CHAPTER IV
RESEARCH FINDING AND DISCUSSION

A. Description of The Result Research

To find out the difference between the students who are taught by using picture message and the students who are not taught by using picture message in vocabulary, especially in MTs. Mafatihut Thullab An-Nawawy Surodadi Jepara, the writer did an analysis of quantitative data. The data is obtained by giving test to the experimental class and control class after giving a different learning both classes.

The subjects of this research were divided into three classes. They are experimental class (VIII A), control class (VIII B) and try out class (VIII C). Before items were given to the students, the writer gave tryout test to analyze validity, reliability, difficulty level and also the discrimination power of each item. The writer prepared 25 items as the instrument of the test. Test was given before and after the students follow the learning process that was provided by the writer.

Before the activities are conducted, the writer determined the materials and lesson plan of learning. Learning in the experiment class used picture message, while the control class without used picture message.

After the data were collected, the writer analyzed it. The first analysis data is from the beginning of control class and experimental class that is taken from the pre test value. It is the normality test and homogeneity test. It is used to know that two groups are normal and have same variant. Another analysis data is from the ending of control class and experimental class. It is used to prove the truth of hypothesis that has been planned.

B. The Data Analysis and Test of Hypothesis

1. The Data Analysis

a. The Data Analysis of Try-out

This discussion covers validity, reliability, level of difficulty and discriminating power.

1) Validity of Instrument

As mentioned in chapter III, validity refers to the precise measurement of the test. In this study, item validity is used to know the index validity of the test. To know the validity of instrument, the writer used the Pearson product moment formula to analyze each item.

It is obtained that from 25 test items; there are 20 test items which are valid and 5 test items which are invalid. They are on number 3, 6, 9, 14, 16. They are to invalid with the reason the computation result of their r_{xy} value (the correlation of score each item) is lower than their r_{table} value.

The following is the example of item validity computation for item number 1 and for the other items would use the same formula.

N = 36 $\quad\quad$ $\sum Y = 653$

$\sum XY = 558$ $\quad\quad$ $\sum X^2 = 25$

$\sum X = 25$ $\quad\quad$ $\sum Y^2 = 12489$

$$r_{xy} = \frac{N\sum XY - \sum(X)\sum(Y)}{\sqrt{\left\{N\sum X^2 - \left(\sum X\right)^2\right\}\left\{N\sum Y^2 - \left(\sum Y\right)^2\right\}}}$$

$$r_{xy} = \frac{36(558) - 25(653)}{\sqrt{\left\{36(25) - (25)^2\right\}\left\{36(12489) - (653)^2\right\}}}$$

$$r_{xy} = 0{,}613$$

From the computation above, the result of computing validity of the item number 1 is 0,621. After that, the writer consulted the result to the table of r Product Moment with the number of subject (N) = 36 and significance level 5% it is 0,316. Since the result of the computation is higher than r in table, the index of validity of the item number 1 is considered to be valid.

2) Reliability of Instrument

A good test must be valid and reliable. Besides the index of validity, the writer calculated the reliability of the test using Kuder-Richarson formula 20(K-R 20).

Before computing the reliability, the writer had to compute Varian (S^2) with the formula below:

N = 36 $\sum Y$ =653

$\sum Y^2$ = 12489 $\sum pq$ = 5,492

$$S^2 = \frac{\sum y^2 - \frac{(\sum y)^2}{N}}{N}$$

$$S^2 = \frac{12489 - \frac{(653)^2}{36}}{36}$$

$$S^2 = \frac{12489 - 11845}{36}$$

$$S^2 = \frac{664}{36}$$

S^2 = 17,88

The computation of the Varian (S^2) is 17, 88. After finding the Varian (S^2) the writer computed the reliability of the test as follows:

Formula:

$$r_{11} = \left(\frac{n}{n-1}\right)\left(\frac{S - \sum pq}{S^2}\right)$$

$$r_{11} = \left(\frac{25}{25-1}\right)\left(\frac{17.88 - 6.49}{17.88}\right)$$

$$r_{11} = 1{,}04\left(\frac{11.39}{17.88}\right)$$

= 0.663

The result shows that 0.663 is more than 0,329, it meant that the items of instrument were reliable.

3) The level of Difficulty

The following is the computation of the level difficulty for item number 1 and for the other items would use the same formula.

B=18+11=29

JS= 36

$$P = \frac{B}{JS} \qquad P = \frac{29}{36}$$

$$P = 0{,}81$$

It is proper to say that the index difficulty of the item number 1 above can be said as the easy category, because the calculation result of the item number 1 is in the interval $0{,}70 \leq p \leq 1{,}00$.

After computing 25 items of the try-out test, there are 20 items are considered to be easy, 5 items are enough. The whole computation result of difficulty level can be seen in appendix 6.

4) The Discriminating Power

The discrimination power of an item indicated the extent to which the item discriminated between the testees, separating the more able testees from the less able. The index of discriminating power told us whether those students who performed well on the whole test

tended to do well or badly on each item in the test. To do this analysis, the number of try-out subjects was divided into two groups, upper and lower groups.

The following is the computation of the discriminating power for item number 1, and for other items would use the same formula.

BA=18 BB=11

JA=18 JB=18

$$D = \frac{BA}{JA} - \frac{BB}{JB}$$

$$D = \frac{18}{18} - \frac{11}{18}$$

D = 0,38

According to the criteria, the item number 1 above is good category, because the calculation result of the item number 1 is in the interval $0{,}40 \leq D \leq 0{,}70$.

After computing 25 items of try –out test, there are 1 item is considered to be good, 14 items are good, 6 items are enough, and 5 items are poor. The result of the discriminating power of each item could be seen appendix 6.

Based on the analysis of validity, reliability, difficulty level, and discriminating power, finally 20 items are accepted. They are number 1 2 4 7 8 10 11 12 13 15 17 18 19 20 21 22 23 24 25.

2. Second Analysis

The second analysis represents the result of pre-test and post-test that was done both in experimental and control group. This analysis will answer the research question “How are picture message effective to improve students’ vocabulary in concrete noun? We can conclude picture message is effective when the result of post test of the experimental class (using picture message

technique) and control class (using conventional technique) has significant differences or the assumption that those classes is equal is not fulfilled.

Before the researcher tested the hypothesis that had been mentioned in the chapter two, the researcher analyzed and tested hypothesis prerequisites which contained of normality test and homogeneity test. Second analysis dealt with normality test, homogeneity test, and t-test (test of difference two variants) in pre-test and post-test.

a. Analysis of Pre-test

The experimental group (Class VIII A) was given a pre-test on 24 February, 2011 and control group (Class VIII B) was given a pre-test on 25 February, 2011.

1. Test of Normality

Test of normality was used to find out whether data of control and experimental group which had been collected from the research come from normal distribution normal or not. The result computation of Chi-square (X^2_{score}) then was compared with table of Chi-Square (X^2_{table}) by using 5% alpha of significance. If $X^2_{score} < X^2_{table}$ meant that the data spread of research result distributed normally.

Based on the result of VIII A students in experimental group, before they were taught vocabulary of concrete noun by using picture message, was found that the maximum score was 75 and minimal score was 45. The stretches of score were 30. So, there were 7 classes with length of classes 5. From the computation of frequency distribution, it was found ($\Sigma f_{i.} x_i$) = 2535 and ($\Sigma f_{i.} x_i^{\,2}$) = 164405. So, the average score ($\overline{X}$) was 63.375 and the standard deviation (S) was 9.80499. After

counting the average score and standard deviation, table of observation frequency was needed to measure Chi-Square (X^2_{score}).

Table 2. Table of the Observation Frequency of Class VIII A

Class	Bk	Z_i	$P(Z_i)$	Ld	Ei	Oi	$\frac{(O_i - E_i)^2}{E_i}$
	44.5	-1.93	-0.4729				
45 – 49				0.0514	2.1	4	1.8372
	49.5	-1.42	-0.4215				
50 – 54				0.1042	4.2	6	0.8064
	54.5	-0.91	-0.3173				
55 – 59				0.1637	6.5	4	0.9903
	59.5	-0.40	-0.1537				
60 – 64				0.1993	8.0	6	0.4883
	64.5	0.11	0.0457				
65– 69				0.1882	7.5	7	0.0372
	69.5	0.62	0.2339				
70 – 74				0.1378	5.5	7	0.4011
	74.5	1.13	0.3717				
75 – 79				0.0782	3.1295	6	2.6330
	79.5	1.64	0.4500				
					X^2	=	4.5606

Based on the Chi-Square table (X^2_{table}) for 5% alpha of significance with df 7– 3 = 4, it was found X^2_{table} = 9.49. Because of $X^2_{score} < X^2_{table}$, so the initial data of control group distributed normally.

While from the research result of VIII B students in the control group before they were taught vocabulary without song lyrics they reached the maximum score 80 and minimum score 50. The stretches of score were 30. So, there were 7 classes with length of classes 5. From the computation of frequency distribution, it was found ($\Sigma f_{i.} x_i$) = 2615, and ($\Sigma f_{i.} x_i^2$) = 174225. So, the average score ($\overline{X}$) was 65.375 and the standard deviation (S) was 9.15588. After counting the average score and

standard deviation, table of observation frequency was needed to measure Chi-Square (X^2_{score}).

Table 3. Table of the Observation Frequency of Class VIII B

Class	Bk	Z_i	$P(Z_i)$	Ld	Ei	Oi	$\frac{(O_i - E_i)^2}{E_i}$
	49.5	-1.73	-0.4585				
50 – 54				0.0760	3.0	6	2.8830
	54.5	-1.19	-0.3825				
55 – 59				0.1431	5.7	5	0.0914
	59.5	-0.64	-0.2395				
60 – 64				0.2014	8.1	8	0.0004
	64.5	-0.10	-0.0381				
65– 69				0.2119	8.5	9	0.0324
	69.5	0.45	0.1738				
70 – 74				0.1667	6.7	5	0.4171
	74.5	1.00	0.3405				
75 – 79				0.0980	3.9	3	0.2163
	79.5	1.54	0.4386				
80 – 84				0.0431	1.72348	4	3.0070
	84.5	2.09	0.4816				
					X^2	=	6.6476

Based on the Chi-Square table (X^2_{table}) for 5% alpha of significance with df 7 – 3 = 4, it was found X^2_{table} = 9.49. Because of $X^2_{score} < X^2_{table}$, so the initial data of control group distributed normally.

2. Test of Homogeneity

Test of homogeneity was done to know whether sample in the research come from population that had same variance or not. In this study, the homogeneity of the test was measured by comparing the obtained score (F_{score}) with F_{table}. Thus, if the obtained score (F_{score}) was lower than the F_{table} or equal, it could be said that the Ho was accepted. It meant that the variance was homogeneous.

Table 4. Test of Homogeneity

Variant Sources	Control G	Experimental G
Sum	2535.0	2455.0
N	40	40
$\bar{X}$	63.38	61.38
Variants (s2)	83.8301	96.1378
Standart deviation (s)	9.16	9.80

By knowing the mean and the variance, the writer was able to test the similarity of the two variants in the pre-test between experimental and control group. The formula of the test of homogeneity as follows:

$$F = \frac{Biggest\ Variance}{Smallest\ Variance}$$

$$= 96.1378/83.8301$$

$$= 1.147$$

On a 5% with df numerator (nb - 1) = 40 – 1 = 39 and df denominator (nk – 1) = 40 – 1 = 39, it was found F_{table} = 1.70. Because of $F_{score} \leq F_{table}$, so it could be concluded that both experimental and control group had no differences. The result showed both groups had similar variants (homogenous).

3. Test of Difference Two Variants in Pre-test Between Experiment and Control Group

After counting standard deviation and variance, it could be concluded that both group have no differences in the test of similarity between two variances in pre-test score. So, to differentiate whether the students' results of vocabulary in experimental and control group were significant or not, the writer used t-test to test the hypothesis that had been mentioned in the chapter two. The writer used formula:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Where:

$$S = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Based on table IV. First the writer had to find out S by using the formula above:

$$S = \sqrt{\frac{(40-1)\ 96.1378 + (40-1)83.8301}{40 + 40 - 2}}$$

$$= 9.48599$$

After S was found, the next step was to measure t-test:

$$t = \frac{61.38 - 63.38}{9.48599\sqrt{\frac{1}{40} + \frac{1}{40}}}$$

$$= -0.943$$

After getting t-test result, then it would be consulted to the critical score of t_{table} to check whether the difference is significant or not. For a = 5% with df 40 + 40 – 2 = 78, it was found $t_{table(0.975)(78)}$ = 1.99 Because of $t_{score} < t_{table}$, so it could be concluded that there was no significance of difference between the experimental and control group. It meant that both experimental and control group had same condition before getting treatments.

b. Analysis of Post-test

The experimental group was given post test on March 12, 2011 and control group was given a post test on March 14, 2011. Post-test was conducted after all treatments were done. Picture message were used as media in the teaching of vocabulary to students in experimental group. While for students in control group, they were given treatments without picture message. Post-test was aimed to measure students' ability after they got treatments.

1. Test of Normality

Test of normality was used to find out whether data of control and experimental group, which had been collected after they got treatments, come from normal distribution normal or not. The formula, that was used, was Chi-Square. The result computation of Chi-Square (X^2_{score}) then was compared with table of Chi-square (X^2_{table}) by using 5% alpha of significance. If $X^2_{score} < X^2_{table}$ meant that the data spread of research result distributed normally.

Based on the result of VIII A students in experimental group, after they were taught vocabulary of concrete noun by picture message, was found that the maximum score was 95 and minimal score was 55. The stretches of score were 40. So, there were 7 classes with length of classes 7. From the computation of frequency distribution, it was found ($\Sigma f_{i.} x_i$) = 3008 and ($\Sigma f_{i.} x_i^2$) = 231094 So, the average score ($\overline{X}$) was 75.2and the standard deviation (S) was 11.2003. After counting the average score and standard deviation, table of observation frequency was needed to measure Chi-Square (X^2_{score}).

Table 5. Table of the Observation Frequency of Class VIII A

Class	Bk	Z_i	$P(Z_i)$	Ld	Ei	Oi	$\frac{(O_i - E_i)^2}{E_i}$
	54.5	-1.85	-0.4677				
55 – 60				0.0624	2.5	3	0.1019
	60.5	-1.31	-0.4053				
61 – 66				0.1240	5.0	9	3.2937
	66.5	-0.78	-0.2814				
67 – 72				0.1861	7.4	6	0.2802
	72.5	-0.24	-0.0952				
73 – 78				0.2111	8.4	5	1.4050
	78.5	0.29	0.1159				
79– 84				0.1810	7.2	7	0.0079
	84.5	0.83	0.2968				
85 – 90				0.1172	4.7	6	0.3669
	90.5	1.37	0.4140				
91 – 96				0.0574	2.2944	4	1.2678
	96.5	1.90	0.4714				
					X^2	=	5.4555

Based on the Chi-Square table (X^2_{table}) for 5% alpha of significance with df 7– 3 = 4, it was found X^2_{table} = 9.49. Because of $X^2_{score} < X^2_{table}$, so the initial data of control group distributed normally.

While from the result of VIII B students in the control group after they got usual treatments, they reached the maximum score 90 and minimum score 60. The stretches of score were 30. So, there were 7 classes with length of classes 7. From the computation of frequency distribution, it was found ($\Sigma f_i x_i$) = 2935, and ($\Sigma f_i x_i^2$) = 217805. So, the average score ($\overline{X}$) was 73.375 and the standard deviation (S) was 7.92493. It meant that there was an improvement of students' score after they got treatments. After counting the average score and standard

deviation, table of observation frequency was needed to measure Chi-Square (X^2_{score}).

Table 6. Table of the Observation Frequency of Class VIII B

Class	Bk	Z_i	$P(Z_i)$	Ld	Ei	Oi	$\frac{(O_i - E_i)^2}{E_i}$
	59.5	-1.75	-0.4600				
60 – 64				0.0914	3.7	6	1.5034
	64.5	-1.12	-0.3686				
65 – 69				0.1811	7.2	8	0.0793
	69.5	-0.49	-0.1876				
70 – 74				0.2440	9.8	9	0.0592
	74.5	0.14	0.0564				
75 – 79				0.2238	9.0	7	0.4250
	79.5	0.77	0.2802				
80 – 84				0.1396	5.6	7	0.3589
	84.5	1.40	0.4198				
85 – 89				0.0592	2.4	2	0.0578
	89.5	2.03	0.4791				
90 – 94				0.0171	0.6839	1	0.1461
	94.5	2.67	0.4962				
					X^2	=	2.4836
				0.2238	9.0	7	0.4250

Based on the Chi-Square table (X^2_{table}) for 5% alpha of significance with df 7– 3 = 4, it was found X^2_{table} = 9.49. Because of $X^2_{score} < X^2_{table}$, so the initial data of control group distributed normally.

2. Test of Homogeneity

The writer determined the mean and variance of the students' score either in experimental or control group. By knowing the mean and variance, the writer was able to test the similarity of the two variance in the post-test between experimental and control group.

Table 7. Test of Homogeneity

Varians Sources	Control G	Experimental G
Sum	2855.0	3010.0.
N	40	40
$\bar{X}$	71.38	75.25
Variants (S^2)	62.8045	112.7564
Standart deviation (S)	7.92	10.62

The formula of the test of homogeneity as follows:

$$F = \frac{Biggest\ Variance}{Smallest\ Variance}$$

$$= 112.7564/62.8045$$

$$= 1.795$$

On a 5% with df numerator (nb - 1) = 40 – 1 = 39 and df denominator (nk – 1) = 40 – 1 = 39, it was found $F_{table\ (0.025)(39:39)} = 1.89$ Because of $F_{score} \leq F_{table}$, so it could be concluded that both experimental and control group had no differences. The result showed both groups had similar variance (homogenous).

3. Test of Difference Two Variants in Post-test Between Experiment and Control Group

After counting standard deviation and variance, it could be concluded that both group have no differences in the test of similarity between two variances in post-test score. So, to differentiate if the students' results of vocabulary in experimental and control group after getting treatments were significant or not, the writer used t-test to test the hypothesis that had been mentioned in the chapter two. To see the

difference between the experimental and control group, the writer used formula:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Where:

$$S = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Based on table VI, first the writer had to find out S by using the formula above.

$$S = \sqrt{\frac{(40-1)112.7564 + (40-1)62.8045}{40+40-2}}$$

$$= 9.36912$$

After S was found, the next step was to measure t-test:

$$t = \frac{75.25 - 71.38}{9.36912\sqrt{\frac{1}{40} + \frac{1}{40}}}$$

$$= 1.850$$

After getting t-test result, then it would be consulted to the critical score of t_{table} to check whether the difference is significant or not. For a = 5% with df 40 + 40 – 2 = 78, it was found $t_{table(0.95)(78)}$ = 1.66. Because of $t_{score} > t_{table}$, so it could be concluded that there was significance of difference between the experimental and control group. It meant that experimental group was better that control group after getting treatments.

Since the obtained t-score was higher than the critical score on the table, the difference was statistically significance. Therefore, based on the computation there was a significance difference in vocabulary of concrete

noun achievement score between students were taught using picture message and those were taught without using picture message for the eighth grade students of MTs. Mafatihut Thullab An-Nawawy Surodadi Jepara. So it can be said that using picture message is effective to improve students' vocabulary of concrete noun, and so the action hypothesis is accepted.

C. Discussion of Research Finding

The result of the research shows that the experimental class (the students who are taught using picture message) has the mean value 75.25. Meanwhile, the control class (the students who are taught without using picture message) has the mean value 71.38. It can be said that the vocabulary achievement score of experiment class is higher than the control class.

On the other hand, the test of hypothesis using t-test formula shows the value of the t-test is higher than the critical value. The value of t-test is 1.850, while the critical value on $t_{s0,05}$ is 1,66. It means that there is a significant difference the vocabulary in noun score between students taught picture message and those taught without using picture message. In this case, the use of picture message is necessary needed in teaching vocabulary.

Picture message have some positive influences for the students in improving vocabulary. There are some reasons why the students can improve their vocabulary by using picture message. They are as follows:

1. By using picture message, the students will have encouragement and curiosity to find out the meaning of unfamiliar words. It is caused by the plot; character and setting that are presented in picture message that incite the reader imagination. The students will try to look up the dictionary or guessing the meaning.

2. The use of picture in young learners' classrooms would seem to offer similar rich of opportunities for learning vocabulary from context indirectly. So, students not only understand the meaning of vocabulary, but also they can use it.
3. The teaching of vocabulary using picture message can give opportunities for students to study grammar indirectly.
4. By using picture message, the students can learn vocabulary relaxes and enjoy.

In contrast, not all students have good English vocabulary. Those are caused by some factors that influence the students in learning English. They are as follows:

1. The perception that English is the difficult lesson in school.
2. A poor motivation from the students to learn English seriously
3. The difficulties in memorizing the new words influenced by the culture, pronunciation and grammar.
4. There is no big willingness to learn English

In this research, the writer used the picture message to improve the students' vocabulary in noun at MTs. Mafatihut Thullab An-Nawawy Surodadi Jepara. So, the research findings are only representative in that school. The writer hopes that more researches will be done by the others to prove this method in improving students' vocabulary and to find out other methods in learning and teaching English.

D. Limitation of the Research

The researcher realizes that this research had not been done optimally. There were obstacles faced during the research process.

Some limitations of this research are:

1. The researcher's ability

   The researcher realizes that the implementation of the research process was less smooth; this was more due to lack of the researcher's experience and knowledge.

2. Limitation of time

   Based on the regulation of Tarbiyah Faculty, the research must be done 21 days. So, the relative short time made this research could not be done maximally.

3. Limitation of application

   In this research, the researcher only gave three times treatment to the experiment class, so the result of the research was not maximal.

4. Limitation of the design

   In this research, the researcher used short design. So the research can not be done maximally.

   Considering all those limitations, there is a need to do more research about teaching simple past by using matching game technique so that the more optimal result will be gained.