CHAPTER III

RESEARCH METHOD

This chapter presented the research method. It discused the method was used in this study. The discussion covered research design, variable of the study, population, data and data source, data collecting method, research instrument, and method of data analysis, and hypothesis testing.

A. Research Design

In this research, the researcher chose experimental method. The experimental method is the only method of research that can truly test hypotheses concerning cause and effect relationship. It is represent the most valid approach to the solution of educational problems, both practical and theoretical, and to the advancement of education as a science. In an experimental study, the researchers manipulated at least one independent variable, controlled over relevant variables, and observed the effect on one or more dependent variable. The researcher determined "who get what," which group of subjects gets which treatment. This manipulation of the independent variable was the one characteristic that differentiates all experimental research from the other methods of research (Gay, 1992: 298).

The experimental research design is classified into pre-experimental design, true experimental, and quasi-experimental. Pre-experimental research do not have random assignment of subjects to groups or other strategies to control

extraneous variables. True-experimental research use randomization and provides maximum control of extraneous variables. Whether quasi-experimental research lack randomization but employ other strategies to provide some control over extraneous variables (Ary, 2002: 302). In this research, the researcher used pre-experimental research. Pre-experiment research involved only one group which was pretested, exposed to treatment, and post tested. The success of the treatment was determined by comparing the pretest and the posttest score.

The researcher gave treatment to the student. Before and after treatment, students were tested using (pre-test and post test). These tests will chose as they were available. To know the improvement of students' vocabulary listening ability, the scores of pre-test and post-test were compared. If the result of the pos-test after the treatment was higher than the pre-test, it means that the treatment can improve the students' vocabulary listening ability.

The research design is as follow:

| X1 |——————| X2 |
     | X |
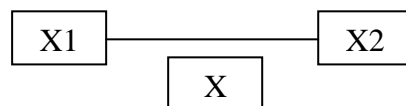
Notes:

X1   : Pre-test

X2   : Post-test

X   : Treatment

B.   Population, Sample and Sampling

1. Population

Gay (1992) states "regardless of the technique to be used in selecting a sample, the first step in sampling is definition of the population. The population is the group of interest to the researcher, the group to which she or he would like the results of the study to be generalizable". A population is defined as all members of any well-defined class of people, events, or objects. It means that population is the largest group. In this research, the population of the research included all students of Perwanida Kindergarten. In this case, Perwanida kindergarten only had two classes (A and B class). A class consist 32 students, while B class consist 30 students. Total population here 62 students included two classes (A and B class).

2. Sample

Sample is a portion of a population. Since conducting a study generally requires a great deal of time and energy, non-generaliable results were extremely wasteful. If all result were true only for the group on which they were based, researcher could never benefit from anyone else's work and each and every study would have to be replicated and almost infinite number of time. That was why the researcher took a sample that represented the whole of population.

In this research, the researcher chose 30 students at B class to be an experimental group and the control group. The researcher chose B group as sample with consideration. Based on the teacher recommended, students at B class were better in English achievement than A class. Therefore, the researcher to choose B class to be a sample of the research.

3. Sampling

Sampling is the process of selecting a number of individuals for a study in such a way that the individuals represent the larger group from which they were selected. The individual selected comprise a simple and the larger group is referred to as a population. The purpose of sampling is to gain information about a population; rarely is a study conducted that includes the total population interest as subject (Gay, 1992: 123). Here, sampling was very important way to obtain a group of a subjects who would be representative of the larger population or would provided specific information needed.

In this research, the researcher used purposive sampling. Based on Ary (2006: 156), Purposive sampling also referred to as judgment sampling-sample elements judged to be typical, or representative, are chosen from the population. The assumption is that errors of judgment in the selection will counterbalance one another. The critical question in purposive sampling is the extent to which judgment can be relied on to arrive at a typical sample. There is no reason to assume that the units judged to be typical of the population will continue to be typical over a period of time. Consequently, the results of study using purposive sampling may be misleading. Because of its low cost and convenience, purposive sampling has been useful in attitude and opinion surveys. Be aware of the limitation, however, and use the method with extreme caution.

In other words, the researcher should be sure that the sample was representative and suitable with the purpose of research. As the process of sampling, the researcher finally decided to choose B class that consists of 30 students with consideration. They were assumed to be homogeneous by recommended of teacher that handles B group Students' of Perwanida Kindergarten. In order to apply the experimental stage, the samples must be good in English achievement, especially in listening ability. It's intended to reduce the extraneous variable may appears since the design was pre-experimental research without control group.

C.   Research Variable

A variable is the characteristic or attribute of an individual, group, educational system, or the environment that is of interest in reseach study. Experimental research includes two types of changeable factors, or variables: independent and dependent. An independent variable is a manipulated, influential, experimental factor. It is a potential cause. The independent is used because this variable can be manipulated independently of other factors to determine its effect. Researcher have a vast array of options open to them in selecting independent variables, and one experiment way include several independent variable (Santrock, 2004: 47). While a dependent variable is a factor that can change in an experimental research, in response to changes in the independent variable. As researcher manipulate the independent variable, they measure the dependent for any resulting effect.

Dependent variable on this research was students' vocabulary listening ability, while the independent variable was drawing technique.

D.   Data Collection Method

Data collection method is the technique to collect the data that is needed by the researcher. In this research, the technique that is used by the researcher is administering test. The test was in the form of multiple choices. The students listened and checked the correct answer based on what researcher said.

The researcher used written test to get the data. The researcher administered the test to all of respondents. The tests were conducted two times before and after getting treatment. Then it would be give scored in order to find out the score of the test. In this research, the techniques to collect the data were pretest, treatment, and posttest. The researcher conducted pretest before gave treatment to the students.

Treatment was conducted after the researcher gave pretest to the students. The researcher conducted treatment two meetings by giving drawing technique related to theme of kindergarten school. More detail about the theme *(see appendix 2).* Here the researcher gave treatment at XVI weeks, the theme was about Nature. After knew what the theme that would be taught on that week, the researcher gave treatment to students.

E.   Research Instrument

Research instrument refers to any equipment used to collect the data (Arikunto, 2010:262). Based on Ary (2006: 200) selecting appropriate and useful measure instruments is critical to the success of any research study. One valuable technique to measuring instruments for educational research is use tests. A test is a set of stimuli presented to an individual in order to elicit responses on the basis of which a numerical score can be assigned. This score, based on a representative sample of the individuals' behavior, was an indicator of extend to which the subject has the characteristic being measured.

There were two kinds of test that was used by researcher. The tests were called by pretest (X1) and posttest (X2). In this research, the researcher also gave treatment (X) to the students. As in the explanation in the research design, the researcher gave pretest first and continued by conducting treatment then the last were posttest.

- Pre-test is a test which is given before the treatment. The purpose of giving pre-test was to know about the students' vocabulary listening ability. The test was in the form of written test (see appendix 4). The test would administrated in a classroom during school hours. The research instructed the students to do the test. There were 10 number of tests that was tested before conducted the treatment. The researcher gave 30 minutes to do the pretest. The students had to listen the instruction before doing the test. The instruction was to listen carefully and check ($\sqrt{}$) the picture below based on what teacher says. Here the researcher made teacher guide (see

appendix 6). Teacher guide was the list of vocabulary that would be heard by students to do the test. Teacher read the list one by one, the students listened it then check below the picture. In this section, students had to focus and listen carefully what teacher said and check the picture, which was suitable with the vocabulary that had been said by teacher.

- Post-test is a test which is given at the end of the research. The purpose was to measure how far the students understand the lesson and also to know whether the technique used mime story was effective or not in improving students' vocabulary listening ability in B class of TK Perwanida Tulungagung academic year 2013 / 2014. Same as pretest, the researcher gave 10 number test and limit the time 30 minutes to do the entire test. (*For more detail about posttest, see appendix 5*).

F. Validity and Reliability Testing

In collecting data process, the researcher used instrument to get the data that was needed. Developing an instrument for a particular study also had several major drawbacks. The development of a good instrument required considerable time, effort, and skill. In this case, the researcher should make a good test that can really measure the students' listening comprehension skill, not in other skill. The researcher was always dependent upon measurement. There were two important characteristics to measure instrument. The instrument should get validity and reliability.

1. Validity

The most simplistic definition of validity is that is the degree to which a test measures what it is supposed to measure. A common misconception is that a test is, or is not, valid. Since tests are designed for a variety of purposes, and since validity can be evaluated only in terms of purpose, it is not supervising that there are several different types of validity: content validity, construct validity, face validity.

Based on the Allison (1999: 85) The validity of a test concerns whether it is measuring what we think and say it is measuring; for example, is a reading comprehension test item testing reading comprehension ability or something else, such as knowledge of grammar without reference to the reading text.

Based on Ary (2006: 228) Here are the explanations each types of validity:

a. Face validity

    Face validity is a term sometimes used in connection with a test's content. Face validity refers to extent to which examinees believe the instrument is measuring what it is supposed to measure. In this research, the researcher ensured that the test items were valid in term of face validity because the test was in the form of objective test (multiple choices) which was consulted by advisor and teacher of B class at Perwanida Kindergarten Tulungagung.

b. Content Validity

    Content validity is a degree to which a test measures an intended content area. Content validity is of prime importance for achievement

test. A test score can not accurately reflect a students' achievement if it does not measure what the student was supposed to learn. Content validity is determined by expert judgment. There is no formula by which it can be computed and there is no way to express it quantitatively. Usually experts in the area covered by the test are asked to assess its content validity. These experts carefully review the process used in developing the test as well as the test itself and make a judgment concerning how well items represent the intended content area.

The researcher ensured that the test items were valid in term of content validity, because the test items were made based on curriculum of Kindergarten level. Based on the curriculum the students had to be able to listen and tell simple story or experience. Here, the researcher used drawing technique which is suitable with the indicator of the curriculum. In addition, the researcher conducted the research at 16th weeks. Based on the curriculum, the theme of 16th weeks was Nature. So, the researcher used appropriate vocabulary based on the theme (Nature).

c. Construct Validity

Construct Validity is the degree to which a test measures an intended hypothetical construct. If a test has construct validity, it is capable of measuring certain specific characteristic in accordance with a theory of language behavior and learning.

According to Buck at Allison (1999: 121) it is important for many purposes to ensure that some parts of language tests focus primarily on listening abilities. In this research, the researcher focused on how to test students' vocabulary listening ability. In this case, the researcher used multiple choices to test the students' comprehension. The tests were to choose kind of pictures, which were suitable with what teacher said. Based on the statement above, the researcher ensured that the tests were valid in the term of construct validity.

2. **Reliability**

In everyday English, reliability means dependability, or trustworthiness. The term means essentially the same thing with respect to measurement. Basically, reliability is the degree to which the test consistently measures whatever it measures. The more reliable a test is, the more confidence we can have that the scores obtained from the administration of the test are essentially the same scores that would be obtained if the test were read ministered. An unreliable test is essentially useless; if a test is unreliable, then scores for a given sample would be expected to be different every time the test was administered (Gay, 1992: 161). Based on Allison (1999: 85) The reliability of a test concerns the accuracy and trustworthiness of its result; if we could erase the test from students' memories and then repeat it, how similar would the results be? Reliable test results will accurately reflect each students' understanding of whatever is being tested.

Reliability is expressed numerically, usually as a coefficient, a high coefficient indicates high reliable. If a test were perfectly reliable, the coefficient would be 1.00, this would mean that a students' score perfectly reflected her or his true status with respect to the variable being measured. However, no test is perfectly reliable.

In this research, the researcher conducted tests (try out) before conducted pretest and posttest to the students. The researcher gave tryout tests to 10 students who would be a volunteer to know whether the tests were reliable or not. Here, the researcher analyzed the result of students' tried out by using SPSS 16.00 version. The researcher also analyzed the tryout score to support it whether the tests were reliable or not. *(For more detail about students' tryout analysis of pretest and posttest, see appendix 6).* The result of computing try out 1 (pretest) and 2 (posttest) could be seen as below:

Table 3.1 Reliability statistic of tryout 1 pretest

Reliability Statistics

Cronbach's Alpha[a]	N of Items
-.672	10

Table 3.2 Reliability statistic of tryout 2 posttest

Reliability Statistics

Cronbach's Alpha[a]	N of Items
-.777	10

According to triton in sujianto (2009:97) the value of cronbach's alpha can be interpreted as follow:

Table 3.3 Cronbach's Alpha Interpretation Based on Triton

Cronbach's alpha	Interpretation
0,00 – 0,20	Less reliable
0,21 – 0, 40	Rather reliable
0,41 – 0, 60	Quite reliable
0,61 – 0,80	Reliable
0,81 – 1,00	Very reliable

Based on the Table 3.1 and 3.2, the result of calculation shows that reliability coefficient of tryout 1(pretest) is 0.672 and tryout 2 (posttest) is 0.777. It shows that the test (tryout 1 and 2) is reliable according to Triton's Cronbach's Alpha Interpretation (Table 3.3).

G. Normality and Homogeneity Testing

1. Normality Testing

Normality tests are used to determine whether a data set is well-modeled by a normal distribution or no, or to compute how likely an underlying random variable is to be normally distributed population.

In this research, the researcher used *Kolmogorov-Smirnove Test* with SPSS 16.00 version. Kolmogorov-*Smirnove Test* was a test of normality for large samples. If the results are significant, then the null hypothesis (H_0) was rejected.

In this research, the researcher analyzed the result of students' pretest and posttest score. *(For more detail about the result of students' pretest and*

posttest score for normality testing, see appendix 7). In this case, the researcher tried to calculate the students, pretest and posttest score used *Kolmogorov-Smirnove Test* with SPSS 16.00 version. *Kolmogorov-Smirnove Test* is a test of normality for large samples. If the results are significant, then the null hypothesis (H_0) is rejected. Simply put a value less than 0.05 indicates that the data are non-normal. If the value more than 0.05 it's indicates that the data are normal. The result of the test could be seen in the table below:

Table 3.4 Table of result pretest and posttest in normality testing

One-Sample Kolmogorov-Smirnov Test

		pretest	posttest
N		30	30
Normal Parameters[a]	Mean	52.00	81.00
	Std. Deviation	13.493	8.847
Most Extreme Differences	Absolute	.174	.198
	Positive	.159	.168
	Negative	-.174	-.198
Kolmogorov-Smirnov Z		.955	1.086
Asymp. Sig. (2-tailed)		.321	.189

a. Test distribution is Normal.

Based on the table 3.5 and 3.6, it was showed that both tests (pretest and posttest) were normal.

2. **Homogeneity Testing**

Homogenity test is intended to show that two or more groups of data samples come from populations having the same variance. To test the

variance of homogeneity the researcher had to know f_{max} value. In homogeneity test f_{value} should be lower than $f_{table.}$ The result of f_{max} value was calculated from students' pretest and posttest score. Before calculated students' score, the researcher analyzed it. *(For more detail about analysis of students' pretest and posttest to test homogeneity, see appendix 8).* Then, the researcher calculated the result of students' pretest and posttest score that were formulated as follows:

$$SD_1^2 = \frac{\sum X_1^2}{N_1} - \left(\overline{X}_1\right)^2$$

$$= \frac{86400}{30} - 2704$$

$$= 176$$

$$SD_2^2 = \frac{\sum X_2^2}{N_2} - \left(\overline{X}_2\right)^2$$

$$= \frac{187100}{30} - 6136.1$$

$$= 100.6$$

$$F_{max} = \frac{S\ max}{S\ min}$$

$$F_{max} = \frac{176}{100.6}$$

$$F_{max} = 1.749$$

$$df_1 = 30 - 1 = 29$$

$$df_2 = 30 - 1 = 29$$

The calculation showed that the result of F_{max} value is 1.749. While df_1 is 29 and df_2 is 29 the value of F_{table} in 5% level is 1.84. In homogeneity test F_{value} should be lower than F_{table}. In this calculation, it can be said that F_{max} value $<$ F_{table} value (1.749 $<$ 1.84). It meant that the variance value of pretest and posttest score is homogeneous.

H.   Data Analysis

To analyze the data statistically, the researcher used matched t-test. It was used since matched t-test was probably the most widely used statistical test for comparison of two means. It could be used with very small sample sizes. It was used as the data coming from the same sample or known as paired data.

This research only one class for the experimental class and there was no control group. Therefore, match-test was used. Match t-test was used to analyze the data to improve the students' listening comprehension through drawing technique. It compared two kinds of data or mean(average score of the students)from similar sample. The procedure of analysis the data both used descriptive and inferential statistic.

1. Descriptive statistic

   In this research, the researcher determined the descriptive statistics that was used to calculate students' score into four steps. First, the calculated mean, median, mode, and the last was standard deviation.

❖ Mean

The first steps were calculating the mean of students' score. Mean is average number of students' score. According to Sudijono (2008) To find the means of students' score, a following simple statistic formula that was used was as follow:

$$M_x = \frac{\Sigma x}{N}$$

Notes :

M_x : Mean (average score)

$\sum x$: Total number of the students' scores

N : Number of Case

❖ Median

Then, after got students' mean then calculated median of the score. According to Sudjiono (2008) "The median is value or a number that divided distribution data into two parts but the number is the same". Here was the formula that the researcher used:

$$\text{Median} = Rall + \frac{\frac{1}{2}N - fkb}{f}$$

Notes :

N : Number of Cases

$Rall$: Real Apparent Lower Limit

f : Frequency

fkb : Frequency lower limit under the mean score class

❖ Mode

According to Sudjiono (2008) " Mode is the score that most frequently arises in each data.

❖ Standard Deviation

The formulation of standard deviation was as follow:

$$SD = \frac{\sqrt{\Sigma x^2}}{N}$$

Notes:

SD   : Standard Deviation

Σx^2 : Total score after getting quadratic process

N   : Number of Case

2. Inferential statistic

Then after calculating the means of pretest and posttest, the data was analyzed by using matched t-test. It was used as the dats come from the same sample or known as paired data. Here the researcher used T-test formula according to Sudijono (2008:284), as follow:

$$t = \frac{M_1 - M_2}{SE_{M1-M2}}$$

Notes:

t   : t value

M_1 : Mean (pretest)

M_2 : Mean (posttest)

SEm   : Standard Error Mean