 and it was allocated within 80 minutes.

2. Pretest

The pretest was administered to measure the students' basic reading comprehension ability before they are 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 CTL in their reading. There were 30 items of multiple choices and it takes 60 minutes.

3.4 Variables

There were three variables in this research. They were:

1. Reciprocal Teaching Technique as independent variable 1(X1)
2. Contextual Teaching and Learning (CTL) as independent variable 2 (X2)
3. Reading Comprehension as dependent variable (Y)

3.5 The Criteria of Good Test

The testes were said to have good quality if it has a good validity, reliability, and level of difficulty and discrimination power.

3.5.1 Validity of the Test

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. In this research, the writer use content validity and construct 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 layout 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 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).

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

Table1. Specification of Data Collecting Instrument (Try-Out Test)

No.	Reading Skills	Item Number	% 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 and Farhady, 1982: 251).

3.5.2 Reliability of the Test

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 was made as if the whole tests had been taken twice. The correlation between those two parts encounters the reliability of half test by using Pearson Product Moment (Henning, 1987: 60). After researcher has obtained the reliability of half test, the researcher then uses Spearmen 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:

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

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)

Then the researcher used Spearman Browns Prophecy formula (Hatch and

Farhady, 1982; 246) to determine the reliability of test as follow:

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

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

3.5.3 Level of Difficulty

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 the students who answer correctly

N   : the total number of the students

The criteria are:

< 0.30   = difficult

0.030 – 0.70   = average

> 0.70   = easy

<p style="text-align: right;">(Shohamy, 1985:79)</p>

3.5.4 Discrimination Power

The discrimination power is used to discriminate between weak and strong examines 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 question correctly were more than the lower student who answer the question correctly (upper students' > lower students' score).

To see the discrimination power, the writer used the following formula:

$$DP = \frac{\text{The proportion of upper SS} - \text{the proportion of lower SS}}{\text{½ total number students}}$$

<p style="text-align: right;">(Shohamy, 1985:81)</p>

The criteria are:

1. If the value is positive, it has discrimination because a larger number or more knowledgeable students than poor students get the item correct. If the value is zero, it means no discrimination.

2. If the value is negative, it has negative discrimination because more low – level students than high level students get the item correct.

3. In general, the higher discrimination index, the better, in classroom situation most items should be higher than 0.20 indexes.

3.6 Scoring System

The researcher in scoring the students work, the researcher uses Arikunto's
formula. The ideal highest score will 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 Research Procedure

The procedures in administering the research were as follow:

1. Determining the sample of the research

The sample was 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 5 Metro.
They were 8.F as experimental class and 8.G as control class. One class, 8.B 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 is 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 is recount text.

3. Administering the try-out test

It was conducted to measure the reliability of pretest and posttest and to make sure whether the test was good or bad for students. The test was tried out to the students whose level was 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 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 and Contextual Teaching and learning (CTL) for the control class. The researcher gave three times of treatments in three meetings, which took 2X40 minutes in every meeting. The texts were taken from the students' English textbook that are 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 his step, the pretest and posttest results in experimental and control class were analyzed by using independent group T-Test to compare the data of the two means scores (Hatch and Farhady. 1982: 108).

3.8 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.

3.9 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 was used in this research. Two means of two different groups (experimental group and control group) were compared. The data was statistically computed through the Statistical Package for Social Science (SPSS).

3.10 Data Treatment

According to Hatch and Farhady is quoted by 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:

3.10.1 Normality Test

Normality test is used to measure whether the data in experimental group and control group are normally distributed or not (Hatch and Farhady is quoted by 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 uses level of significance 0.05.

Based on the Appendix 15, the normality value of pretest and posttest in experimental class were 0.78>0.05, 0.51>0.05. The normality value of pretest and posttest in control class were 0.23>0.05, 0.21>0.05. It meant that H_1 was accepted, and H_0 was rejected. In short, the distribution of the data in this research was normal.

3.10.2 Homogeneity Test

This kind of test is used to know the data in experimental class and control class 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:

H_0 : there is no significant difference (equal)

H_1 : there is a significant difference (not equal)

In this case, H_0 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.

Look at the Table 4 below to know the result of results of homogeneity test and the comparison of the pretest scores in both classes.

Table 2. The Analysis of the Homogeneity Test

Independent Samples Test				Pre-test	
				Equal variances assumed	Equal variances not assumed
Levene's Test for Equality of Variances		F		.115	
		Sig.		.736	
t-test for Equality of Means		T		.297	.297
		Df		58	57.979
		Sig. (2-tailed)		.767	.767
		Mean Difference		.63333	.63333
		Std. Error Difference		2.13191	2.13191
	95% Confidence Interval of the Difference	Lower		-3.63414	-3.63417
		Upper		4.90080	4.90084

The result showed that the value of two-tailed significance (p) was 0.77. It meant that $p > \alpha$ (0.77 > 0.05). It can be said that H_0 was accepted and H_1 was rejected. In other word, the both classes had the same level of reading comprehension ability before treatments.

3.10.3   Hypothesis Test

The hypothesis analyzed at the significant level of 0.05 in which the hypothesis would approve if sign $< \alpha$. It meant that the probability of error in the hypothesis was only about 5%.

The hypotheses were:

Ho: There is no significant difference of students' reading comprehension

achievement between those taught through RTT and those taught through

CTL.

H_1:   There is a significant difference of students' reading comprehension

achievement between those taught through RTT and those taught through

CTL.

(Hatch and Farhady, 1982: 111)

Look at the Table 7 below to know the result of results of hypothesis test and the

comparison of the posttest scores in both classes.

Table 3.The Analysis of the Hypothesis Test

Independent Samples Test				
			post-test	
			Equal variances assumed	Equal variances not assumed
Levene's Test for Equality of Variances		F	1.075	
		Sig.	.304	
t-test for Equality of Means		T	9.168	9.168
		Df	58	57.112
		Sig. (2-tailed)	.000	.000
		Mean Difference	12.36667	12.36667
		Std. Error Difference	1.34889	1.34889
	95% Confidence Interval of the Difference	Lower	9.66657	9.66568
		Upper	15.06676	15.06766

Based on the table above, it can be assumed that the significant score of students

was 0.000. It meant that $p < \alpha$ (0.000<0.05). It can be determined that H_0 was

rejected and H_1 was accepted. In other word, the experimental class and control

class had a significant difference of reading comprehension ability after

treatments.