CHAPTER IV
RESEARCH FINDINGS AND DISCUSSIONS

A. Descriptions of Research Findings

To find out the effectiveness of using pictures to improve listening comprehension skill of narrative texts on the students' achievement in class VIII SMPN 31 Semarang, the writer did an analysis of quantitative data. After conducting the research, she got the data of research finding that is obtained by using the test of experimental class and control class after conducting different treatment of learning process in both classes.

The implementation of this study was divided in two classes, namely the experimental class (VIII D) and the control class (VIII F). Before the activities were conducted, the writer determine the materials and lesson plan of learning. Learning in the experimental class was conducted by using pictures as the media, while in control class using the conventional learning (without using pictures as media).

Test was given before and after the students follow the learning process that was provided by the writer. After the data were collected, the writer analyzed them to prove the truth of the hypothesis that had been formulated. However, before the analysis was done, first the writer scored the results of the test that had been given to the students.

Before items were given to the students, the researcher gave tryout test for try-out class on 7th April 2012 to analyze validity, reliability, difficulty level and also the discrimination power of each item. The researcher prepared 25 items as the instrument of the test. Test was given to know the validity, reliability, degree of test difficulty, and discriminating power of test items of try-out test in control class that was provided by the researcher.

In this research finding of try out test, the researcher used r-*product* formula to analyze validity. The researcher applied the *spearman-brown* formula which was combined with *product- moment* formula to analyze

reliability instrument. The degree of test difficulty used difficulty level formula by considered five levels of difficulty. The last analysis of try-out test was discriminating power by divided into two groups; lower group and upper group.

The researcher gave pre-test on 24th April 2012 in control group and 14th April 2012 in experimental students. The questions consisted of 20 items were stated valid according to try-out analysis. After giving pre-test, the researcher determined the materials and lesson plans of learning activities. Pre-test conducted to both groups to know that two groups were normal and homogeny.

After knowing the control class and experimental class had same variant. The researcher conducted treatment in experimental class twice in week for 40 minutes each meeting. The first treatment conducted on 16th April 2012 and the second treatment conducted on 22nd April 2012 by using pictures to teach listening comprehension skill of narrative texts.

The control class was not taught using pictures; just explaining the material orally based on the teacher's lesson plan without gave variation in learning process. The teacher also asked students just to do the assignment until they felt bored in the class. The teaching also conducted twice a week on 17th April 2012 and 20th April 2012 for 40 minutes for each meeting.

The evaluation of the research found some obstacles in teaching and learning process in control class. The first was the experimental research conducted when the English teacher can not present in the class, so the students felt bad mood to build the better atmosphere because they had not recognized the researcher yet. Moreover, the students did not concentrate into the material because they regard that researcher was not their teacher. Students in experimental class also felt bored in beginning of teaching and learning atmosphere, but they got a great potential to build creativity and could accept materials of the lessons easily in warm atmosphere of the classroom using pictures as the media in teaching and learning process.

From the different situation, the researcher evaluated that the researcher should be humorist to recognize students personally. The teacher also had to know the names each student and they will do the teacher's instruction if the teacher points them. This evaluation was done in the second meeting of teaching in control class and giving treatment in experimental class and could be as reference on the other occasion of the future teaching.

After the researcher gave treatments in experimental class and conventional teaching in control class, the researcher gave post-test which consisted 20 test items which approximately finished on 30 minutes. Giving post test on 22th April 2012 both experimental class and control class.

To analyze the data of test result, the first known the beginning of data from experimental class and control class that is taken from the pre-test value. And after the control and experimental conduct the learning process, then both of the class is given a test to obtain the data that will be analyzed.

B. The Data Analysis and Hypothesis Test

1. The Data Analysis

a. The Data Analysis of Try-out Test

This discussion covered 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 was used to know the index validity of the test. To know the validity of instrument, the researcher used the Pearson product moment formula to analyze each item.

It was obtained that from 25 test items; there were 20 test items which were valid and 5 test items which were invalid. They were on number 5, 6, 17, 22 and 25. They were invalid with the reason the computation result of their r_{xy} value (the correlation of score each item) was lower than their r_{table} value.

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

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

N = 30 $\sum Y = 413$

$\sum XY = 313$ $\sum X^2 = 20$

$\sum X = 20$ $\sum Y^2 = 6449$

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

$$r_{xy} = \frac{30(313) - 20(413)}{\sqrt{\left\{30(20) - (20)^2\right\}\left\{30(6449) - (413)^2\right\}}}$$

$$r_{xy} = 0{,}3321$$

From the computation above, the result of computing validity of the item number 1 was 0.3321. After that, the researcher consulted the result to the table of r Product Moment with the number of subject (N) =30 and significance level 5% it was 0.361. Since the result of the computation was higher than r in table, the index of validity of the item number 1 was considered to be valid. The list of the validity of each item can be seen in appendix 9.

Before computing the reliability, the researcher had to compute *product moment* formula (r_{xy}) with the formula below:

N = 30 $\sum XY = 313$

$\sum Y = 413$ $\sum X^2 = 20$

$\sum Y^2 = 6449$ $\sum X = 20$

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

$$r_{xy} = \frac{30(313) - (20)(413)}{\sqrt{\left\{30(20) - (20)^2\right\}\left\{30(6449) - (413)^2\right\}}}$$

$$r_{xy} = 0{,}3321$$

2. Reliability of Instrument

A good test must be valid and reliable. To get the coefficient of correlation, the researcher applied the *product-moment* formula and then continued to the *spearman-brown* formula. The formula of product moment as follow:

Before computing the reliability, the researcher had to compute *product moment* formula (r_{xy}) with the formula below:

N = 30 $\sum XY = 1611$

$\sum Y = 195$ $\sum X^2 = 1925$

$\sum Y^2 = 1455$ $\sum X = 227$

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

$$r_{xy} = \frac{30(1611) - (227)(195)}{\sqrt{\left\{30(1925) - (227)^2\right\}\left\{30(1455) - (195)^2\right\}}}$$

$$r_{xy} = 0{,}6872$$

After finding *product moment* formula (r_{XY}) the computation was continued to the *spearman-brown* formula as follow:

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

$$r_{11} = \frac{2x0{,}687}{\sqrt{1 + 0{,}69}}$$

$r_{11} = 0{,}815$

From the computation above, it was found out that r_{11} (the total of reliability test) was 0.815 whereas the number of subjects was 30 and the critical value for r-table with significance level 5% was 0.361. Thus, the value resulted from the computation was higher than its critical value. It could be concluded that the instrument used in this research was reliable.

3. The level of Difficulty

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

R = 13+7

N = 30

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

$$FR = \frac{20}{30}$$

$$FR = 0{,}67$$

It was proper to say that the index difficulty of the item number 1 above can be said as the medium category, because the calculation result of the item number 1 was in the interval 0.70 $\leq FR \leq 1{,}00$

4. The Discriminating Power

The discrimination power of an item indicated the extent to which the item discriminated between the tested, separating the more able tested 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. They were upper and lower group.

Table 3

The Table of Discriminating Power of Item Number 1

Upper Group			Lower Group		
No	Code	Score	No	Code	Score
1	T-30	1	1	T-20	1
2	T-03	1	2	T-21	0
3	T-09	1	3	T-23	1
4	T-17	1	4	T-26	1
5	T-22	1	5	T- 12	0
6	T-02	0	6	T- 28	1
7	T-05	1	7	T-06	1
8	T-07	1	8	T-27	0
9	T-11	1	9	T-29	0
10	T-14	1	10	T-13	1
11	T-19	1	11	T-08	1
12	T-04	0	12	T-15	1
13	T-18	1	13	T-16	0
14	T-01	1	14	T-25	0
15	T-10	1	15	T-24	0
Jumlah		13	Jumlah		7

T : Try Out Student

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

This was the analysis of discriminating power for item number 1:

n =15

U=13

L=7

$$D = \frac{Correct\ U - Correct\ L}{n}$$

$$D = \frac{13 - 7}{15}$$

D = 0,40

According to the criteria, the item number 1 above was medium category, because the calculation result of the item number 1 was in the interval $0.20 \leq D \leq 0.40$.

The result of the discriminating power of each item could be seen in appendix 7.

Based on the analysis of validity, reliability, difficulty level, and discriminating power, finally 20 items of test.

The data in this study were gotten from the test result, as follow:

a. The data of score pre-test of the experimental class

Based on the result of research in class VIII D before being taught by using pictures in listening skill of narrative texts the highest score achieved is 70, the lowest is 45, the range (R) is 25, the number of class (K) is 6, and the class interval is 4,2 , so the mean $(\bar{x})$ = 56,8 with standard deviation (S) = 6,2. The result of the calculation above is, then inputted into the table of frequency distribution as follow:

Table 4

List of Frequency Distribution Score of Pre-test of the Experiment Class

No.	**Interval**	Absolute frequency
1	45– 49	1
2	50 – 54	8
3	55 – 59	7
4	60 – 64	8
5	65 – 69	5
6	70 – 74	1

b. The data of score pre-test of the control class

Based on the result of research in class VIII F before being taught by using conventional learning (without pictures as media) in listening comprehension skill of narrative text the highest score achieved is 70, the lowest score is 45, range (R) = 25, the number of class (K) = 6, and

the class interval is 4,2, so the mean $(\bar{x})$ = 56,7 with standard deviation (S) = 7,0. The result of the calculation above is, then inputted into the table of frequency distribution as follow:

Table 5

List of Frequency Distribution Score of Pre-test of the Control Class

No.	**Interval**	Absolute frequency
1	45– 49	2
2	50 – 54	8
3	55 – 59	7
4	60 – 64	6
5	65 – 69	5
6	70 – 74	2

c. The data of score post-test of the experimental class

Based on the result of research in class VIII D after being taught by using pictures as media in teaching listening skill of narrative texts the highest score achieved is 90, the lowest score is 50, range (R) = 40, the number of class (K) = 6, and the class interval = 6,7, so the mean $(\bar{x})$ = 68,0 with standard deviation (S) = 9,9. The result of the calculation above is, then inputted into the table of frequency distribution as follow:

Table 6

List of Frequency Distribution Score of Post-test of the Experiment Class

No	Interval	Absolute Frequency
1	50 – 56	4
2	57 – 63	5

3	64 – 70	11
4	71 – 77	5
5	78 – 84	2
6	85 – 91	3

d. The data of score post-test of the control class.

Based on the result of research in class VIII F after being taught by using conventional learning (without pictures as media) in listening comprehension skill of narrative texts the highest score achieved is 75, the lowest score is 50, range (R)= 25, the number of class (K) = 6, and the class interval is 4,2, so the mean $(\bar{x})$ = 62,3 with standard deviation (S) = 6,0. The result of the calculation above is, then inputted into the table of frequency distribution as follow:

Table 7

List of Frequency Distribution Score of Post-test of the Control Class

No	Interval	Absolute Frequency
1	50 – 54	1
2	55 – 59	5
3	60 – 64	10
4	65 – 69	9
5	70 – 74	3
6	75 – 79	2

e. The average score of pre test and post test of the experimental class and control class.

The data were obtained from the students' achievement scores of the listening comprehension skill of narrative texts. They were pre test and post test scores from the experimental and control classes. The average score from the experimental class was 56.83 for the pre test and

68.00 for the post test. While the average score for the control class was 56.67 for the pre test and 62.33 for the post test. The following was the simple table for the pre test and post test students' average scores:

Table 8

The Result Average Score of the Pre-test and Post-test of the Experimental and Control Class

Class	The average score of the pre test	The average score of the post test
Experimental	56.83	68.00
Control	56.67	62.33

The more calculation can be seen in appendix 3 dan 4

Based on the table above, it can be seen that there was an improvement of the students' achievement in listening comprehension skill on narrative text. Each class had different achievement. The achievement of the experimental class was higher than the control class**.**

1. Analysis Phase First

It was done to know the normality and homogeneity of the initial data in the experimental class and control class.

Table 9

Score of pre-test experimental and control class

No	**Source of variance**	**Experimental**	**Control**
1	N	30	30
2	Average	56.83	56.67
3	Variance	38.77	48.85
4	Standard deviation	6.23	6.99
5	Maximal score	70	70
6	Minimal score	45	45

The more calculations can be seen in appendix 3 dan 4

a) Normality Test of Pre-test

The normality test is used to know whether the data is normally distributed or not. To find out the distribution data is used normality test with Chi-square.

Ho : the data of normal distribution

Ha : the data of un normal distribution

With criteria, Ho accepted if $x^2_{count} < x^2_{table}$ with α = 5% and df = k-3.

Table 10

The result of normality test of experimental and control class pre-test

No	Class	Test	x^2_{count}	x^2_{table}	Criteria
1	Experiment	Pre-test	4,617	7.81	Normal
2	Control	Pre-test	4,083	7.81	Normal

The more calculations can be seen in appendix 10 dan 11

Based on the analysis above it can be seen that x^2_{count} both of class is lower than x^2_{table} ($x^2_{count} < x^2_{table}$), so Ho accepted. It can be concluded that the distribution data of experiment and control class are normal.

b) Homogeneity Test of Pre-test

The homogeneity test is used to know whether the group sample that was taken from population is homogeneous or not.

Ho = $\sigma_1^2 = \sigma_2^2$ (homogeny variance)

Ha = $\sigma_1^2 \neq \sigma_2^2$ (non homogeny variance)

With criteria, Ho accepted if $F_{count} < F_{table}$ with $\alpha = 0.05$ and df = k-1.

Table 11

The result of homogeneity test of experimental and control class pre-test

No	Class	Variance	N	F_{count}	F_{table}	Criteria
1	Experiment	38.77	30	1.26	2.1	Homogen
2	Control	48.85	30			

The more calculation can be seen in appendix 14.

Based on the formula:

$$F_{count} = \frac{\text{maximum variance}}{\text{minimum variance}}$$

F_{count} = 1.26

Based on the computation above it is obtained that F_{count} is lower than F_{table}, so Ho accepted. It can be concluded that the data of pre test from experimental and control class have the same variance or homogeneous.

c) Testing the similarity of average of the initial data between experimental and control class.

To test the difference of average, the writer used t-test.

Ho: $\mu_1 = \mu_2$

Ha: $\mu_1 \neq \mu_2$

Where:

μ_1 : average data of experimental group

μ_2: average data of control group

Table 12

The average similarity test of pre-test of experimental and control class

Source of variance	Experimental	Control	Criteria
Sum	1705	1700	Same
N	30	30	
Average	56.83	56.67	
Variance (S^2)	38.76	48.85	
Standard deviation (S)	6.23	6.99	

The more calculations can be seen in appendix 16

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

= 6.618

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt[5]{\frac{1}{n_1} + \frac{1}{n_2}}}$$

t = 0.098

Ho is accepted if $-t_{(1-\frac{1}{2}\alpha)(n1+n2-2)} < t < t_{(1-\frac{1}{2}\alpha)(n1+n2-2)}$. Based on the computation above, by α = 5% and df = 30+30- 2 = 58 is obtained t_{table} = 1,67 and t_{count} = 0.098. Ho is accepted if $-t_{table} < t_{count} < t_{table}$. So, it can be concluded that there is not significant different of the average pre-test between experimental and control class, because t_{count} at the reception area of Ho.

2. Analysis Phase End

It is done to answer hypothesis of this research. The data used are the result of post tests of both classes. The experimental class taught by using pictures as media and the control class taught without pictures.

The final analysis contains the normality test, homogeneity test and the difference average test of post test.

a. Normality test of the Post-test

Ho : the data of normal distribution

Ha : the data of un normal distribution

With criteria, Ho accepted if $x^2_{count} < x^2_{table}$ with α = 5% and df = k-3.

Table 13

The result of normality test of experimental and control class post-test

No	Class	x^2_{count}	x^2_{table}	Criteria
1	Experiment	5.842	7,81	Normal
2	Control	5.263	7,81	Normal

The more calculations can be seen in appendix 12 and 13

Based on the computation above it is obtained that x^2_{count} is lower than x^2_{table} by α = 5% with df = 6-3 = 3. So it can be concluded that the distribution data of post test of experimental and control class are normal.

b. Homogeneity test of the post-test

Ho = $\sigma_1^2 = \sigma_2^2$ (homogeny variance)

Ha = $\sigma_1^2 \neq \sigma_2^2$ (non homogeny variance)

With criteria, Ho accepted if $F_{count} < F_{table}$ with $\alpha = 0.05$ and df = k-1.

Table 14

The result of homogeneity post-test of experiment and control class

No	Class	Variance	n	F_{count}	F_{table}	Criteria
1	Experiment	97.586	30	2.73	2.1	Non Homogen
2	Control	35.747	30			

The more calculation can be seen in appendix 15.

Based on the formula:

$$F_{count} = \frac{\text{maximum variance}}{\text{minimum variance}}$$

$F_{count} = 2.73$

Based on the computation above it is obtained that F_{count} is higher than F_{table}, it means that Ho rejected. It can be concluded that data of post test of experimental and control classes have not the same variance or non homogeneous.

2. Hypothesis Test

Hypothesis test is used to know whether there is a difference average on post test of experimental class and control class. The data which is used to test the hypothesis is score post test both of class. To test the difference of average used t-test.

Ho: $\mu_1 = \mu_2$: it means there is no significant difference between the listening comprehension skill improvement of students who were taught by using pictures and who were taught by lecturing (without using pictures)

Ha: $\mu_1 \neq \mu_2$: it means there is significant difference between the listening comprehension skill improvement of students who were taught by using pictures and who were taught by lecturing (without using pictures)

Ha is accepted if $t_{count} > t_{(1-\alpha)(n1+n2-2)}$

Table 15

The score of experimental and control class post test

No	Source of variance	Experimental	Control
1	N	30	30
2	Average	68.00	62.33
3	Variance	97.59	35.75
4	Standard deviation	9.88	5.98
5	Maximal score	90	75
6	Minimal score	50	50

The more calculations can be seen in appendix 4

Table 16

Result of computation t-test

Class	N	Average ($\bar{X}$)	Variance (S^2)	Standard Deviation (s)	t_{table}	t_{count}	Criteria
Experiment	30	68.00	97.59	9.88	1.67	2.688	Ha accepted
Control	30	62.33	35.75	5.98			

The more calculation can be seen in appendix 17.

Based on the computation above, it is obtained that the average of post test of the experimental class who are taught by using pictures is 68.00 and standard deviation (s) is 9.88. While the average of post test of the control class who are taught by lecturing or conventional learning is 35.75 and standard deviation (s) is 5.98 with df = 30+30-2 = 58 by α = 5%, so obtained t_{table} = 1.67 from the result of calculation t-test t_{count} = 2.688. It means that t_{count} is higher than t_{table} **(**$t_{count} > t_{table}$**).** So Ho is rejected and Ha is accepted.

Because $t_{count} > t_{table}$**,** it can be concluded that there is a significant difference between experimental and control classes on post test, the score of the experimental class is higher than the control class.

C. Discussion of The Research Findings

1. The score of initial ability (Pre test)

   Based on the calculations of normality and homogeneity test from class VIII D as the experimental class and class VIII F as the control class, both of classes are normal distribution and homogeneous.

2. The score of final ability (Post test)

   The result of this research is obtained the average score of experimental class was 68.00 which were higher than the result of control class 62.33.

   The average score of experimental class was 68.00 and standard deviation (s) was 9.88. Teaching listening in experimental class by using pictures as media to teach narrative texts can encourage the students to be more active and motivated. Pictures as a teaching media that can create situation in teaching listening more interesting and make the students easier to understand the material. It can be seen on average score of experimental class which better result than control class.

   The average score of control class was 62.33 and standard deviation (s) was 5.68. Teaching listening in control class by using conventional learning or lecturing to teach listening of narrative texts make the students feel bored with the material that is presented because the method too monotone.

   Based on the result of calculation t-test is obtained t_{count}: 2,688 and t_{table}: 1,67 with $\alpha = 5\ \%$ and $df = 58$. It shows that $t_{count} > t_{table}$ (t_{count} higher than t_{table}). So it means that there is a significant difference between listening skill improvement of students taught by using pictures and taught by lecturing or conventional learning in listening of narrative texts.

D. Limitation of the Research

The researcher realizes that this research had not been done optimally. There were constraints and obstacles faced during the research process. Some limits of this research were:

1. Relative short time of research makes this research could not be done maximum.
2. The research was limited at SMPN 31 Semarang in the academic year of 2011/ 2012. So that when the same research will be gone in other schools, it was still possible to get different result.
3. The implementation of the research process was less smooth; this was more due to lack of experience and knowledge of the researcher.

Considering all those limitations, there is a need to do more research about teaching listening skill on narrative text using the same or different media. In the hope there will be more optimal result.