

III. RESEARCH METHOD

3.1 Research Design

In conducting the research, the writer used the experimental research with *one group pre test and post test design*. In the form of objective test, the researcher gave a pre test before treatments and a post test after the three treatments (Hatch and Farhady, 1982). The sample of the research was one class only and the design can be presented as follow:

(T1 X1 X2 X3 T2)

Note:

T1	: Pre test
T2	: Post test
X1, 2, 3	: Treatments

The treatment in this research divided into three meetings of activities. In each meeting, the length of each treatment was 2 x 45 minutes. The overall treatments had been given about three weeks. In a week, English lesson taught two days in 3 hours. On February 07th, 2011 the researcher started the research by trying out the test and in February 09th, the researcher administered the pre test to measure the students' basic ability before the treatments. The research was continued on the next meeting on February 12th, 14th, and on February 16th, 2011. At the last meeting, on February 18th, 2011 the researcher conducted the post test to know the improvement of students' vocabulary after being taught through flashcard.

3.2 Population and Sample of the Research

The population of this research was the fifth year elementary students of Madrasah Mathla'ul Anwar SG Bandar Lampung in academic year of 2010-2011.

There has been one class as the sample, namely experimental class. The sample of the research has been chosen randomly from two classes, class A and class B by using lottery because the participants have similar chance to be chosen and in order to avoid the subjectivity in the research (Setiyadi, 2006:39). Class V B of the Fifth Year is used as sample of this research.

3.3 Variables

The research consists of the following variables:

1. Students' achievement on vocabulary as dependent variable (Y).
2. Flashcard as independent variable (X).

3.4 Data Collecting Technique

In order to collect the data, the researcher applied some techniques as follows:

1) Pre test

This test gave in order to find out the pre test score of the students and to know the ability of the students in vocabulary before the students being taught

by using flashcard. The test has been in the multiple-choice form with four options (A, B, C, and D) and the sum of the test was about 50 items with 30 multiple choices and 20 in form of matching tests. The time allocation has been 80 minutes. The result of the pre test has been compared with the post test result to find out their achievement. The scoring system was the load of each correct answer is two point five. Therefore, if one participant answers all the items correctly, s/he will get 100 points (2 X 50).

2) Post test

After conducting the teaching or the treatment, a post test had been administered. The post test consisted of 50 items in the form of multiple choices with four options and time allocation was 80 minutes. It has been done in order to know the students' vocabulary achievement after having the treatment. The questions in post test have been the same as the pre test. However, the researcher has changed the questions and the distracters in other number differ than those in pretest. This test has had the same difficulty as pre test. In order to know the student' progress in mastering the vocabularies that have been tested, the writer has computed the students' score by using Statistical package for Social Sciences (SPSS) version 15.0.

3) Conducting the treatments

The researcher has applied three treatments in presenting the materials by using flashcard in teaching vocabulary. The first treatment was about kinds of things at home. The second treatment was about things at school. The third treatment was review of the first and the second treatment. The three steps of treatments have been necessary in order to increase the student's vocabulary.

3.5 Research Instrument

In this research, the researcher has administered three tests: try out test, pre test and post test. Try out test was given to know how the quality of the test which has been used as the instrument of the research. Pre test was given in order to know the students' vocabulary before the treatments. Post test was given in order to know the students' vocabulary achievement after the treatments. The form of the try out test, pretest, and the post test has been multiple-choice and matching tests. The total number of the items of the tryout test has been 50 items, and the total number of the pre test and post test has been 50 items. The try out test has been administered about 80 minutes, and pre test - post test has been administered about 80 minutes.

3.6 Try Out Test

The try out has been needed to be done to prove whether the test had good quality or not. The test was said to have a good quality if it has a good validity, reliability, level of difficulty and discrimination power. The try out was held to different class from the experimental class. There have been some elements that have been tested as follows:

1. Validity

The test can be said valid if the test measured the objective to be measured and suitable with the criteria (Hatch and Farhady, 1982:250). To measure

whether the test had a good validity, the writer has seen the content and the constructs validity.

a. Content validity

Content validity is extended to which a test measures representative sample of the subject matter contents, the focus of the content validity is adequacy of the sample and simply on the appearance of the test (Hatch and Farhady, 1982:251). It means that the test should be correct and represent the materials that have been taught such as concrete and abstract nouns. To get the content validity, the test has been adapted from the students' book. Then, the test has been determined according to the material that has been taught to the students. In other words, the writer has made the test based on the materials in the 2004 English Curriculum for the fifth grade of Elementary School students. Beside that, the writer also makes **a table of specification** to judge the content validity already good or not).

Table 1. Specification that was used to judge the content validity of the vocabulary test concerning concrete nouns

Topic/materi	Aspect to be measured			
	Knowledge (50%)	Comprehension (30%)	Application (20%)	Sum (100%)
Things at Home (50%)	(13 items) 1., 4., 8., 14., 20., 22., 31., 32., 35., 37., 44., 46., 49.,	(7 items) 7., 11., 18., 21., 40., 45., 48.	(5 items) 17., 19., 26., 41., 50.	20
Things at School (50%)	(13 items) 2, 5, 10, 12, 16, 24., 25., 30., 34., 36., 39., 42., 47.,	(7 items) 3., 6., 9., 15., 29., 33., 43.,	(5 items) 13., 23., 27., 28., 38.,	20
Total (100%)	26	14	10	50

(Adopted and developed from Arikunto 2005:196)

b. Construct validity

Construct validity focused on the kind of test that was used to measure the ability. It was used to the research that has many indicators.

According to Setiyadi (2006:26), if the instrument just measures one aspect, for example vocabulary, the construct validity can be measure by evaluate all items in the test. If all items have measured vocabulary mastery, this instrument has fulfilled construct validity. In this research, the all items had measure vocabulary mastery, so it has fulfilled construct validity. It can be seen from the Specification table of Construct validity below. Moreover to decide the construct validity of the test, the researcher has used three raters, they are, teacher, researcher and colleague. This is also called as an inter-rater. By using this method, it was hoped that the items of the test has fulfilled the construct validity.

Table 2. Specification that was used to judge the construct validity of the vocabulary test concerning noun, verb, adjective

No	Aspect to be Measured	Spread of Item	Percentage
1	Vocabularies of Noun: a. Abstract nouns	2., 3., 4., 5., 6., 7., 8., 9., 12., 13., 14., 21., 22., 23., 24., 25., 26., 27., 30., 31., 32., 33., 34., 37., 39., 40., 44. (23 Items)	46 %
	b. Concrete Nouns	1., 11., 16., 17., 18., 20., 29., 35., 38., 45., 49. (11 Items)	22 %
2	Vocabularies of Verb	10., 36., 41., 42., 47. (5 items)	10 %
3	Vocabularies of Adjective	15., 19., 22., 28., 43., 48., 50. (7 Items)	14 %

(Adopted and developed from Arikunto 2005:196)

2. Reliability

Shohamy (1985:70) writes, "That reliability refers to the extent to which is consistent in its score, and it gives one an indication of how accurate the test score are". The concept of reliability stems from the idea that no measurement is perfect, even if one goes on the same scale today and then again tomorrow, there will always be differences in his weight which is a result of the fact that measuring instrument are not perfect. A score of a student on a test consists of a true score and error. Reliability helps us estimate the error part of the scores. Since there are different sources of error for different types of tests, there are also different types of reliability.

To find the reliability of this test the writer used Split Half Method, which is has two steps, they are:

1. **First**, using Pearson Product Moment Correlation, we should find the coefficient correlation between Odd and Even Number of the items.

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

Where:

- r_{xy} : Coefficient of reliability between odd and even number.
- $\sum x^2$: Total Square of x (total score of odd number).
- $\sum y^2$: Total square of y (total score of even number).
- $\sum xy$: Total score of odd and even number items.

2. **Second**, after we get the coefficient Correlation between Odd and even Number, we continue to put them into the Reliability Formula.

The Reliability Formula is below:

$$r_{11} = \frac{2 \times r_{xy}}{1 + r_{xy}}$$

Where:

r_{11} : Reliability Coefficient

r_{xy} : Coefficient between odd and even number.

The criteria of reliability

0.00 – 0.20 : very low.

0.21 – 0.39 : low.

0.40 – 0.59 : average.

0.60 – 0.79 : high.

0.80 – 1.00 : very high.

(Arikunto; 1989:167)

In this research, the result of the reliability was 0.99. it can be concluded that the test has very high reliability in which criteria for very high reliability is in the range 0.80 – 1.00. It indicated that this instrument would produce consistent result when administered under similar condition, to the same participant and in different time (Hatch & Farhady, 1982:286). So, it can be stated that the test has fulfilled the criteria of reliability. In other words, the test was reliable. The calculation of reliability of try out test is shown on Appendix 5.

3. Difficulty Level

In determining the level of difficulty of the test, the writer has used the following formula:

$$LD = \frac{R}{N}$$

Where:

LD : Level of Difficulty.

R : The number of the students who answer correctly.

N : The total number of the students.

The criteria are:

0.00 – 0.30 : difficult.

0.31 – 0.70 : average.

0.71 – 1.00 : easy.

(Arikunto; 1997:121)

Based on the try out test result related to those criteria there were 29 easy items (2., 3., 6., 8., 10., 12., 14., 15., 16., 17., 19., 20., 21., 22., 25., 26., 31., 32., 33., 36., 37., 39., 40., 42., 44., 45., 48., 49., 50.), 15 average items (4., 5., 7., 9., 11., 13., 18., 23., 29., 30., 38., 41., 43., 46., 47.), and 6 difficult items (1., 24., 27., 28., 34., 35.). Some items which were easy and difficult were dropped and replaced, meanwhile for the items which were average in difficulty level used in pre test and post test. The results of try out test difficulty level are shown in Appendix 3.

4. Discrimination level

To estimate the discrimination power of the tests, the writer has used the following formula:

$$DP = \frac{U - L}{\frac{1}{2}N}$$

Where:

DP : Discrimination power.

U : The number of upper group students who answer correctly.

L : The number of lower group students who answer correctly.

N : The total number of the students.

The criteria are:

0.00 – 0.20 : poor.

0.21 – 0.40 : satisfactory.

0.41 – 0.70 : good.

0.71 – 1.00 : excellent.

(Arikunto; 1992:221)

Based on the try out test result related to those criteria there were 23 items (1., 6., 7., 8., 14., 15., 16., 17., 19., 23., 26., 27., 28., 30., 31., 35., 37., 38., 40., 41., 45., 46., and 49.) were poor, 6 items (4., 12., 18., 20, 25., and 47.,) were good, 14 items (2., 5., 9., 10., 13., 21., 22., 29., 32., 33., 39., 43., 48., and 50.) were satisfactory and 7 items (3., 11., 24., 34., 36., 42., and 44.,) were have negative discrimination power.

Negative discrimination items were dropped and replaced, meanwhile for the items good and satisfactory were administered. Some of items, which have poor discrimination power, were revised and used in the pre test and post test. The total items that were administered for pre test and post test were 25 items (2., 4., 5., 6., 9., 10., 12., 13., 18., 20., 21., 22., 29., 31., 32., 33., 37., 38., 39., 43., 45., 46., 47., 48., and 50.).

3.7 Research Procedure

The procedures of this research are as follows:

1. Selecting and determining the population and the sample of the research.

The writer chose two classes of the fifth year at Madrasah Mathla'ul Anwar SG Bandar Lampung, which was divided as experimental class, and tryout class.

2. Trying out the instrument

The instrument, vocabulary test, was tried previously in the tryout class. It was aimed at making sure the tests in this study are valid and reliable. The number of the items was 50 items in form of multiple choices test. The scoring system for each correct answer was two. So, if one participant answered all items correctly, s/he got 100 points (2 x 50). The time allocation for the try out test was 80 minutes. The try out was held to find out the quality of the test and to determine the items that should be revised for the pre test and the post test.

3. Administering Pre test

The purpose in giving pre test was to measure the students' mastery of vocabulary before being taught by using flashcard. So that, the writer able to compare the result with the post test results later. The pre test was an objective test in form of multiple choices with 50 items. In addition, it has been administered for 80 minutes in the experimental class.

4. Arranging the materials to be taught

The writer arranged the material that has been taught based on the 2004 curriculum for the fifth year of Elementary students. The material was about things around us which were divided into two topics, things at school and things at home. The vocabularies taught to the student classified into content words, especially concrete nouns.

5. Implementing the treatments

The experimental class has been taught by using flashcard. While the teaching and learning occurs, the writer observed the whole process.

6. Administering post test

The post test was administered to the experimental class. The result of the test was compared with the pre test to find out whether there was any significant increase from the post test and the pre test result. It was in objective test in form of multiple choices that consist of 50 items. The post test has been administered about 80 minutes.

7. Analyzing the data

The data was analyzed by comparing the average score (mean) of pre test and post test to know whether there was a significant increase in students' vocabulary achievement taught through flashcard.

8. Reporting the result

In reporting the data, the data has been arranged systematically based on the pre test and post test to see whether there was an increase of the students' vocabulary achievement.

3.8 Data Analysis

After collecting the data, the researcher has calculated the student's achievement by:

1. Scoring pre test and post test.
2. Tabulating the result of the test and calculating the mean of the pre test and post test.

To compute the average score or mean of the pre test and post test, the writer used a very simple formula as follows:

$$M = \frac{\sum X}{N}$$

In which:

M : (Mean) Average score

$\sum X$: Total student's score

N : Total number of the students.

The average is total student's score divided by total number of the students.

(Hatch and Farhady, 1982:55)

3. Drawing conclusion from the tabulated result of the pre test and post test administered. The data has been analyzed by using statistical computerization

Repeated Measures t-test of SPSS 15 for Windows i.e. $t = \frac{\bar{x}_1 - \bar{x}_2}{S_D}$ to test

whether the difference between pre test and post test was significant or not, in which the significance was determined by $p < 0,05$ (Hatch and Farhady, 1982:114). Since the data was gained from one group and the research was intended to find out whether there was a significant increase of the students' vocabulary achievement.

3.9 Data Treatment

According to Setiyadi (2006:168-169), using Repeated Measures t-test for hypothesis testing has three basic assumptions, namely;

- The data is interval or ratio
- The data is taken from random sample in a population

- The data is distributed normally.

Therefore, the writer will use these following procedures:

1. Random test

Run test was used to make sure whether the data random or not. The writer used SPSS version 15 to help her. In this case, the writer used the mean as the cut point run t-test. The hypothesis for the random test was formulated as follows:

H_o : The data are random

H_a : The data are not random

In this script, the criteria for the hypothesis are:

H_o is accepted if Sig. $> \alpha$. In this case, the researcher used the level of significance 0,05.

From the result (see Appendix 14), we can see that Sig. $> \alpha$ in all test (pre test and post test) in the experimental class. It proved that H_o was accepted. In other words, all the data were random.

2. Normality test

The writer used normality test to know whether the data in the pre test and the post test are distributed normally. The hypothesis for the normality test was as follows:

H_o : The data are distributed normally.

H_a : The data are not distributed normally.

In this script, the criteria for the hypothesis are:

H_o is accepted if Sig. $> \alpha$. In this case, the researcher used the level of significance 0,05.

From the result (see Appendix 15), we can see that $\text{Sig.} > \alpha$ in all test (pre test and post test) in the experimental class. It proved that H_o was accepted. In other words, all the data were distributed normally.

3.10 Hypothesis Testing

The researcher compared pretest and posttest to know the gain. The researcher did the t-test (Repeated Measured T-test) toward the average score of pretest and posttest. The result of t-test also used to know the significance of treatment effect and to prove whether the proposed hypothesis was accepted or rejected.

H_a : "There is a significant difference between the pre test and post test score of the students' vocabulary achievement concerning noun, verb, and adjective after being taught using flashcard."

The criterion for accepting the hypothesis:

H_o is accepted if t-ratio is higher than t-table

H_o is rejected if t-ratio is lower than t-table

The hypothesis has been analyzed at the significant level of 0,05 which means that the probability of error in the hypothesis is only about 5%.