CHAPTER III

RESEARCH METHODOLOGY

This chapter elaborates the methodology applied in the study. The discussion of this chapter includes research design, population and sample of the research, data collection, research procedures, and data analysis.

3.1 Research Design

This study conducted quasi-experimental design. It was taken because the participants of this study were not chosen randomly, but they were already part of the groups (Jackson, 2009, p.318). In this case, the groups were class VIII-A and VIII-B. Furthermore, this study provided pre-test - post-test to both experimental and control group. It was considered as nonequivalent control group pre-test – post-test design. "Nonequivalent control group pre-test – post-test design is an experimental design in which at least two nonequivalent groups are given a pre-test, then a treatment, and then a post-test measure" (Jackson, 2009, p.323). Schematically quasi-experimental research design using nonequivalent control group pre-test – post-test design is depicted as follow:

Group	**Pre-test**	**Treatment**	**Post-test**
Experimental Group	O_1	X_1	O_2
	.....		
Control Group	O_3	X_2	O_4

Note: - $\mathbf{O_1}$ and $\mathbf{O_3}$ show the initial ability of students' vocabulary mastery.

- $\mathbf{O_2}$ shows the ability of students' vocabulary mastery after given a special treatment.
- $\mathbf{O_4}$ shows the ability of students' vocabulary mastery without given a special treatment.
- $\mathbf{X_1}$ and $\mathbf{X_2}$ refers to certain treatment that given to the students.

(Adapted from Sugiyono, 2010: 79)

Based on the figure above, in order to find out whether the treatment has effect on the performance of the groups, the differences between $\mathbf{O_1 - O_2}$ and $\mathbf{O_3 - O_4}$ were further compared.

3.2 Population and Sample

Population of this study was all students of second grade in one junior high school in Bandung enrolled in academic year 2012/2013.

To determine the sample of this study, the sampling technique used was purposive sampling. Purposive sampling is a sampling technique with particular consideration (Sugiyono, 2010: 85). This study took two classes as the sample of the study. They were class VIII-A and VIII-B, however, only 30 students from each classes were decided to be the sample of this study. The first 30 students from class VIII-A participated as experimental group and other 30 students taken from class VIII-B participated as control group. The focus of this study was on improving students' vocabulary mastery in composing descriptive text. Based on that condition, the chosen

subjects were eighth-grade students where they study descriptive texts. It was taught in odd semester which was coincided with the time of this study taken.

3.3 Data Collection

3.3.1 Research Instrument

Pre-test – post-test in the form of writing test were employed as the instruments to collect the data for this study in order to be able answering the first research question. The writing test given was focused on collecting the data of students' vocabulary mastery. That was in line with the aim of this present study. The writing test given was consisted of two kinds of test. They were "fill in the blank" form and describing some pictures given.

In addition, another instrument used in this study was questionnaire. The questionnaire was given to the experimental group and the results of it were used to support the data of the first research question. In detail, the instruments of this study could be more explained as below:

3.3.1.1 Pre-test

Pre-test was given to both experimental and control group in the beginning. It was aimed to find out the initial ability of students' vocabulary mastery before the groups was given a different treatment. The test consisted of 20 questions in the form of "fill in the blank" and 20 questions of describing some pictures given.

3.3.1.2 Post-test

Post-test was given to both experimental and control group after the different treatment given to the both groups. The post-test given was aimed to find out the students' progress of vocabulary mastery after the treatment conducted to the both groups. The test was similar to pre-test; it consisted of 20 questions in the form of "fill in the blank" and 20 questions of describing some pictures given.

3.3.1.3 Questionnaire

Questionnaire was given only to the experimental group who has experienced the treatment of the use of audio visual aids during their learning process. The questionnaire given was aimed to find out the students' responses toward the vocabulary learning and the use of audio visual aids in their teaching learning process. The questionnaire consisted of 10 closed-ended questions. Closed-ended question is a question that contains a set of answer that a respondent chooses (Beins, 2012: 264). Every single of question item was related to the topic and condition when the students used audio visual aids in their learning process. The format of questionnaire was based on Likert Scale. Likert Scale is used to measure the attitudes, opinions, and perceptions of a person or group of people about social phenomena (Sugiyono, 2010: 93). Likert Scale consisted of five responses category: strongly agree, agree, neutral, disagree, and strongly disagree.

3.4 Research Procedures

3.4.1 Preparation

The schedule program was planned for nine meetings. The first meeting was used to conduct a tryout of the instruments of this study. The next meeting was used to conduct the pre-test and the third to seventh meetings were allocated to implement the lesson in the classroom. The post-test and questionnaire were administered in the rest two meetings.

A preparation in doing the research in this study involved designing lesson plan and preparing materials. There were five lesson plans provided to be implemented in the treatment of the use of audio visual aids for experimental group. The lesson plans were designed based on 2006 English School-Based Curriculum for eighth grade students. The five lesson plans were all based on the standard competence and competence of descriptive text learning. In addition, for the control group, the lesson plans were designed similar to the experimental one. The differences were on the learning media and materials used.

Learning materials in teaching learning process of the experimental group included descriptive texts, pictures (in slides), audio files, and the limited authentic materials available in the classroom that were related to the descriptive texts topic. On the other side, the learning materials used in the teaching learning process in the control group included descriptive texts, textbook that were related to the descriptive texts topic.

3.4.2 Instrument of Try Out

In the very beginning, the instrument test was administered to students outside the experimental and control group in order to measure the validity, reliability, difficulty level, and discrimination level of the instruments employed. So, after it's proven then it could be surely say that the experimental and control group used good instrument test to assessed their vocabulary mastery.

3.4.3 Pre-test

Pre-test was administered to both experimental and control group in the first meeting. It was aimed to find out the initial ability of students' vocabulary mastery. As many 60 students from the experimental and control group taken the test.

3.4.4 Treatment

The treatment program was done in five meetings in the experimental group. The use of audio visual aids in the teaching learning process was provided as the treatment to the experimental group. It was aimed to improve students' vocabulary mastery. The audio visual aids employed in this study were pictures (in slides and/or printed-out) and audio files (voice-recorded and spoken text). In addition, both audio and visual aids were not regularly used as a combination. However some of teaching learning process just used either audio aids or visual aids.

The treatment program in the control group was done in five meetings. Five meetings in the control group showed the same pattern of teaching. The descriptive text materials and vocabularies presented in text alone.

3.4.5 Post-test

Post-test was administered to 60 students of both experimental and control group in the last two meetings. It was held after the treatment had done, because it was aimed to find out the students' progress in term of their vocabulary mastery after getting the treatment. In other word, the post-test could be the data to help this study in finding out whether there are any differences between the experimental group who was treated with the use of audio visual aids and the control group who was treated with the text alone.

3.4.6 Questionnaire

In the last meeting from the whole eight meetings, the chance was used to administer the questionnaire to students. The questionnaire discussed about vocabulary learning and the use of audio visual aids in the teaching learning process. So, the questionnaire was administered to the experimental group only who had experienced the treatment of the use of audio visual aids in their learning process.

3.5 Data Analysis

The data needed which were collected, were further analyzed to find out the result of this study. The data analysis included the scoring technique, data analysis on instrument tryout, data analysis on pre-test – post-test results, and data analysis on the results of the questionnaire.

3.5.1 Scoring on instruments try out, pre-test and post-test

Since the test was in two different forms of instructions, the way to score each of them needed two kinds of scoring technique. In detail, the scoring technique used for "fill in the blank" question form, are as follow:

S = R

Where: S = Score
R = Right Answer

(Arikunto, 2003)

To score the students' work on the second test with the instruction to describing some pictures given, a writing rubric was applied. The detail of the writing rubric could be seen in Appendix C.

3.5.2 Data Analysis on Instrument Try Out

The aim of analyzing the instrument of try out was to investigate its validity, reliability, difficulty level, and discrimination level. It was conducted since good instrument test was required to measure the proficiency level of English language learners. In this case was students' vocabulary mastery.

3.5.2.1 Validity

To find out the validity of instrument items, this study conducted Pearson' Product Moment Correlation. The calculation process used SPSS v.17.0. Correlation technique is one of techniques that is mostly used until now by researchers to find out the validity of instrument items (Masrun, 1979 cited in Sugiyono 2010: 133).

In interpreting the result, the minimum criterion of a valid instrument was at 0,3 correlation coefficient. "The instrument item which has high and positive correlation with the total score, it shows that the item has high validity too. Usually, the minimum requirement to qualify is if $r = 0,3$" (Masrun, 1979 cited in Sugiyono 2010: 134). So, if the correlation coefficient between item and total score is bigger than 0,3, then the instrument item is stated valid. Whereas, if the correlation coefficient is lower than 0,3, then the instrument item is stated not valid. "all items that reach a minimum correlation coefficient of 0,30 is considered satisfactory on its distinguishing power" (Azwar, 1999 cited in Priyatno, 2010: 90). The following table is the result of validity test on instrument try-out:

Table 3.1

The Result of Validity Test

Item Number	Validity Score	Interpretation
1,2,3,4,5,6,7,8,9,10,11,12,13,14, 15,16,17,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36, 37,38,39,40.	Correlation coefficient > 0,3	Valid
18	Correlation coefficient < 0,3	Not Valid

Based on the table above, 39 items were valid and 1 item was stated not valid. The valid items could be used as research instrument. So, the total items that could be used as the instrument in this study were 39 items. Moreover, the detail results of validity test of those items could be seen in Appendix C.

3.5.2.2 Reliability

A good instrument is instruments which consistently can give data that matched with the reality (Arikunto, 2003: 86). Therefore, a reliability test was required to find out whether the instrument test of this study could provide consistent result of the data or not.

To find out the reliability of the instrument test, this study conducted two kinds of method. It was because the instruments in this study had two kinds of questions type. It was objective and subjective questions. To test the reliability of subjective questions (essay), it could not use the same way as objective questions in reliability test (Arikunto, 2003: 109). Therefore, Spearman-Brown formula was used to test the reliability of objective questions, number 1 to 20 and Cronbach's Alpha was used for subjective questions from number 21 to 40. In calculating the data, SPSS v.17.0 was used in this study. Furthermore, in interpreting the result, the criteria of correlation coefficient interpretation by Sugiyono (2010) was used in this study.

The result of reliability test for objective questions by conducting Spearman-Brown formula in SPSS v.17.0 showed that the instrument has moderate level of reliability. The score was at 0,532. While, the result of reliability test for subjective questions by using Cronbach's Alpha in SPSS v.17.0 showed that the instrument had very high level of reliability. The score was at 0,967. In addition, the detail results of reliability test of instrument items may be seen in Appendix C.

3.5.2.3 Difficulty Index

In order to give an appropriate and acceptable test for students, analyzing the level of difficulty test items was required to be done in this study. The computation of difficulty index test on instrument try out employed Anates v.4. In interpreting the result, the criteria of difficulty index interpretation by Arikunto (2009) was used in this study. The following table is the result of the difficulty index test:

Table 3.2

The Result of Difficulty Index Test

Item Number	Raw Score	Interpretation
3,4,13,15,24,28,31.	$0.00 < P \leq 0.30$	Difficult item
1,2,5,6,7,8,9,10,11,12,14,16,17,21,22,23,25,26,27,29,30,32,33,34,35,36,37,38,39,40.	$0.30 < P \leq 0.70$	Moderate item
19,20.	$0.70 < P \leq 1.00$	Easy item

Based on the table above, 7 items had difficult level of difficulty index, 30 items had moderate level of difficulty index, and 2 items had easy level of difficulty index. In sum, the most items on the instrument of this study were in the level of moderate. For detail, the result of difficulty index test of instrument items could be seen in Appendix C.

3.5.2.4 Discrimination Index

The discrimination index was used to measure the extent of which items that differentiate the students who gained higher scores and those who gained lower scores. To find out the discrimination index of the instrument in this study, Anates v.4 was used to calculate it. Furthermore, in interpreting the result, the criteria of discrimination index

interpretation by Arikunto (2009) was used in this study. The following table is the result of the discrimination index test:

Table 3.3

The Result of Discrimination Index Test

Item Number	Raw Score	Interpretation
-	$0.00 < D \leq 0.20$	Poor
5,7,12,13,14,15,16.	$0.20 < D \leq 0.40$	Moderate
1,3,4,9,19,20	$0.40 < D \leq 0.70$	Good
2,6,8,10,11,17,21,22,23,24,25,26,27,28,29, 30,31,32,33,34,35,36,37,38,39,40.	$0.70 < D \leq 1.00$	Excellent

Based on the table above, there was no item that was marked "poor". While, 7 items were marked "moderate" in discrimination index test, 6 items were stated "good" in discrimination index test, and 26 items were stated "excellent" in discrimination index test. In sum, the most items on the instrument of this study were stated "excellent" in discrimination index based on the criteria of Arikunto (2009). In addition, the detail results of discrimination index test of the instrument could be seen in Appendix C.

3.5.3 Data Analysis on Pre-test and Post-test

In order to find out whether there was significant different between the experimental and control group, the data of pre-test – post-test from both group had to be analyzed. The analyzing processes in this study intended to use t-test formula. However, before performing the t-test formula, the data was ensured to meet the conditions required. The requirements of using t-test are as follow: the data are interval-

ratio scale, the underlying distributions are bell-shaped (normally distributed), the observations are independent, and the variance of the two groups must be homogenous (Jackson, 2009: 234). In addition, the conditions required of using dependent t-test were actually similar to the independent t-test, except the requirement that the observations are independent. The assumption for dependent t-test was the observations were not independent, but the observations were correlated or dependent (Jackson, 2009: 239).

3.5.3.1 Normality Distribution Test

In the aim to find out if the data were normally distributed, Kolmogrov - Smirnov's formula in SPSS v.17.0 was used. In addition, the hypotheses to test are as follow:

Hypothesis in experimental group:

Ho = The scores of experimental group are normally distributed.

H_1 = The scores of experimental group are not normally distributed.

Hypothesis in control group:

Ho = The scores of control group are normally distributed.

H_1 = The scores of control group are not normally distributed.

The interpretation of the result were based on the level of significant 0,05. If the result is higher than 0,05, then H_0 is not rejected or in the other word that the data are normally distributed. Whereas, if the result is lower than 0,05, H_0 is rejected (Priyatno, 2010: 71).

3.5.3.2 Variance Homogeneity Test

The next step to meet the requirement of using t-test formula was to find out whether the data were homogenous or not. Levene's test in SPSS v.17.0 was used to find it out. In addition, the hypothesis to test is as follow:

Ho = The scores of both experimental and control group are homogenous.

H_1 = The scores of both experimental and control group are not homogenous.

The interpretation of the result to find out whether the data were homogenous or not were based on the level of significant 0,05. If the result is higher than significant level 0,05, then H_0 is not rejected. Meanwhile, if the result is lower than 0,05, then H_0 is rejected or in the other word that the data are not homogenous (Priyatno, 2010: 76).

3.5.3.3 t-test Computation

This study was conducted some statistical process of t-test to analyze the data, such as: independent t-test for pre-test, independent t-test for post-test, dependent t-test for experimental group, and dependent t-test for control group. This t-test is a parametric inferential statistical test of the null hypothesis (Jackson, 2009: 184). In addition, null hypothesis or H_0 is the hypothesis that predicting no difference exists between the groups being compared (Jackson, 2009: 166).

3.5.3.3.1 Independent t-test Computation on Pre-test

Independent t-test was conducted to calculate the data from pre-test of both experimental and control group. In detail, the independent t-test was used to compare means of pre-test scores of both experimental and control group. Independent t-test is a

parametric statistical test that compares the means of two different samples of participants (Jackson, 2009: 227). It was aimed to find out whether there was significant different or not between both groups' initial ability before the groups were given the treatments.

In this study, the independent t-test was analyzed by using SPSS v.17.0. The significance level used in independent t-test was 0,05 (two-tailed test of significance). In addition, the hypothesis to test is as follow:

Ho = There is no significant different between the means of pre-test scores of both experimental and control group.

H_a = There is significant different between the means of pre-test scores of both experimental and control group.

The testing criterions are as follow:

Ho is not rejected if $-t_{crit} \leq -t_{obt}$ or $t_{crit} \geq t_{obt}$

Ho is rejected if $-t_{obt} < -t_{crit}$ or $t_{obt} > t_{crit}$

And

Ho is not rejected if significant > 0,05

Ho is rejected if significant < 0,05

(Priyatno, 2010: 36)

3.5.3.3.2 Independent t-test Computation on Post-test

The use of independent t-test in post-test was to compare means of post-test scores of both experimental and control group. It was aimed to find out whether there

was significant difference or not between both groups' ability after the groups experienced the treatments in their learning process.

As in the pre-test analysis, SPSS v.17.0 was also used this computation to analyze the independent t-test. The significance level used was at 0,05 (two-tailed test of significance). The hypothesis to test is as follow:

H_0 = There is no significant different between the means of post-test scores of both experimental and control group.

H_a = There is significant different between the means of post-test scores of both experimental and control group.

The testing criterions are as follow:

Ho is not rejected if $-t_{crit} \leq -t_{obt}$ or $t_{crit} \geq t_{obt}$

Ho is rejected if $-t_{obt} < -t_{crit}$ or $t_{obt} > t_{crit}$

And

Ho is not rejected if significant > 0,05

Ho is rejected if significant < 0,05

(Priyatno, 2010: 36)

3.5.3.3.3 Dependent t-test

The dependent t-test was used to find out whether there was significance different or not between pre-test and post-test. Dependent t-test is a parametric statistical test that compares the means of two related (within or matched-participants) samples (Jackson, 2009: 234).

3.5.3.3.3.1 Dependent t-test on the Experimental Group Score

The dependent t-test on the experimental group was used to find out whether there was significance different or not between pre-test and post-test after the group had experienced the treatment of the use of audio visual aids in this study.

The dependent t-test computation used SPSS v.17.0 in calculated the data. The significance level used was at 0,05 (two-tailed test of significance). In addition, the hypothesis to test is as follow:

Ho = There is no significant different between pre-test and post-test after treatment.

H_a = There is significant different between pre-test and post-test after treatment.

The testing criterions are as follow:

Ho is not rejected if $-t_{crit} \leq -t_{obt}$ or $t_{crit} \geq t_{obt}$

Ho is rejected if $-t_{obt} < -t_{crit}$ or $t_{obt} > t_{crit}$

And

Ho is not rejected if significant > 0,05

Ho is rejected if significant < 0,05

(Priyatno, 2010: 41)

3.5.3.3.3.2 Dependent t-test on the Control Group Score

The dependent t-test on the control group was used to find out whether there was significance different or not between pre-test and post-test after the group had experienced the treatment of the use of text alone in this study.

As in the experimental group, the dependent t-test computation on control group was used SPSS v.17.0 too in calculated the data. The significance level used was at 0,05 (two-tailed test of significance). Moreover, the hypothesis to test is as follow:

Ho=There is no significant different between pre-test and post-test after treatment.

H_a =There is significant different between pre-test and post-test after treatment.

The testing criterions are as follow:

Ho is not rejected if $-t_{crit} \leq -t_{obt}$ or $t_{crit} \geq t_{obt}$

Ho is rejected if $-t_{obt} < -t_{crit}$ or $t_{obt} > t_{crit}$

And

Ho is not rejected if significant > 0,05

Ho is rejected if significant < 0,05

(Priyatno, 2010: 41)

3.5.3.4 Calculation of the Effect Size

In order to measure how well the treatment worked in this study, calculation of the effect size was conducted to measure. Effect size has positive correlation to its value, the larger effect size value, the larger impact of treatment (Coolidge, 2000).

3.5.3.4.1 Calculation of the Effect Size on the Dependent t-test

The computation of effect size should be conducted for the dependent t-test (Jackson, 2009: 238). Cohen's d formula was used to calculate the effect size on dependent t-test. The Cohen's d formula for dependent t-test is as follow:

$$d = \frac{\overline{D}}{S_D}$$

Where: $\overline{D}$ = The mean of the difference scores

S_D = The standard deviation of the difference scores

(Jackson, 2009: 238)

In interpreting the result, the criteria of effect size by Cohen (1992) was adopted. The criteria are as follow: a small effect size is one of at least 0,20, a medium effect size is at least 0,50, and a large effect size is at least 0,80 (Cohen, 1992 cited in Jackson, 2009: 238).

3.5.4 Data Analysis on Questionnaire

In analyzing the students' responses about the vocabulary learning and the use of audio visual aids in their English learning process, the questionnaire was analyzed by using this below percentage formula:

$$P = \frac{f_o}{n} \, x \, 100\%$$

Where: P = Percentage

f_o = Frequency observed

n = Number of sample

(Riduwan, 2009)

In interpreting the result of percentage computation above, the classification percentage by Ali (1985:184, cited in Rahmawati, 2011) was used to classify it.