

## CHAPTER III

### RESEARCH METHODOLOGY

#### A. Research Method and Design

##### 1. Research Method

The research method that used in this research was quasi-experimental method. A quasi-experimental design involves the use of an intervention, but not random assignment of participants to groups because the experimenter cannot artificially create group for experiment (Creswell, 2012). This method was appropriate with the objective of this study which was investigating the role of Reciprocal Teaching foster students' comprehension and questioning ability in learning ecosystem.

##### 2. Research Design

The design that used in this research was pre-test and post-test design (Creswell, 2012). The research assigned intact group the experimental and control treatment, administered a pre-test to both group, conducted experimental treatment activities, with the experimental group only, and then administered a post-test to assess the differences between the two groups.

**Table 3.1 Pre-test and Post-test Design**

Select Control Group	Pretest	No treatment	Posttest
Select Experimental	Group Pretest	Experimental Treatment	Posttest

*(Creswell, 2012)*

#### B. Population and Sample

The location held in this research was in Private Junior High School "X" Jakarta. The population in this research was all 7<sup>th</sup> grade students at Junior High School "X" Jakarta. The samples were 7<sup>th</sup> grade students from two different

classes in Junior high School “X” Jakarta. The sampling technique was cluster random sampling. Frankel and Wallen (2007) stated that cluster random sampling is obtained by using group as the sampling unit rather than individuals.

### **C. Operational Definition**

In order to avoid the misconception in this research there are some operational definitions are explained in this research. Those terminologies are explained as follow.

a. Reciprocal Teaching is a reading instructional model that involves teachers and learners engaging in dialogues vis-à-vis a text. This strategy involves four components: summarizing, questioning, clarifying and predicting. This strategy was conducted by lesson plan that represent the reciprocal teaching indicator and measure by using observation sheet.

b. Students’ comprehension

Students’ comprehension that is the competence of students that cover the cognitive level of remembering (C1), understanding (C2), applying (C3) and analyzing (C4). This was measured by using multiple. This competence will measure by using multiple choice questions (C1-C4) through pretest and posttest.

c. Questioning ability.

Questioning ability is an integral part of meaningful learning and scientific inquiry. This research is according to Question Category System of Science (QCSS) which is consists of two levels of thinking expected by questions which are closed questions and open questions (level I) and in level II of QCSS the question divided into four types, which are cognitive-memory operation and convergent thinking operation (closed question), divergent thinking operations and evaluative thinking operation (open question). This was analyzed by QCSS. This will analyze their questioning ability through table of QCSS.

## **D. Assumption**

The assumption as the foundation of this study as follow:

- a. Reciprocal Teaching showed that individual students' reading comprehension was enhanced by remedial instruction incorporating different functionalities of the Reciprocal Teaching system.
- b. Reading helps in almost every area of smarts. Those that read have higher GPA's or Score, higher intelligence, and general knowledge than those that do not.
- c. Readers improve their general knowledge, and more importantly are able to spot patterns quicker. If can spot patterns quicker, analytical skills receive a boost, so makes readers more critic it can be see when readers propose the question.

## **E. Hypothesis**

Hypothesis that is tested in this study are as follow:

H<sub>0</sub>: There are no differences in student comprehension in learning Ecosystem through Reciprocal Teaching.

H<sub>1</sub>: There are differences in student comprehension in learning Ecosystem through Reciprocal Teaching.

## **F. Research Instrument**

There are three instruments that were used in this research namely objective test, observation sheet and Question Category System of Science (QCSS) rubric. Those instruments are described below.

### **a. Objective Test (cognitive)**

This test used to evaluate concept comprehension and questioning ability of students in learning Ecosystem. Objective test consist of four options multiple choices and some essay question. The test formulated according to Bloom (in

Clark, 2013) cognitive value of remembering (C1), understanding (C2), applying (C3) and analyzing (C4).

## 1. Validity

Cronbach (1971) in Crocker and Algina (2008) described validation as the process by which a test developer or test user collects evidence to support the types of inferences that are to be drawn from test scores.

Validity is defined as the extent to which an instrument is measured what it claimed to measure (Ary et al, 2010). Validity can be defined as the agreement between a test score or measure and the quality it is believed to measure (Kaplan, 2012). Hence, by having the validity test it can measure data is resulted from the test is valid with some variable and interpreted. In this study to determine the instrument is valid or not (validity) is done by using software ANATES.

The maximum validity coefficient ( $r_{12max}$ ) between two variables is equal to the square root of the product of their reliability.

Where:

$$r_{12max} = \sqrt{r_{11}r_{22}}$$

$r_{11}$  and  $r_{22}$  are the reliabilities for the two variables

(Kaplan and Saccuzzo, 2012)

## 2. Reliability

Measurement always involves some errors. However, physiological and educational specialists devoted to measure the errors and its effect. They provide evidence of the advanced scientific status of the field (Kaplan and Saccuzzo, 2012).

Ary et al. (2010) stated that Reliability is concerned with the effect of error on the consistency of scores. There are two kinds of errors: random errors of measurement and systematic errors of measurement. Reliability

show the tendency toward consistency found in repeated measurement of the same phenomena is referred (Carmines and Zeller, 1979).

The reliability coefficient is the ratio of the variance of the true scores on a test to the variance of the observed scores:

Where

$$r = \frac{\sigma^2 T}{\sigma^2 x}$$

$r$  = the theoretical reliability of the test

$\sigma^2 T$  = the variance of the true score

$\sigma^2 x$  = the variance of the observed score

(Kaplan and Saccuzzo, 2012)

**Table 3.2 Factors that Affecting the Reliability of a Test**

Factor	Potential effect
Length of the test	The longer the test, the greater the reliability.
Heterogeneity of Group	The more heterogeneous the group, the greater the reliability.
Ability level of group	A test that is too easy or too difficult for a group results in lower reliability.
Techniques used to estimate reliability	Test-retest and split-half give higher estimates. Equivalent forms give lower estimates.
Nature of the variable	Tests of variables those are easier to measure yield higher reliability estimates.
Objectivity of scoring	The more objective the scoring, the greater the reliability

(Ary *et al*, 2010)

### 3. Difficulty Level

Difficulty level means how hard students' ability to answer the question. Difficulty is defined as the percentage of test takers who pass the item (Kaplan and Saccuzzo, 2012). Thus, difficulty level is the number of people who get an item correctly.

The optimal difficulty level for items is usually about halfway between 100% of the respondents that getting the item correct. Item difficulty is only one way to evaluate test items (Kaplan and Saccuzzo, 2012).

According to Crocker and Algina (2008) when an item is dichotomously scored the mean item score corresponds to the proportion of examinees who answer the item correctly. This proportion for item  $i$  is usually denoted as  $p_i$  and is called the item difficulty. Recall that the value of  $P_i$  may range from .00 to 1 .00.

Then, according to Crocker and Algina (2006) stated Thus, the easier item actually has the higher item difficulty value. This can be formulated as following below:

$$P_i = \frac{\text{number of person with a score of 1 item test}}{N_i}$$

Where:

$P_i$  = Difficulty index of item  $i$

$N_i$  = Number of correct answer plus number of incorrect answer of item  $i$

(Crocker and Algina, 2006)

#### 4. Discriminating Power

The way to examine the relationship between performance on particular items and performance on the whole test is known as discriminability (Kaplan and Saccuzzo, 2012).

Item discriminability determines whether the people who have done well on particular items have also done well on the whole test. One can evaluate the discriminability of test items in many ways (Kaplan and Saccuzzo, 2012).

One simple discrimination parameter, called the index of discrimination (Crocker and Algina, 2008). Crocker and Algina (2008) stated that to measure the discrimination index, first identified upper and lower group, so index the index of discrimination (D) is computed as:

$$D = p_u - p_l$$

(Crocker and Algina, 2008)

Where  $p_u$  is the proportion in the upper group who answered the item correctly and  $p_l$  is the proportion in the lower group who answered the item correctly. Values of D may range from - 1 .00 to 1 .00. Positive values indicate that the item discriminates in favor of the upper group; negative values indicate that the item is a reverse discriminator, favoring the lower-scoring group.

## 5. Distractor

When taking a multiple-choice examination, must determine which of several alternatives is “correct.” Incorrect choices are called distractors. If demonstrate in the section on item analysis, the choice of distractors is critically important. Studies have shown that it is rare to find items for which more than three or four distractors operate efficiently. Ineffective distractors actually may hurt the reliability of the test because of the time-consuming to read and can limit the number of good items that can be included in a test (Kaplan and Saccuzzo, 2012). Because test takers got correct answer by guessing so we can formulate the guessing on test is:

$$\text{Corrected score} = R - \frac{W}{n-1}$$

Where

R = the number of right responses

W = the number of wrong responses

n = the number of choices for each item

(Kaplan and Saccuzzo, 2012)

**b. Questioning ability**

The questioning ability is another variable that measured in this study. The questioning ability can be measured by question that proposed. The question that asking by students was analyzed by using rubric that based on Question Category System of Science (Table 3.3)

**Table 3.3 Rubric of students' question ability based on Question Category System for System**

	<b>Cognitive-Memory Operation</b>	<b>Convergent Thinking Operation</b>
<b>Closed Question</b>	<p>Cognitive-memory questions are those questions in which the students' memory is stimulated involving factual recall and involve recognition.</p> <p>1. Recall includes repeat, duplicate, or memorized definitions</p> <p>Examples of question:</p> <ul style="list-style-type: none"> <li>• Apa yang dimaksud dengan rantai makanan?</li> <li>• Siapa yang meduduki posisi paling dasar dalam piramida makanan?</li> <li>• Apa yang dimaksud dengan simbiosis mutualism?</li> <li>• Apa arti dari herbivora?</li> </ul>	<p>Convergent thinking questions are those questions in which the student utilizes analysis or integration of given or remembered data and focuses thinking toward a possible answer. (Associate, reformulate, synthesize, closed predication, and critical judgment levels.</p> <p>1. The associate level includes questions that are utilized to discriminate and classify responses.</p> <p>2. The reformulate level requires reformation of a response from the asked question. Apply level uses previously gathered information to solve a new or different problem.</p> <p>3. The synthesize level requires the student to combine parts of answers to</p>



		<p>formulate a response.</p> <p>4. Closed prediction questions have limitations imposed by conditions or the evidence.</p> <p>5. Critical judgment questions call for a student to make judgment responses by using standards commonly known by the class)</p> <p>Examples of question:</p> <ul style="list-style-type: none"> <li>• Jika populasi ulat berkurang, siapakah yang organisme pertama yang mengalami penderitaan?</li> <li>• Mengapa bentuk piramida makanan seperti segitiga, semakin atas semakin sempit? Jelaskan hubungannya dengan transfer energi.</li> <li>• Apa perbedaan antara rantai makanan dan jaring-jaring makanan?</li> <li>• Elang adalah karnivora Elang memakan bangkai singa Posisi elang pada piramida makanan ada di posisi paling atas piramida Berdasarkan informasi diatas dapat disimpulkan tingkat tropik pada hewan adalah?</li> </ul>
<b>Open Question</b>	<b>Divergent thinking Operations</b>	<b>Evaluative Thinking Operation</b>
	<p>Divergent thinking question are those with many possible responses used to stimulate original responses (divergent thinking questions, are further subdivided into the classifications of opinion, open prediction, and inference.</p> <ol style="list-style-type: none"> <li>1. Opinion questions require the student to state his or her opinion without a specific desired response.</li> <li>2. Open prediction questions occur when there is data insufficient to limit a response.</li> </ol>	<p>Evaluative thinking questions are those that involve the use of standards or criteria and deal with matters of value, cognitive and or affective domains (justify, design, judge A, and judge B).</p> <ol style="list-style-type: none"> <li>1. Justify questions are those in which the student responds with a plan of action, behavior, or a position taken. Design questions require the student to utilize new methods to formulate hypotheses, and</li> </ol>

	<p>3. The infer classification requires a student to conclude by reasoning from something previously known.)</p> <p>Examples of question:</p> <ol style="list-style-type: none"> <li>1. Terdapat populasi hewan herbivora tinggal di savana, pada saat itu keadaan alam di savanna sangatlah tandus dan tidak ada lagi tumbuhan hijau semua kering dan mati, bagaimana hewan herbivora bertahan hidup dan beradaptasi dengan lingkungannya?</li> <li>2. Bagaimana jika semua organisme di bumi ini memiliki "eating habits" yang sama?</li> <li>3. Faktor apa saja yang mempengaruhi adanya rantai makanan?</li> <li>4. Bagaimana energi dapat berpindah dari satu organisme ke organisme lain? Melalui proses apa?</li> </ol>	<p>conclusions.</p> <ol style="list-style-type: none"> <li>2. Judge A questions require the student to respond within the affective domain by making judgments as a matter of value.</li> <li>3. Judge B questions require the student to evaluate and make judgments based upon the students' cognitive behaviors.)</li> </ol> <p>Examples of question:</p> <ol style="list-style-type: none"> <li>1. Menurut kamu, bagaimana caranya untuk tetap untuk menjaga keselarasan ekosistem? Dan bagaimana agar proses makan memakan (jaring-jaring makanan) tetap terjadi?</li> <li>2. Jika populasi konsumen tertier lebih banyak jumlahnya dibandingkan jumlah populasi produsen, apakah ekosistem dan proses makan memakan tetap terjadi? Dan keselarasan ekosistem tetap terjaga?</li> </ol>
--	--	--

### c. Observation Sheet.

In this research Observation Sheet used to make sure the teaching procedure of Reciprocal Teaching is implemented. It will specifically assess whether the step conducting by the researcher is appropriate or not.

## G. Research Procedure

The research procedure arranged based on the syntax of Reciprocal Teaching. There are three stages of research procedure which consist of preparation stage, implementation stage and analysis and conclusion stage. Those stages will be explained as follow:

### 1. Preparation stage

This stage, several stages are decided to become starting point of doing research and the stages are explained as follow:

- 1) Formulating problem that is investigate.
- 2) Analyzing of KTSP Curriculum
- 3) Determining focus of variable research.
- 4) Arranging literature review and analysis of Reciprocal Teaching, concept comprehension and questioning ability.
- 5) Arranging literature review and analysis of concept Ecosystem.
- 6) Arranging research proposal which is will presenting in proposal seminar.
- 7) Revising all and prepare and make research instrument after having suggestion.
- 8) Research instrument validation by expert.
- 9) Revising research instruments.

## **2. Implementation Stage**

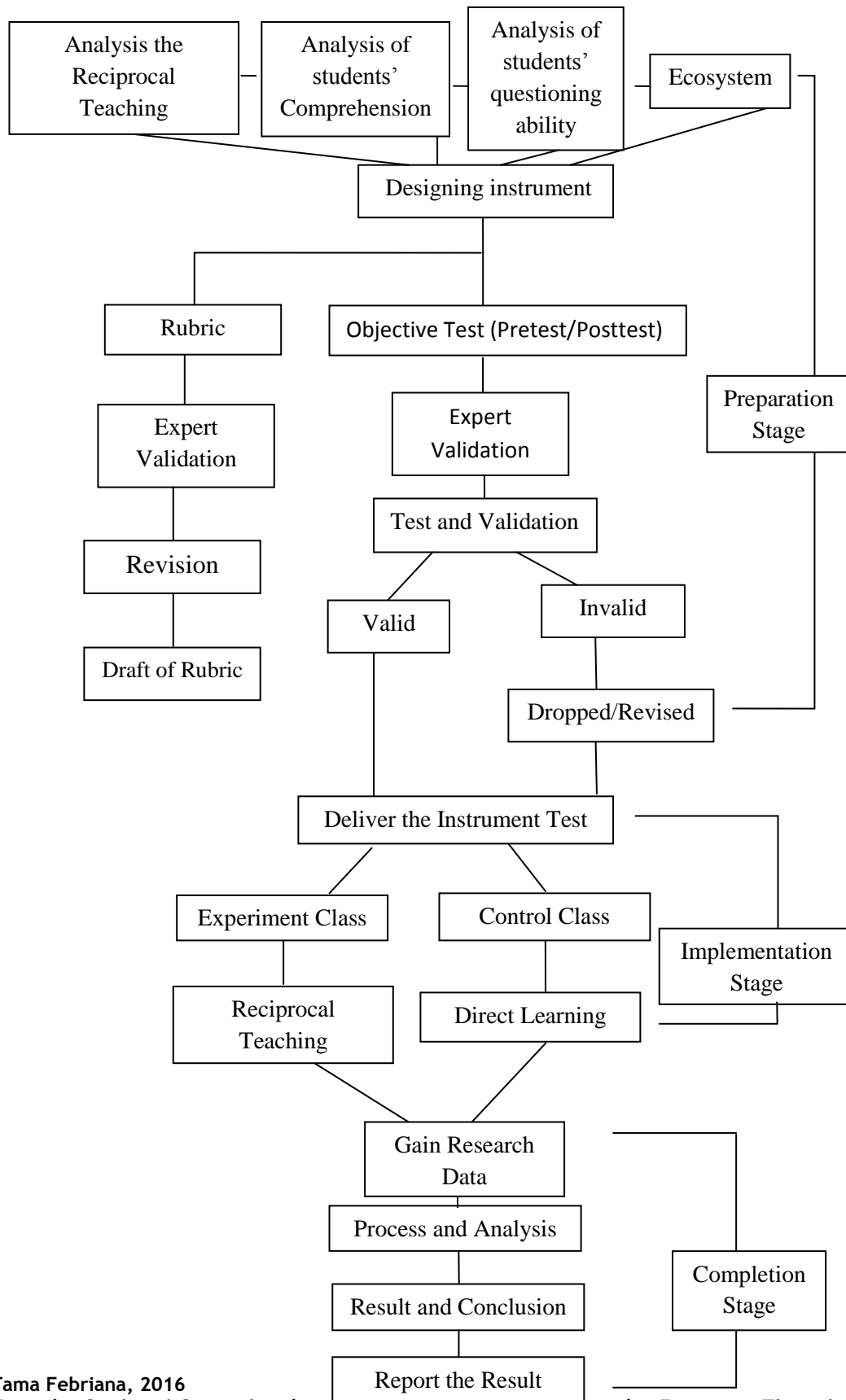
This stage explains about how research is implemented and stages can describe as follow:

- 1) Determining experiment class and control class.
- 2) Giving pre-test to sample class to know the initial condition of student (prior knowledge).
- 3) Processing the pre-test result.
- 4) Conducting the Reciprocal Teaching in experiment class while in control class students learn the Ecosystem using common lecturing (direct learning).
- 5) Collecting the worksheet to measure students' questioning.
- 6) Giving post-test in the experiment class to recognize the improvement students' comprehension.

## **3. Completion Stage**

This is the last stage of research procedure. There are several steps that are conducted in this stage, such as:

- 1) Analyzing the data of the whole research.
- 2) Discussing the result and make conclusion from the data of analysis result.
- 3) Arranging the report of the research



Tama Febriana, 2016  
*Fostering Students' Comprehension and Questioning Ability in Learning Ecosystem Through Reciprocal Teaching*  
 Universitas Pendidikan Indonesia Figure 3.1 Flowchart of Research Plot

## H. Research Instrument Judgment

### a. Objective Test Validation

In this study, pretest and posttest were used to measure the improvement of students' concept mastery. Before implemented the objective test on experiment and control classes, the objective test needs be validated. The reliability, difficulty level, discriminating power, and distraction point of the objective test were investigated. Therefore, the test consists of 33 multiple choice questions which tested to 7<sup>th</sup> grade students in secondary level. The questions which accepted were 20 questions. Table 3.4 presents the recapitulation of test items based on cognitive level. Meanwhile Table 3.5 presents the recapitulation of test items validation by ANATES. Reliability test was 0.75 (Very high degree).

**Table 3.4 Recapitulation of Test Items Based on Cognitive Level**

No	Cognitive Level	Total of Item Test	The position in number
1	C1 (Remembering)	5	1,2,9,10,16
2	C2 (Understanding)	10	3,4,5,11,12,13,14,17,18,19
3	C3 (Applying)	2	7,8,
4	C4 (Analyzing)	3	6, 15,20
<b>Total</b>		20	

**Table 3.5 Recapitulation of Test Items Validation by ANATES**

No	Discriminating Power	Difficulty level	Validity	Sign. Correlation	Decision
1	Poor	Very easy	Invalid	-	Dropped
2	Poor	Very difficult	Invalid	-	Dropped
3	Poor	Very easy	Moderate	Very significant	Used
4	Good	Medium	Moderate	Very significant	Used
5	Satisfactory	Medium	Moderate	Very significant	Used
6	Poor	Medium	Low	Significant	Used

No	Discriminating Power	Difficulty level	Validity	Sign. Correlation	Decision
7	Poor	Medium	Invalid	-	Dropped
8	Poor	Very easy	Invalid	-	Dropped
9	Poor	Very easy	Low	Significant	Used
10	Good	Medium	Moderate	Very significant	Used
11	Satisfactory	Medium	Low	-	Dropped
12	Satisfactory	Medium	Low	Very significant	Used
13	Satisfactory	Medium	Low	Significant	Used
14	Poor	Easy	Invalid	-	Dropped
15	Poor	Medium	Invalid	-	Dropped
16	Poor	Very easy	Low	Significant	Used
17	Poor	Easy	Invalid	-	Dropped
18	Poor	Medium	Invalid	-	Dropped
19	Good	Easy	Moderate	Very significant	Used
20	Satisfactory	Easy	Moderate	Very significant	Used
21	Satisfactory	Medium	Low	Significant	Used
22	Good	Medium	Low	Significant	Used
23	Excellent	Medium	High	Very significant	Used
24	Good	Medium	Moderate	Very significant	Used
25	Satisfactory	Very easy	Moderate	Very significant	Used
26	Poor	Very easy	NAN	NAN	Dropped
27	Poor	Very easy	Invalid	-	Dropped
28	Poor	Medium	Invalid	-	Dropped
29	Satisfactory	Easy	Low	Significant	Used
30	Satisfactory	Easy	Low	Significant	Used
31	Poor	Very easy	Low	Significant	Used
32	Poor	Medium	Invalid	-	Dropped
33	Satisfactory	Medium	Moderate	Very significant	Used

### b. Instrument Non-Test

Instruments that used in this study not only in form of test, the instrument non-test was used in order to get qualitative data. Those instruments are rubric and observation sheet. Rubric of question category system for science is used to measured students' questioning ability. Questions was proposed by students are assessed into some

criteria which are available in the rubrics. Meanwhile, observation sheet used to assess whether the step conducting by the researcher is appropriate or not.

### c. Objective Test Categorization

After the data of the test item score is obtained, then the data was processed by calculating gain score and normalize gain score. Gain score reflects improvement result from pretest to the posttest. It could be assumed as the effect of treatment. After the gain score is obtained, then Normalized Gain (N-Gain) can be obtained. The normalized gain score is used to determine the categories of students' achievement improvement.

According to Hake (1999), Formula of gain is

$$G = S_f - S_i$$

Where

$G$ = Gain Score

$S_f$ = Posttest score

$S_i$ = Pretest score

(Hake, 1999)

Meanwhile, the effectiveness of reciprocal teaching implantation in increasing students' comprehension in learning ecosystem can be seen from the result of the normalized gain that achieved by students. The normalized gain better to show the improvement rather than actual gain score because the improvement of higher and lower achiever students can be shown clearly.

The N-Gain can be calculated by the following formula.

$$\langle g \rangle = \frac{S_f - S_i}{S_{max} - S_i}$$



Where

$\langle g \rangle$  = Normalized Gain Score

$S_f$  = Posttest score

$S_i$  = Pretest score

$S_{max}$  = Maximum score

(Hake, 1999)

Then, the Normalized gain value is determined based on Table 3.7.

**Table 3.6 N-gain Categorization**

N-gain Score Category	
Result	Category
$g > 0,7$	High
$0,3 > g \geq 0,7$	Medium
$g \leq 0,3$	Low

(Hake, 1999)

### I. Data Statistical Analysis

Data Statistical Analysis needed to support the data of improvement students' comprehension and to prove the hypothesis test. Data statistical analysis is helped by IBM SPSS Statistics version 23. The steps were obtained as follow.

#### 1) Normality test

The first to process data statistic is testing the normality of the data. Normality test aims to know the sample which comes from population has normal population distribution or not. Normality statistic test is done by using SPSS 23, Kolmogorov-Smirnov or Shapiro-Wilk with significance level ( $\alpha$ ) is 0.05.

In this study the researcher used Shapiro-Wilk to test the normality. The Shapiro-Wilk test is used to calculates a  $W$  statistic that tests is to know whether a random samples,  $x_1, x_2, \dots, x_n$  comes from (specifically) a normal distribution. The small values of  $W$  statistic indicate of departure from normality and percentages points from the  $W$  statistic (Shapiro and Wilk, 1965).

The  $W$  Statistic is calculated as follows:

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{(\sum_{i=1}^n (x_i - \bar{x})^2)}$$

(Shapiro and Wilk, 1965)

Hypothesis that is tested in this study are as follow:

$H_0$ : Data (control and experiment class) comes from normal distribution.

$H_a$ : Data (control and experiment class) comes from not normal distribution.

The normality test is tested using Shapiro-Wilk at the significance level ( $\alpha = 0.05$ ), with test criteria are as follow:

- a) If significance score  $< 0.05$ , thus data is not from normal distribution
- b) If significance score  $\geq 0.05$ , thus data is from normal distribution

The ways to know the data are normal just see the significant of Shapiro-Wilk or Kolmogorov-Smirnov, if the result  $> 0.05$  indicates the data are normally distributed.

## 2) Homogeneity Test

This research used Levene's test to test the homogeneity. Levene's test (Levene, 1960) is used to tests if  $k$  samples have equal variances. The equal variance across sample is called homogeneity variance. Homogeneity test is needed because to determine a sample from population that is originated from two classes that homogenous.

The homogeneity test used significance level ( $\alpha$ ) is 0.05. When significance value is  $\geq 0.05$ , data is considered as homogenous (Sudjana, 2005).

The Levene test is defined as:

$$H_0 : \sigma_1 = \sigma_2 = \dots = \sigma_k$$

$H_a$  :  $\sigma^2_i \neq \sigma^2_j$  for at least one pair (i,j).

Test Statistic : given a variable a Y with sample of size N divided into k subgroup, where  $N_i$  is the sample size of the  $i^{th}$  subgroups, the levene test statistic is defined as:

$$W = \frac{(N - k)(k - 1) \sum_{i=1}^k N_i (\bar{Z}_i - \bar{Z}_{..})^2}{(k - 1) \sum_{i=1}^k \sum_{j=1}^{N_i} (\bar{Z}_{ij} - \bar{Z}_i)^2}$$

(Levene, 1960)

Hypothesis that is tested in this study are as follow:

$H_0$ : data (control and experiment class) have variances homogenous.

$H_a$ : data (control and experiment class) haven't variances homogenous.

The homogeneity test is tested using ANOVA (one-way) at the significance level ( $\alpha = 0.05$ ), with test criteria are as follow:

- a) If significance score  $< 0.05$ , thus data haven't variances homogenous.
- b) If significance score  $\geq 0.05$ , thus data have variances homogenous.

If two samples of are homogenous variance hence difference of both means can be done by using T test. On the contrary, if the data are not homogenous variance, the differences of two means can be done using non-parametric test.

### 3) Hypothesis Test : Parametric Test

Parametric test is used in certain condition of result data research; the normal data distribution and homogeneity data. Parametric test

method are classify in to one sample T-test, independent sample T-test, Summary Independent sample T-test and paired sample T-test. Meanwhile, this research was adapted the Independent sample T-test because the data is two unrelated groups on the same continuous, dependent variable.

#### 4) Independent Sample T-test

The independent-samples t-test is one of type parametric test. Independent sample t-test is compares the means between two unrelated groups on the same continuous, dependent variable. The Independent t method tests the null hypothesis that the population means related to two unpaired data, random samples from an approximately normal distribution are equal (Altman, 1991; Armitage and Berry, 1994).

Assuming equal variances, the test statistic is calculated as:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$s^2 = \frac{\sum_{j=1}^{n_1} (x_j - \bar{x}_1)^2 + \sum_{i=1}^{n_2} (x_i - \bar{x}_1)^2}{n_1 + n_2 - 2}$$

Where

$\bar{x}$  = the sample means

$s^2$  = the pooled sample variance

$n_1$  and  $n_2$  = the sample sizes

$t$  = is a Student t quantile with  $n_1 + n_2 - 2$  degrees of freedom

(Minium et al., 1993)

According to Sudjana (2005), if the level of significance (sig) is  $\leq 0.05$   $H_0$  is rejected. Meanwhile if the level of significance (sig)  $\geq 0.05$   $H_1$  is retained.

a) If significance score  $< 0.05$ , Thus, there are no differences in student comprehension in learning ecosystem through Reciprocal Teaching

b) If significance score  $\geq 0.05$ , Thus, there are differences in student comprehension in learning ecosystem through Reciprocal Teaching

The hypothesis of the difference of the average posttest result in experiment class is determined as follow.

$H_0$ : There are no differences in student comprehension in learning ecosystem through Reciprocal Teaching.

$H_1$ : There are differences in student comprehension in learning ecosystem through Reciprocal Teaching.