CHAPTER III

RESEARCH METHODOLOGY

Introduction

This chapter presents the explanation about procedures which are taken in this study in order to find out the answer to the research questions. This chapter includes research design, Population and sample, Research hypothesis, data collection, trying out the research instruments, and data analysis.

3.1 Research Design

In this present research, quantitative approach with correlation method is employed. Quantitative research is used since this research focuses on analyzing the data through systematic process by using certain computation. Nunan (2003) and Arikunto (2003) state that quantitative research is an attempt to investigate an issue by using numerical data and statistical processing.

Correlation method is considered appropriate, since this research concerns on the investigation to find out the correlation between students' vocabulary mastery and their reading comprehension. It is in line with Sudjana (1996) who state that correlation method is a process to find out the relationship between two or more variables and how strong the relationship is. Another statement comes from Hatch & Farhady (1982) explain that correlation is a statistical technique that can show whether and how strong pairs of variables are

related. Therefore in this study, the variables to be correlated are vocabulary mastery and reading comprehension.

3.2 Population and Sample

3.2.1 Population

Arikunto (2003) state that population is the whole subject of the research. The population of the research was the second grade of a junior high school in Bandung. There were 7 classes consisting of 250 students.

3.2.2 Sample

Sample is a part of the investigated population (Arikunto 2003). The subject of this study was the second grade of a junior high school. This selection was based on the reason that those students have learned several types of text and the researcher had taught the second grade in this school. The present study chooses one class randomly; the class consisted of thirty four (34) students. It is in line with Gay, Mills, and Airasian (2006) who suggest that there should be at least 30 (thirty) participants in correlation study to establish a relationship.

3.3 Research Hypothesis

Hypothesis is a prediction about what you expect to happen in the study (Sudjana, 1996). There are two types of hypothesis, alternative hypothesis (H_a) and null hypothesis (H_0).

According to Weaver (2005) Alternative Hypothesis (H_a) is the hypothesis that states that there is a relation between the phenomena under investigation. Null hypothesis (H_0) is the opposite of alternative hypothesis, in order words there is no relation between the phenomena under investigation.

When there is a correlation between students' vocabulary mastery and their reading comprehension, the alternative hypothesis is accepted and null hypothesis is rejected.

3.4 Data Collection

The technique used to collect data in this study is achievement test. Through this technique, the information about students' ability in reading comprehension and vocabulary mastery are expected to be obtained.

3.4.1 Research Instrument

According to Arikunto (2003) research instrument is a tool used by the researcher to find out or to measure ability with certain rules. Achievement test was given to the participants in order to measure the ability of students' vocabulary mastery and their reading comprehension.

The test contains 50 questions, twenty five questions are the questions to measure students' reading comprehension and the other twenty five are to measure students' vocabulary mastery. The test was taken from 2006-2010 UAN test items (UAN questions is a standardized test for Indonesian students). So, it is

reasonable to be an instrument in this research. It was constructed in multiple-choice questions.

The aspects related to reading comprehension are finding the main idea of the text, and getting the information about the text. The aspects related to vocabulary mastery test are finding the synonym or antonym, answering notice, completing the sentence and correcting the spelling.

In scoring the test, the right answer was marked one (1) point and the wrong answer was marked zero (0) point, so the overall raw score from the right answers in this achievement test is 50 points. After marking the test, the present study tried to gain the final scores by using ***S*** formula as below.

$$S = \frac{CA}{N} \times 100$$

Where:

S : final test score

N : number of questions

CA : number of correct answers

(Arikunto, 2003)

The scores were interpreted in order to classify participants' reading ability (Arikunto, 2003). The classifications are presented in Table 3.1

Table 3.1 Classification of students' achievement

Score Range	Classifications
80 – 100	Excellent
66 – 79	Good
56 – 65	Average
30 – 55	Poor
0 – 29	Fail

In collecting the data, this study took several procedures:

a. Preparing Research instrument (achievement test).
b. Trying out the research instruments to the students in order to check its validity, reliability, difficulty index and discrimination index.
c. Giving the achievement test to the participants
d. Scoring the participants' achievement test.
e. Calculating the data by using a median formula and Pearson Product Moment formula.
f. Analyzing the result through the relevant theories and drawing a conclusion of this research.

Those procedures above should be taken carefully one by one to prevent the emergence of mistakes during the research. Furthermore, explanation about the procedures is clearly presented in the next sections.

3.5 Trying Out the Research Instruments

A good test at least possesses two qualities, which are validity and reliability (Arikunto 2003). The questions of the test were collected from UAN

2006 – 2010. In order to get the requirement of a good test, the test was first tried out before it was actually administered. Then its results then were calculated in order to find out its validity and reliability.

Besides measuring validity and reliability of the instrument, difficulty level and discrimination power are also calculated. Difficulty level is about classifying the test into easy or difficult, while discrimination power is calculated to find out the significance of test items (Arikunto 2003).

3.5.1 Validity

Arikunto (2003) state that a test is valid if it measures what it is supposed to measure. In order to find out the validity of the achievement test, the test item was tried out and the result was computed with Pearson Product Moment correlation formula. The formula is.

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

Where:

r : correlation coefficient

x : item which its validity is assessed

y : item which its validity is assessed

N : the number of participants

(Arikunto 2003:72)

The criteria of validity were shown in the table.

Table 3.2 r Coefficient Correlation (Validity)

Raw Score	Interpretation
0.8 – 1.0	Very high
0.6 – 0.8	High
0.4 – 0.6	Moderate
0.2 – 0.4	Low
0.0 – 0.2	Very Low

(Arikunto, 2003)

The result of statistical computation on test shows that there were 22 vocabulary items and 21 reading comprehension items that were valid and could be used as the research instrument. See appendix A.1

Table 3.3 The Result of Vocabulary Validity Test

Item Number	**Raw Score**	**Interpretation**
46, 49, 50	0.000 – 0.200	Very Low
9, 10, 15, 20, 24, 25, 30, 31, 35, 36, 37, 39, 40, 41, 42, 43, 45, 47, 48	0.200 – 0.400	Low
6, 38, 44	0.400 – 0.600	Moderate
-	0.600 – 0.800	High
-	0.800 – 1.000	Very High

Table 3.4 The Result of reading Comprehension Validity Test

Item Number	Raw Score	Interpretation
11, 26, 29, 33	0.000 – 0.200	Very Low
1, 2, 3, 4, 5, 7, 8, 12, 13, 14, 16, 17, 18, 19, 21, 22, 23, 28, 32, 34,	0.200 – 0.400	Low
27	0.400 – 0.600	Moderate
-	0.600 – 0.800	High
-	0.800 – 1.000	Very High

In this correlation research, the items of research instrument must be equal between two variables. So, one item from vocabulary test must not be used. The lowest validity score in vocabulary mastery instrument was discarded.

3.5.2 Reliability

According to Arikunto (2003) reliability is related with consistent and stable indication of test. It means a test can be considered reliable if it is consistent with the result when it is used more than once to the same objects in different time.

In finding the reliability of the test, the Split-half method was used. Split-half method is a method that uses one achievement test and test once. There are several steps in this method. The achievement test is equally separated into two parts, first half and second half. Those data are calculated firstly by using Pearson

Product Moment correlation formula. Then the correlation coefficient from calculation is calculated by using Spearman Brown formula (Arikunto, 2003). The Spearman Brown formula is as the following.

$$r_{11} = \frac{2 \times r_{1/2\,1/2}}{(1 + r_{1/2\,1/2})}$$

Where:

r_{11} : reliability coefficient

$r_{1/2\,1/2}$: correlation coefficient for each half of the test item

(Arikunto, 2003:93)

After calculating the reliability (see appendix A.2), it was found that the reliability coefficient of achievement test which is also called r_{11}, is 0.779. After that, the result of reliability coefficient then should be applied to ***r*** value in Product Moment table (Sugiyono, 2008). The Product Moment table can be seen in Appendix A.3. Then the result should apply the interpretation.

if $r_{obtained} > r_{critical}$ = valid

if $r_{obtained} < r_{critical}$ = invalid

(Sugiyono, 2008)

According to Sugiyono (2008) the ***r critical*** for this instrument is 0.339 (see appendix A.3). Since the **r obtain** exceed the ***r critical,*** it means that the test is reliable.

3.5.3 Difficulty Index

Difficulty index needs to be calculated in order to find out the difficulty level of a test. Arikunto (2003) state that the index of difficulty or facility value of an item illustrates how easy or difficult the certain item established in the test. The value around 0.500 was considered to be ideal with an acceptable range from around 0.3 to 0.7. In addition, the following formula is used to calculate the index of difficulty of an item.

$$P = \frac{B}{JS}$$

P = Facility/ Index of difficulty
B = The number of correct answers
JS = The number of students taking the test

(Arikunto, 2003:208)

After obtaining the result, the classifications of result were applied to the table below.

Table 3.5 Criteria of difficulty Index

Index of Difficulty	Difficulty Degree
0.00 – 0.30	Difficult item
0.31 -0.70	Moderate item
0.71 – 1.00	Easy item

(Arikunto, 2003)

The result of computing shows that 4 items were considered difficult, 20 items were considered moderate and 26 items were considered easy.

Table 3.6 The Difficulty Test

Item Number	Index of Difficulty	Difficulty Degree
1, 12, 37, 48	0.00 – 0.30	Difficult Item
3, 9, 11, 15, 19, 26, 27, 29, 30, 31, 33, 36, 38, 41, 44, 45, 46, 47, 49, 50	0.30 -0.70	Moderate Item
2, 4, 5, 6, 7, 8, 10, 21, 13, 14, 16, 17, 18, 20, 22, 23, 24, 25, 28, 32, 34, 35, 39, 40, 42, 43,	0.70 – 1.00	Easy Item

The result shows that 2 reading items were included into difficult items, 7 reading items were included into moderate items and 16 reading items were included into easy items. The vocabulary item result was different from reading result. 2 vocabulary items were included into difficult, 13 vocabulary items were included into moderate items and 10 items were included into easy items. Detail calculation see appendix A.4.

3.5.4 Discrimination Power Index

Discrimination power index needs to be calculated in order to find out the significance of test items in determining participants' skill (Arikunto 2003).

The present study is able to find the discrimination index by conducting the procedures.

1. Arranging students' total score and dividing the score into two groups of equal size (the top half and the bottom half).

2. Counting the number of the students in the upper group who answer each item correctly, then counting the number of lower group students who answer the item correctly.
3. Subtracting the number of correct answer in the upper group to find the difference in the proportion passing in the upper group and the proportion passing the lower group, and
4. Dividing the difference by the total number of students in one group.

The following formula is used to calculate the discrimination index of an item.

$$D = \frac{B_u}{J_u} - \frac{B_l}{J_l}$$

Where:

D : discrimination power index

B_u : participant in upper group answers right

J_u : participants in upper group

B_l : participant in lower group answers right

J_l : participants in lower group

(Arikunto, 2003)

After obtaining the result of discrimination power index, the classification and recommendation should be applied (Arikunto, 2003), as presented in Table.

Table 3.7 Classifications of Discrimination Power Index

Discrimination Index	interpretation
≤ -0,01	Worst (must be discarded)
00.00 – 0.20	Poor
0.21 – 0.40	Moderate
0.41 – 0.70	Good
0.71 – 1.00	Excellent

By taking those steps above, test result was being sorted, divided, and calculated by using ***D*** formula (see Appendix A.5). After obtaining the result, then the classifications were applied, as presented in Table 3.8 below.

Table 3.8 Discrimination Power Index

Items	*D* Score	Classifications
Q3, Q5, Q6, Q7, Q10, Q13, Q14, Q21, Q25, Q31, Q35, Q36, Q42, Q43	00.00 – 0.20	**Poor**
Q1, Q2, Q4, Q8,Q12, Q16, Q17, Q18, Q19, Q20, Q21, Q23,Q24, Q27, Q28,Q32, Q34, Q37, Q39, Q40, Q44,Q45, Q47, Q48	0.21 – 0.40	**Moderate**
Q15, Q22, Q30, Q38, Q41	0.41 – 0.70	**Good**

The result above shows that 14 items were considered poor, 23 items were considered moderate and 5 items were considered good.

3.6 Data Analysis

The following are the steps of how the data were analyzed. First, after the test was taken from the students, the analysis started by scoring the result of the test. Then, the process was about finding the level of participants' vocabulary mastery and reading comprehension. To find out the mastery of the two variables, computing the mean of each variable was necessary. The formula to compute mean is as written below.

$$Mx = \frac{\sum x}{N}$$

$$My = \frac{\sum y}{N}$$

(Arikunto, 2003)

Where: Mx = Mean x (Vocabulary mastery)

My = Mean y (Reading comprehension)

$\sum$x = The sum of x scores

$\sum$y = The sum of y scores

N = Number of participants

Afterward, it is necessary to make sure that the data were normally distributed or not (Sudjana, 1996). This study utilized SPSS 17 (Statistical Package for Social Sciences) as it is one of the oldest and the most widely-used statistical software package. The equations of *Kolmogorov-Smirnov* and *Saphiro-Wilk* were used to find out the normality distribution.

The result of normality distribution determines the formula which is employed to analyze the data. If the data are normally distributed, then Pearson Product Moment formula is applied, as it is also a correlation formula for

parametric statistic and interval data (Sudjana, 1996). The Pearson Product Moment formula is.

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

Where:

r = Correlation Coefficient

N = Number of participants

X and y = variables (variable x and variable y)

∑ = Sum

(Arikunto 2003, p.72)

Then the result of correlation coefficient was interpreted to find out its strength to follow (Arikunto, 2003). The interpretations are presented in Table.

Table 3.9 Correlation Coefficient Interpretation

Raw Score	Interpretation
0.8 – 1.0	Very strong
0.6 – 0.8	Strong
0.4 – 0.6	Moderate
0.2 – 0.4	Weak
0.0 – 0.2	Very weak

(Arikunto, 2003)

In the other hand, Spearman Correlation for Ranked Data formula is employed if the data are not normally distributed, as it is a correlation formula for non-parametric statistic and ranked data (Sudjana, 1996). The formula is.

$$r_s = 1 - \frac{6\sum D^2}{N(N^2 - 1)}$$

Figure 3.7 Spearman Correlation

Where:

r_s : correlation coefficient N : number of participants

D : the difference between participants' ranks

(Sudjana, 1996)

After determining the correlation coefficient, it is necessary to find out whether the hypothesis is accepted or not. The null and alternative hypotheses for Pearson Product Moment correlation are as follows.

H_o : $\rho = 0$

H_a : $\rho \neq 0$

(Arikunto, 2003)