

CHAPTER IV

RESEARCH FINDINGS AND DISCUSSION

This chapter presents about the description of data, hypothesis testing, and also discussion based on the result of this research.

A. The Description of Data

In this study, the researcher used quantitative research by using quasi-experimental research design. The samples of this research were VIII-D as experimental class and VIII-E as control class at SMP Terpadu Al Anwar Durenan Trenggalek. An experimental consisted of 34 students while control class consisted of 38 students. The sampling technique that used by the researcher was purposive sampling. The researcher used instrument to collect the data, namely pre-test and post-test. Before conducting pre-test and post-test for both of classes, the researcher conducted try out to the same grade at SMP Terpadu Al Anwar Durenan Trenggalek of VIII-B that consisted of 33 students.

Every class got four meetings. The first meeting was used to conduct pre-test. In pre-test, the students could choose one of topics that given by the researcher. They were asked to write minimum in ten lines without any maximum lines. The purpose of this pre-test was to know the students' ability in writing recount text before given treatments. After pre-test was administered, the scores of pre-test were used to know whether those classes were equal or not. Moreover, the result of normality testing of pre-test

showed that the data was not normally distributed (see table 3.11). So that was way, the researcher used Mann-Whitney to know those classes were equal or not. Based on the Mann-Whitney result of pre-test, those classes were not equal. They had no the same ability in writing. Because those classes were not equal, the N-Gain score both of classes was calculated. Then, in the second and the third meeting were used to conduct treatments. An experimental was given special treatment by using peer feedback, while control class got treatment but didn't use peer feedback (conventional treatment). After giving the treatments, the researcher gave post-test for both of classes in the fourth meeting. As same as pre-test, the students were asked to choose one of topics given and write it minimum in ten lines without any maximum lines. The purpose of this post-test was to know the students' ability in writing recount text after given the treatments.

The researcher gained the students' scores both of classes by scoring rubric that has been settled with the detail descriptions (see appendix 4). After gaining the students' score by using scoring rubric, the researcher gave the scoring criteria for their scores. The purpose of scoring criteria was to know whether the most of their scores both of control and experimental class included in excellent, good, average, poor, or very poor criteria. For the scores' criteria and analysis of pre-test and post-test scores could be seen in table 4.1 as follows:

Table 4.1 Scores' Criteria

Scores	Description
0-20	Very poor
21-40	Poor
41-60	Average
61-80	Good
81-100	Excellent

a. Control Class

In control class, the researcher collected two data. They were data of pre-test scores and data of post-test scores. It was collected before and after given conventional treatment. The data of pre-test scores was used to know whether both of classes had an equal variance or not. For the detail explanation, it was explained in the following explanation.

1) Pre-Test Scores of Control Class

Control class was class which didn't get special treatment (peer feedback) from the researcher. It got the conventional treatment that conducted by the researcher as usual. Before giving conventional treatment, the researcher gave pre-test for this class. The class that became control class was VIII-E. It consisted of 38 students, but there were 2 students who were absents. So, there were 36 students took the pre-test. The result of pre-test scores of control class could be seen in the following table (see table 4.2).

Table 4.2 Pre-Test Scores of Control Class

NO	NAME	SCORE
1	ADP	75
2	AF	75
3	A	65
4	ACA	75
5	DR	75
6	ESA	70
7	ENS	70
8	FFS	60
9	FAZN	75
10	FLJ	75
11	IDS	80
12	IFS	80
13	LQM	75
14	LAR	75
15	LM	60
16	MR	80
17	ME	60
18	NEA	75
19	NLN	75
20	NBKS	80
21	NP	70
22	RH	70
23	RKM	60
24	RZ	70
25	SAA	70
26	SMP	75
27	UZA	70
28	VAP	50
29	VNAA	75
30	VNP	55
31	WNFA	65
32	YA	75
33	ZCM	85
34	NA	40
35	N	80
36	NHZ	75

Table 4.2 showed that pre-test in control class taken by 36 students. Every student had different scores based on their ability. The highest score was 85 and the lowest score was 40. For the result of statistic calculation of pre-test data by using SPSS 16.0 version for windows could be seen in the table 4.3.

Table 4.3 Statistical Analysis of Pre-Test in Control Class

Statistics		
Pre-Test of Control Class		
N	Valid	36
	Missing	0
Mean		70.56
Std. Error of Mean		1.553
Median		75.00
Mode		75
Std. Deviation		9.318
Variance		86.825
Range		45
Minimum		40
Maximum		85
Sum		2540

Table 4.3 showed that there were 36 students took pre-test. The mean score of pre-test in control class was 70.56. The median was 75.00. The mode of the score was 75 and the standard deviation

was 9.318. The frequencies of the students' scores in pre-test were presented in the table 4.4 as follows.

Table 4.4 Frequency of Students' Pre-test Scores in Control Class

		Pre-test			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	40	1	2.8	2.8	2.8
	50	1	2.8	2.8	5.6
	55	1	2.8	2.8	8.3
	60	4	11.1	11.1	19.4
	65	2	5.6	5.6	25.0
	70	7	19.4	19.4	44.4
	75	14	38.9	38.9	83.3
	80	5	13.9	13.9	97.2
	85	1	2.8	2.8	100.0
	Total	36	100.0	100.0	

From the table 4.4, it showed that 1 student got 40 scores, 1 student got 50 scores, 1 student got 55 scores, 4 students got 60 scores, 2 students got 65 scores, 7 students got 70 scores, 14 students got 75 scores, 5 students got 80 scores, and 1 student got 85 scores. It could be concluded that only one student got the excellent scores based on the score criteria (see table 4.1). Most of their scores included in good criteria.

2) Post-Test Scores of Control Class

After giving conventional treatment (without using peer feedback), the researcher gave post-test for control class. The post-test was taken by 36 students because 2 students were getting sick. The post-test scores of control class could be seen in table 4.5 as follows.

Table 4.5 Post-Test Scores of Control Class

NO	NAME	SCORE
1	ADP	80
2	AF	65
3	A	65
4	ACA	80
5	DR	80
6	ESA	75
7	ENS	65
8	FFS	65
9	FAZN	65
10	FLJ	60
11	IDS	75
12	IFS	85
13	LQM	75
14	LAR	85
15	LM	60
16	MR	80
17	ME	50
18	NEA	85
19	NLN	80
20	NBKS	90
21	NP	85
22	RH	85
23	RKM	60
24	RZ	50
25	SAA	85

Continued

Continuation

26	SMP	70
27	UZA	75
28	VAP	50
29	VNAA	70
30	VNP	55
31	WNFA	75
32	YA	70
33	ZCM	55
34	NA	60
35	N	75
36	NHZ	75

Table 4.5 showed that the highest score of post-test was 90 and the lowest score was 50. For the result of statistic calculation of data by using SPSS 16.0 version for windows could be seen in table 4.6 as follows:

Table 4.6 Statistical Analysis of Post-Test in Control Class

Statistics		
Post-Test of Control		
N	Valid	36
	Missing	0
	Mean	71.11
	Median	75.00
	Mode	75
	Std. Deviation	11.409
	Variance	130.159
	Range	40
	Minimum	50
	Maximum	90
	Sum	2560

Table 4.6 showed that there were 34 students took post-test. The mean score of pre-test in control class was 71.11. The median was 75.00. The mode of the score was 75 and the standard deviation was 11.409.

From the data, it known that there was an improvement in their ability in writing recount text without using peer feedback. It could be seen in their mean of pre-test scores was 70.56 and the mean scores of post-test was 71.11. It meant that the scores of post-test were better than scores in pre-test. The frequencies of the students' scores in post-test were presented in the following table (see table 4.7).

Table 4.7 Frequency of Students' Post-test Scores in Control Class

Post-test					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	50	3	8.3	8.3	8.3
	55	2	5.6	5.6	13.9
	60	4	11.1	11.1	25.0
	65	5	13.9	13.9	38.9
	70	3	8.3	8.3	47.2
	75	7	19.4	19.4	66.7
	80	5	13.9	13.9	80.6
	85	6	16.7	16.7	97.2
	90	1	2.8	2.8	100.0
	Total	36	100.0	100.0	

From the table 4.7, it showed that 3 students got 50 scores, 2 students got 55 scores, 4 students got 60 scores, 5 students got 65 scores, 3 students got 70 scores, 7 students got 75 scores, 5 students got 80 scores, 6 students got 85 scores, and 1 student got 90 score. It could be concluded that there were 7 students got the excellent scores based on the score criteria (see table 4.1). Most of their scores included in good criteria.

b. Experimental Class

As in the control class, the researcher also collected two data in the experimental class. They were pre-test and post-test score. The pre-test was conducted before giving special treatment (peer feedback). The result of pre-test scores was used to know whether both of classes had an equal variance or not. And post-test was conducted after giving special treatment (peer feedback). The researcher gave the detail explanation in the following explanation.

1) Pre-Test Scores of Experimental Class

Experimental class was class which given special treatment (peer feedback). The class who became experimental class was VIII-D which consisted of 34 students. Before giving special treatment, the researcher gave pre-test first. It conducted to know their ability before giving the treatment. It was taken by 33 students because one student was absent. The scores of pre-test in experimental class could be seen in table 4.8.

Table 4.8 Pre-Test Scores of Experimental Class

NO	NAME	SCORE
1	ARPSR	80
2	AN	85
3	ALN	75
4	AKF	80
5	AFM	80
6	AK	85
7	AS	90
8	AMM	90
9	BUH	85
10	ESUN	80
11	ENA	75
12	EWN	75
13	FPR	85
14	FM	85
15	ITQ	85
16	KZ (a)	85
17	KZ (i)	80
18	LDS	85
19	MY	80
20	MZ	75
21	NQA	85
22	NN	75
23	NDM	85
24	NL	75
25	RJ	90
26	SAN	75
27	SNA	85
28	SNN	65
29	SSNP	85
30	SMN	85
31	TJFZ	75
32	WEP	80
33	WIPAH	75

Table 4.8 showed that pre-test in experimental class was taken by 33 students. They had different scores based on their

writing ability. The highest score of pre-test was 90 and the lowest score was 65. For the result of statistic calculation of data by using SPSS 16.0 version for windows could be seen in the table 4.9 as follows:

Table 4.9 Statistical Analysis of Pre-Test in Experimental Class

Statistics		
Pre-test of Experiment		
N	Valid	33
	Missing	0
Mean		81.06
Median		80.00
Mode		85
Std. Deviation		5.695
Variance		32.434
Range		25
Minimum		65
Maximum		90
Sum		2675

Based on the result of statistic calculation by using SPSS in table 4.9, the mean of the students' scores was 81,06. The median was 80. The mode was 85 and the standard deviation was 5.695. The frequencies of the students' scores in pre-test were presented in the following table (see table 4.10) in the next page.

Table 4.10 Frequency of Students' Pre-test Scores in Experimental Class

		Pre-test			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	65	1	3.0	3.0	3.0
	75	9	27.3	27.3	30.3
	80	7	21.2	21.2	51.5
	85	13	39.4	39.4	90.9
	90	3	9.1	9.1	100.0
	Total	33	100.0	100.0	

From the table 4.10, it showed that 1 student got 65 scores, 9 students got 75 scores, 7 students got 80 scores, 13 students got 85 scores, 3 students got 90 scores. It could be concluded that there were 12 students got the excellent scores based on the score criteria (see table 4.1). Most of their scores included in good criteria.

2) Post-Test Scores of Experimental Class

After giving special treatment, the researcher conducted post-test to measure their ability in writing recount text after using peer feedback. The subjects of experimental class were 34 students. But, the post-test was taken by 33 students. The post-test scores of experimental class could be seen in table 4.11.

Table 4.11 Post-Test Scores of Experimental Class

NO	NAME	SCORE
1	ARPSR	80
2	AN	70
3	ALN	75
4	AKF	85
5	AFM	90
6	AK	85
7	AS	90
8	AMM	90
9	BUH	90
10	ESUN	85
11	ENA	70
12	EWN	75
13	FPR	85
14	FM	85
15	ITQ	90
16	KZ (a)	80
17	KZ (i)	85
18	LDS	90
19	MY	90
20	MZ	75
21	NN	90
22	NDM	85
23	NL	80
24	RJ	90
25	SAN	80
26	SNA	80
27	SNN	70
28	SSNP	80
29	SMN	90
30	TJFZ	85
31	WEP	85
32	WIPAH	85
33	WK	90

Table 4.11 showed that the highest score of post-test in experimental class was 90 and the lowest score was 70. For the result of statistic calculation of data by using SPSS 16.0 version for windows could be seen in the table 4.12 as follows:

Table 4.12 Statistical Analysis of Post-Test in Experimental Class

Statistics		
Post-test of Experimental		
N	Valid	33
	Missing	0
Mean		83.48
Std. Error of Mean		1.120
Median		85.00
Mode		90
Std. Deviation		6.433
Variance		41.383
Range		20
Minimum		70
Maximum		90
Sum		2755

Based on the result of statistic calculation by using SPSS in table 4.12, the mean of the students' scores was 83.48. The median was 85.00. The mode was 90 and the standard deviation was 6.433. From the data, it was known that there was improvement in their ability in writing recount text after using pee feedback. It could be seen in their mean of pre-test scores before using peer feedback was 81.06 and the mean scores of post-test after using peer feedback was 83.48. It meant that post-test sores were better than pre-test scores because the students could look their mistakes and they repair or correct their mistakes in writing recount text. So, their writing ability

was getting improvement. The frequencies of the students' scores in post-test were presented in the following table (see table 4.13).

Table 4.13 Frequency of Students' Post-test Scores in Experimental Class

Post-test					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	70	3	9.1	9.1	9.1
	75	3	9.1	9.1	18.2
	80	6	18.2	18.2	36.4
	85	10	30.3	30.3	66.7
	90	11	33.3	33.3	100.0
	Total	33	100.0	100.0	

From the table 4.13, it showed that 3 students got 70 scores, 3 students got 75 scores, 6 students got 80 scores, 10 students got 85 scores, 11 students got 90 scores. It could be concluded that there were 21 students got the excellent scores based on the score criteria (see table 4.1). Most of their scores included in excellent criteria.

c. N-Gain Score of Pre-Test in Experimental Class and Control Class

After the researcher got the statistical analysis of data from the students' pre-test in experimental and control class, the researcher found that the mean score of pre-test in experimental was very different from the mean score of pre-test in control class. It could

be seen that the mean score of pre-test in experimental class was 80.06 while the mean score of pre-test in control class was 70.65. Moreover, the result of Mann-Whitney of pre-test both of classes showed that those classes were not equal or those classes had no same ability in writing (see table 3.4). So that was way, the researcher calculated N-Gain Score of pre-test and post-test both of those classes by using SPSS 16.0 version for windows. The purpose of this N-Gain Scores was to know the effectiveness of using a method, strategy, technique, or treatment in this research both of using control and experimental class or pre-experimental design (one group pre-test post-test design). By using N-Gain scores, we could know whether that treatment (method, strategy, or technique) was effective or not. According to Sundayana (2015: 151) the category of N-Gain Score as follows:

Table 4.14 Category of N-Gain Score

N-Gain Score	Category
$N\text{-Gain} > 0,70$	High
$0,30 \leq N\text{-Gain} \leq 0,70$	Average
$N\text{-Gain} < 0,30$	Low

The result of N-Gain Score could be seen in the following table 4.15 in the next page.

Table 4.15 The Result of N-Gain Score

Group Statistics					
	Kelas	N	Mean	Std. Deviation	Std. Error Mean
NGain_Score	Experiment	33	7.3593	42.02947	7.31639
	Control	36	-1.8221	45.52611	7.58769

From table 4.15, it was found that the mean of experimental class was 7.3593 or if it was rounded off equal to 7.4. Based on the category of N-Gain Score (see table 4.15), it included in high category of effectiveness because the N-Gain score was higher than 0.70 ($7.4 > 0.70$) (see table 4.14 for category of N-Gain Score) . It meant that the use of peer feedback was very effective to be applied in experimental class.

Before calculating to know the significant difference on students writing recount text ability assessed by using peer feedback (experimental class) and without using peer feedback (conventional treatment for control class) using SPSS 16.0 version for windows by parametric or non-parametric test, the normality of N-Gain score had to be calculated to know whether the data of N-Gain score was normally distributed or not. This case used N-Gain score because those classes (experimental and control class) were not equal. The result of normality testing of N-Gain score could be seen in table 4.16 in the next page:

Table 4.16 The Result of Normality Testing of N-Gain Score

Kelas	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
NGain_Score Experiment	.249	33	.000	.833	33	.000
Control	.183	36	.004	.793	36	.000

a. Lilliefors Significance Correction

From table 4.16, it focused in Kolmogorov-Smirnov column because the sample was more than 50 students. In Kolmogorov-Smirnov, it could be known that significance value (Sig.) in experimental class was 0.000 and in control class was 0.004. It indicated that significance value was lower than significance level 0.05 ($0.000 < 0.05$) and ($0.004 < 0.05$). It meant that the data of N-Gain Score was not normally distributed. Because the data of N-Gain score was not normally distributed, so it was not necessary to calculate the homogeneity testing.

B. Hypothesis Testing

The hypothesis testing was used to reveal whether there was a significant difference on the students' writing ability in recount text between assessed with and without using peer feedback. Before the hypothesis was rejected or accepted, it must be changed to the null hypothesis (H_0). The hypothesis testing between experimental and control class could be seen from null hypothesis (H_0) and alternative hypothesis (H_a). Because this research

used one-right tailed, the formula to reject or not reject the null hypothesis was:

H_0 : If P -Value ≤ 0.05

H_a : If P -Value > 0.05

From the formula above, it meant that:

1. If the p -value (significance value) was less than or equal to 0.05 ($\alpha = 5\%$), then the null hypothesis (H_0) was rejected and the alternative hypothesis (H_a) was accepted. It meant that there was significant difference on students' writing recount text ability assessed with and without assessed by using peer feedback.
2. If the p -value (significance value) was greater than 0.05 ($\alpha = 5\%$), then the null hypothesis (H_0) was accepted and the alternative hypothesis (H_a) was rejected. It meant that there was no significant difference on students' writing recount text ability assessed with and without assessed by using peer feedback.

Because the data of N-Gain score was not normally distributed (see table 4.16), the researcher used Mann-Whitney U Test to know the significant difference on students' writing recount text ability assessed with and without using peer feedback. According to Uyanto (2009: 321), Mann-Whitney is non-parametric test. It is used to compare two samples with ordinal or interval scale with non-normally distributed.

The result of hypothesis testing could be seen in table 4.18.

Table 4.17 The Result of Hypothesis Testing

Test Statistics ^a	
	NGain_Score
Mann-Whitney U	489.000
Wilcoxon W	1155.000
Z	-1.274
Asymp. Sig. (2-tailed)	.203

a. Grouping Variable: Peer Feedback

From table 4.17, it could be made interpretation that z-value was -1.274 with p -value 0.203 (2-tailed). Because this research used one right-tailed test, so the p -value has to divided into two. Thus, $0.203 : 2$ equals to 0.1015. It meant that p -value was higher than significance level 0.05 ($0.1015 > 0.05$). In consequence, the null hypothesis could not be rejected and alternative hypothesis was not accepted. It meant that there was no significant difference on students' writing recount text ability assessed with and without assessed by using peer feedback. Then, it could be concluded that null hypothesis (H_0) which stated there was no significant difference on students' writing recount text ability assessed with and without using peer feedback was accepted. In contrary, the alternative hypothesis (H_a) which stated that there was significant difference on students' writing recount text ability assessed with and without using peer feedback was rejected.

C. Discussion

The main purpose of this research was to know the significant different on students' writing recount text ability assessed with and without using peer feedback of the second grade students at SMP Terpadu Al Anwar Durenan Trenggalek. To collect the data both of classes, the researcher used test as instrument. There were two tests, namely pre-test and post-test. The first step that conducted by the researcher was giving pre-test for both of classes. The purpose of this pre-test was to know the students' ability before given the treatments. After the researcher got the students' scores of pre-test, then the researcher calculated to know whether both of classes were equal or not. Because the result of normality testing of pre-test showed that the data was not normally distributed, so the researcher used Mann-Whitney U Test in SPSS 16.0 version for windows. The result showed that sig (2-tailed) was less than significance level 0.05 ($0.000 < 0.05$) (see table 3.4 in page 41). It indicated that that those classes were not equal in writing ability. Moreover, the mean score of experimental class was 81.06 and the mean score of control class was 70.56. It indicated that the mean of experimental class was very different with the mean of control class. Because those classes were not equal, so N-Gain Score of experimental and control class was calculated.

After conducting pre-test for both of classes, the next step was giving treatments in the second and third meeting. For experimental class got the special treatment by using peer feedback in writing recount text. But, in control group got the treatment also but didn't get special treatment by using

peer feedback. After the treatment had been conducted, the next step was conducting post-test for both of classes. Both of experimental and control class got the same question from the researcher. It was about writing recount text. The students were asked to write recount text minimum in 10 lines without any maximum lines based on the topic given.

After the researcher analyzed the post-test scores of experimental and control class, it resulted the mean score of post-test in control class was 71.11 and the mean score of post-test in experimental class was 83.48. It meant that the mean score of experimental class was higher than mean score of control class. As stated in the beginning that those classes were not equal, so the researcher used Mann-Whitney U Test of N-Gain score data both of classes to reject the null hypothesis. The output of Mann-Whitney U Test (see table 4.17) showed that p -value was higher than significance level 0.05 ($0.1015 > 0.05$). It meant that the null hypothesis could not be rejected and alternative hypothesis could not be accepted. It indicated that there was no significant difference on students' writing recount text ability assessed with and without using peer feedback. Based on the research finding, the experimental got the higher scores both in pre-test and post-test than control class. But, it didn't mean that an experimental class was better than control class because those classes were not the same in their ability in writing. It might be caused some factors that happened when the researcher conducted research.

According to Andianto (2014), based on the result of his research, he stated that peer feedback technique was more effective than teacher feedback

technique in writing because the students who were taught by using peer feedback technique had better writing ability than students who were taught without using peer feedback technique. He also stated that peer feedback focused on student center. Peer feedback allowed the students to gain confidence, critical thinking skill from being able to read text by peer writing on similar tasks. Moreover, by applying peer feedback, students learn by themselves, learn more, feel more confidence, enjoy the class, teach each other, and became independent learners.

Moreover, in the previous study, Astuti (2013) conducted research to know the significant difference on students writing ability with and without using peer feedback. Her result showed that the use of peer feedback in teaching learning process of writing made a significant improvement in the students' score. Therefore, it could be stated that the use of peer feedback in teaching writing could be used to solve the students' writing problem and increasing the students writing ability. The students benefited peer feedback process as their peers told the mistakes that they were not aware of and gave them opinions and suggestions (Kurt and Atay: 2007).

Thus, based on the result of this research, it could be concluded that there was no significant difference on students' writing recount text ability assessed with and without using peer feedback of the second grade students at SMP Terpadu Al Anwar Durenan Trenggalek. It could be seen from the result of Mann-Whitney of N-Gain Score showed that p -value was higher than significance level 0.05 ($0.1015 > 0.05$). Then, it could be concluded that null

hypothesis (H_0) which stated there was no significant difference on students' writing recount text ability assessed with and without using peer feedback was accepted. In contrary, the alternative hypothesis (H_a) which stated that there was significant difference on students' writing recount text ability assessed with and without using peer feedback was rejected. This research was contrary with the previous studies that showed that there was significant difference on students' writing ability taught with and without using peer feedback.

In this research, there was no significant different on students' writing recount text ability assessed with and without using peer feedback could be caused by some factors. It caused the treatment was only twice. The researcher only gave two treatments. Every treatment had no the same of time allocation. The first treatment was 45 minutes and the second treatment was 90 minutes. So, the researcher thought that the time allocation was less for them, especially in writing. In writing, the students needed more time in order to they could write and assess their peer's writing optimally. Moreover, this research conducted in the muslim boarding school, so the students had many activities in dormitory until night. So that was way, the students could not focus well when conducted treatment. Thus, they could not do the treatment optimally.