CHAPTER IV

RESEARCH FINDING AND DISCUSSION

The chapter explains the finding of this study from the result of analyzing the data. Hence, this chapter provides the Description of the Data experiment, Hypothesis testing, and Discussion of the Results this study.

A. Description of the Data

This research used treatment which applied at both of samples to manage threading comprehension activity. Collaborative Strategic Reading method was implemented to experimental class. While, the lecture method (teacher-centered) was given to control class (as conventional method). Then, the researcher compared and measured the students' ability according to the score of tests. This study purposed to see the significant difference of students' achievement taught with and without Collaborative Strategic Reading for reading comprehension ability of recount text.

Furthermore, the instrument this study was 25 items multiple choice question for post-test. The post-test was given to all of students in class VIII F (control) and class VIII G (experimental) MTs Ma'arif NU Kota Blitar. The researcher had implemented reliability and validity testing on the instrument and tried out the test before distributing to the samples. When managing validation test, there were 19 numbers from 25 questions in post-test were valid. Therefore,

there were 6 numbers of questions that should be revised. After revising the instrument, in reliability testing, the test was reliable. It was shown from reliability testing in *IBM SPSS 22.0 versions for Windows*, the reliability statistics was 0.428. Based on Cronbach Alpha's, it meant that was quite reliable because it was between 0.40 - 0.60.

The researcher did not use individual scores for comparison the results, but the researcher used the results of class scores (mean of the scores) in reading recount text. The result of class scores or mean was gained from scores of pre and posttest in control class and experimental class. Therefore, the researcher analyzed the data in descriptive statistics using *IBM SPSS 22.0 version for Windows* to know the central tendency (mean, median, mode) and the variability (standard deviation and range) of the students' reading comprehension achievement from pre-test and post-test scores..

1. Experimental Class (8G)

a.) Pre-test Score of Experimental Class

The class got the treatment (Collaborative Strategic Reading) from the researcher was experimental class. Pre-test was conducted for this class before giving the treatment. Class 8G was experimental class which consisted of 24 students. There were 24 students took the pre-test. This Table 4.1 showed students' pre-test scores of **experimental class**:

Table 4.1: Pre-test Scores of Experimental Class

Students (X)	Pre-test
ANJ	50
CYN	60
DEV	70

DYA	70
ENG	60
FAI	50
FEL	40
HEL	70
INT	40
JES	70
KHO	40
NAB	70
NAD	40
NATH	60
NID	40
RE	70
REY	50
RIZ	60
SEM	50
SIL	50
USA	40
VAN	70
WAH	30
WIJ	40

From Table 4.1, it showed the maximum and the minimum scores of pre-test. The maximum score was 70, then the minimum score was 30.Then, the pre-test results that gained from experimental class (VIII G) were described statistically in following table:

Table 4.2: Descriptive Statistics Pre-test of Experimental Class

Statistics

pre_test_ex

N	Valid	24
	Missing	0
Mean		53,75
Std. Error of Mean		2,679
Median		50,00
Mode		40[a]
Std. Deviation		13,126

Variance	172,283
Range	40
Minimum	30
Maximum	70
Sum	1290

a. Multiple modes exist. The smallest value is shown

Table 4.2 showed the data description of experimental class consist of 24 participants. Mean score was the total scores which divided by number of data and the mean of the data was 53.75. Next, median score was defined as the value of numbers divided higher half of the population, sample of data, or possibility distribution of lower half. The median score was 50.00. Then, there was mode that construed by the most frequently appeared element in a given. The number of mode was 40.00. Next, there were set of scores formed from population that was called as variance, which the number of variance was 172.283. Besides that, there was deviation standard which was a deviated measurement in a group as a whole that indicated the extent of quantity. The deviation standard that can be seen from the table was 13.126. Then, the distance between the highest value and the lowest one, called as range which was 40. The lowest score was 30 and the highest score was 70. The total score of the data was 1290.

According to the Table 4.2, the data was also presented in table of frequency distribution which showed in Table 4.3 as follow:

Table 4.3: Frequency Distribution Pre-test of Experimental Class

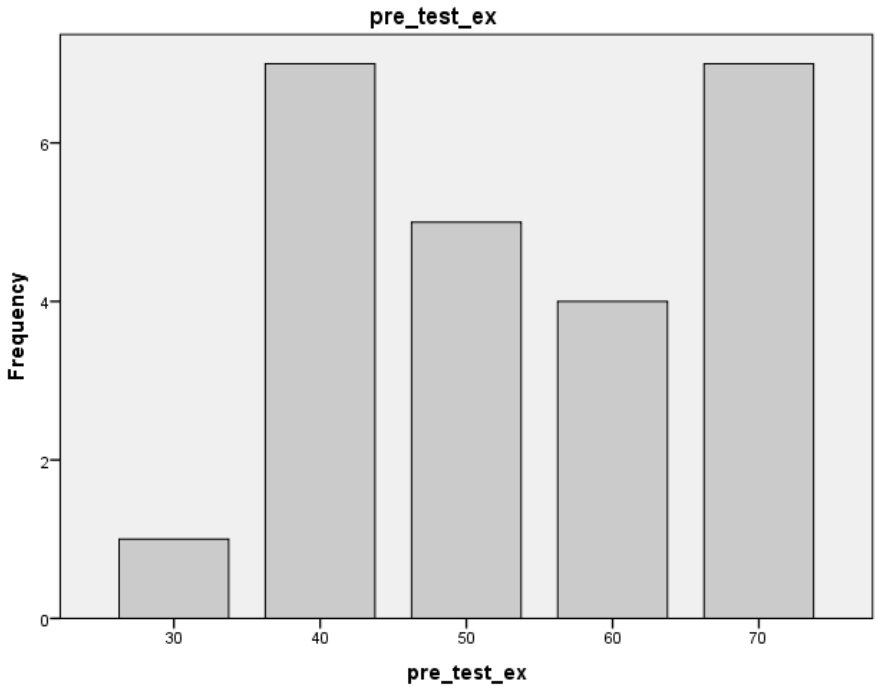
pre_test_ex

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	30	1	4,2	4,2	4,2
	40	7	29,2	29,2	33,3

	50	5	20,8	20,8	54,2
	60	4	16,7	16,7	70,8
	70	7	29,2	29,2	100,0
	Total	24	100,0	100,0	

In Table 4.3, it displayed the frequency of the scores. There were 1 student got the score of 30 with the percent 4.2%, 7 students got the score of 40 with the percent 29.2%, 5 students got the score of 50 with the percent 20.8%, 4 students got the score of 60 with the percent 16.7%, and 7 students got the score of 70 with the percent 29.2%. The table was a tool of presentation as a media to show the consisting of data frequency, the cumulative percent, and the distributed frequency in percentage.

Furthermore, the data from Table 4.3 also can be described by bar diagram which is presented as following picture:



Picture 4.1

The Diagram of Pre-test Scores of Experimental Class

In Picture 4.1, the score that have the greatest frequency was 40 and 70 with the number each score was 7 students. Then, the score that have the most little frequency was 30 with the number of the score was 1 student. It could be concluded that the number of students with the highest score were higher than the median score.

b.) Post-test Score of Experimental Class

Experimental class got the treatment (Collaborative Strategic Reading) from the researcher. The researcher conducted post-test for this class after the treatment. The experimental class was 8G that consisted of 24 students. There was no absent student, so there were 24 students took the post-test. The following table (Table 4.4) showed the score of **experimental class** students in post-test:

Table 4.4: Post-test Scores of Experimental Class

Students (X)	Post-test
ANJ	56
CYN	60
DEV	76
DYA	64
ENG	48
FAI	56
FEL	56
HEL	40
INT	60
JES	60
KHO	52
NAB	44
NAD	68
NATH	52
NID	60
RE	64
REY	40
RIZ	64

SEM	52
SIL	56
USA	48
VAN	64
WAH	48
WIJ	68

From Table 4.1, it showed the maximum, the minimum, and the scores of post-test from 24 students. The maximum score was 76, then the minimum score was 40.Then, the post-test result that gained from experimental class (VIII G) was described statistically in following table:

Table 4.5: Descriptive Statistics Post – test of Experimental Class

Statistics

post_test_ex

N	Valid	24
	Missing	0
Mean		56,50
Std. Error of Mean		1,854
Median		56,00
Mode		56[a]
Std. Deviation		9,084
Variance		82,522
Range		36
Minimum		40
Maximum		76
Sum		1356

a. Multiple modes exist. The smallest value is shown

Table 4.5 showed the data of experimental class which was post-test scores from 24 students. Mean score was the total scores which divided by number of data and the mean of the data was 56.50. Next, median score was defined as the value of numbers divided higher half of the population, sample of

data, or possibility distribution of lower half. The median score was 56.50 Then, there was mode as the most frequently appeared element in a given. The number of mode was 56. Next, there were set of scores formed from population that was called as variance, which the number of variance was 82.552. Besides that, there was deviation standard which was a deviated measurement in a group as a whole that indicated the extent of quantity. The deviation standard that can be seen from the table was 9.084. Then, the distance between the highest value and the lowest one, called as range which was 36. The lowest score was 40 and the highest score was 76. The total score of the data was 1356.

According to Table 4.5, the data was also presented in table of frequency distribution which showed in Table 4.6 below:

Table 4.6: Frequency Distribution Post – test of Experimental Class

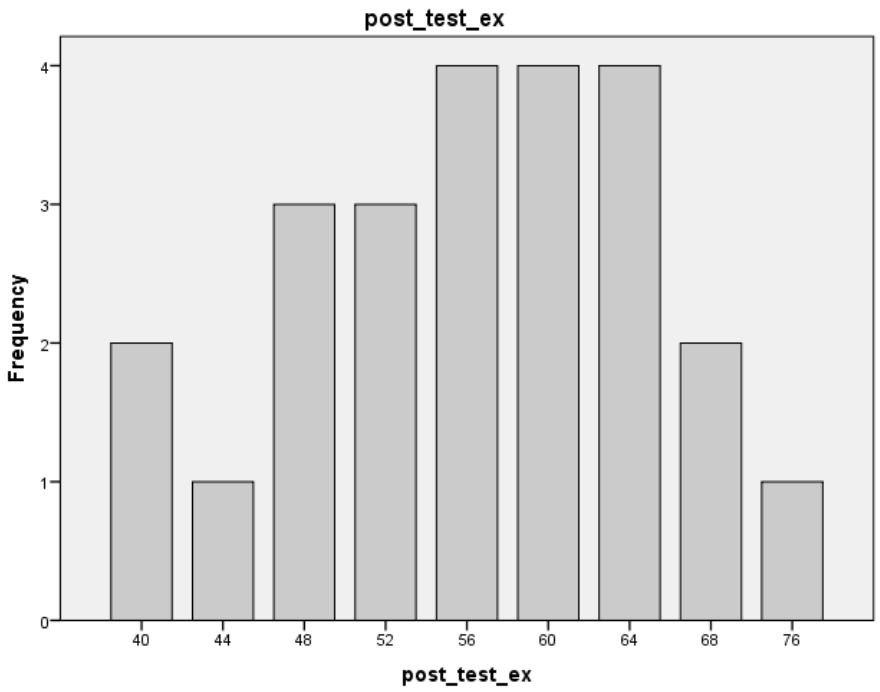
post_test_ex

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	40	2	8,3	8,3	8,3
	44	1	4,2	4,2	12,5
	48	3	12,5	12,5	25,0
	52	3	12,5	12,5	37,5
	56	4	16,7	16,7	54,2
	60	4	16,7	16,7	70,8
	64	4	16,7	16,7	87,5
	68	2	8,3	8,3	95,8
	76	1	4,2	4,2	100,0
	Total	24	100,0	100,0	

From the Table 4.6, it can be seen the frequency of the scores. Apparently, two students got the score of 40 with the percent 8.3%, a student got the score of 44 with the percent 4.2%, 3 students got the score of 48 with the percent 12.5%, 3

students got the score of 52 with the percent 12.5%, 4 students got the score of 56 with the percent 16.7%, 4 students got the score of 60 with the percent 16.7%, 4 students got the score of 64 with the percent 16.7%, 2 students got the score of 68 with the percent 8.3%, and 1 student got the score of 76 with the percent 4.2%. That table was a tool of presentation as a media to show the consisting of data frequency, the cumulative percent, and the distributed frequency in percentage

Based on the frequency distribution of post-test data, the diagram of Table 4.6 was presented as following picture:



Picture 4.2

The Diagram of Post-test Scores of Experimental Class

The score that have the greatest frequency was 56, 60, and 64 with the number each score was 4 students. Then, the score that have the most little frequency was 44 and 76 with the number each score was 1 student. The conclusion was the student with the highest was less than the median score.

2. Control Class (8F)

a.) Pre-test Score of Control Class

The class did not get the treatment (Collaborative Strategic Reading) from the researcher was control class. It got conventional method which used lecture method. The researcher conducted pre-test for this class before the treatment. The control class was 8F that consisted of 25 students. There were 25 students took the pre-test. The following table (Table 4.7), showed the scores of **control class's** students in pre-test.

Table 4.7: Pre-test Scores of Control Class

Students (Y)	Pre-test
AIDI	50
ALY	40
ANG	50
ANI	80
CI	80
DHE	80
DI	50
FIL	70
GLA	50
IRM	80
IRT	60
ITN	60
LUT	60
MEL	60
NABQ	70
NAB	70
NAD	70
NAJ	70
NEZ	80
NUR	60
PUT	50
RIK	60
UNI	60

ZF	80
ZQ	70

Based on the Table 4.7, it showed the scores of pre-test from 25 students. The minimum score from control class was 40, while the maximum score was 80.The result of pre-test from VIII F class as a control class was presented in Table 4.8, as follows:

Table 4.8: Descriptive Statistics Pre-test of Control Class

Statistics

Pre_test_co

N	Valid	25
	Missing	0
Mean		64,40
Std. Error of Mean		2,386
Median		60,00
Mode		60
Std. Deviation		11,930
Variance		142,333
Range		40
Minimum		40
Maximum		80
Sum		1610

Table 4.8 showed that the data was from 25 students of control class. Mean score was the total scores which divided by number of data and the mean of the data was 64.60. Next, median score was defined as the value of numbers divided higher half of the population, sample of data, or possibility distribution of lower half. The median score was 60.00. Then, there was mode as the most frequently appeared element in a given. The number of mode was 60. Next, there were set of scores formed from population that was called as variance. The

number of variance was 142.333. Besides that, there was deviation standard which was a deviated measurement in a group as a whole that indicated the extent of quantity. The deviation standard that can be seen from the table was 11.930. Then, the distance between the highest value and the lowest one, called as range which was 40. The lowest and the highest score were 40 and80. The total score of the data was 1610.

Based on the Table 4.8, the data was also presented in table of frequency distribution which showed in Table 4.9 below:

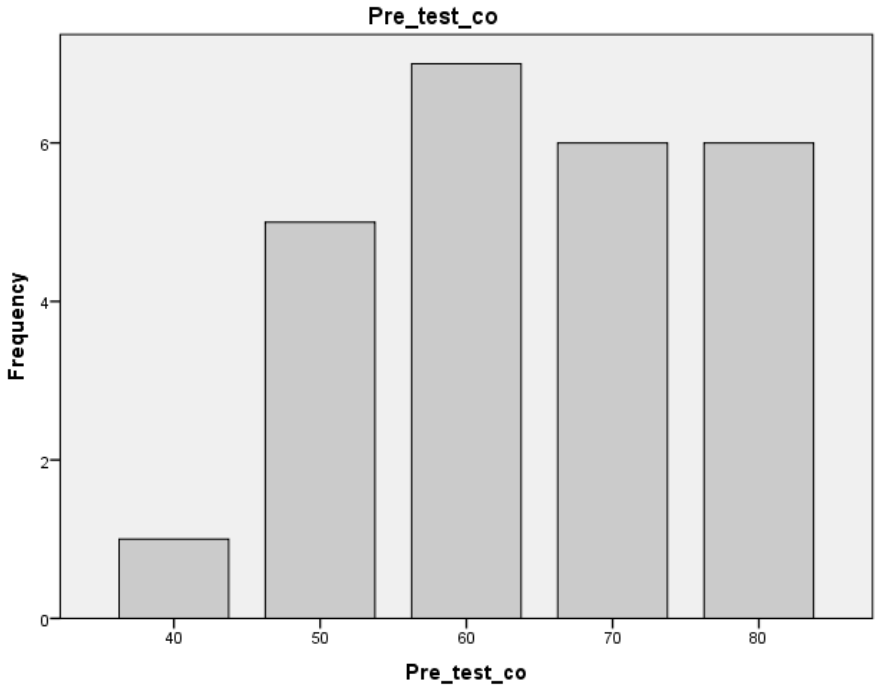
Table 4.9: Frequency Distribution Pre-test of Control Class

Pre_test_co

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	40	1	4,0	4,0	4,0
	50	5	20,0	20,0	24,0
	60	7	28,0	28,0	52,0
	70	6	24,0	24,0	76,0
	80	6	24,0	24,0	100,0
	Total	25	100,0	100,0	

From Table 4.9, it showed the frequency of the scores. There were 1 student got the score of 40 with the percent 4.0%, 5 students got the score of 50 with the percent 20.0%, 7 students got the score of 60 with the percent 28.0%, 6 students got the score of 70 with the percent 24.0%, and 6 students got the score of 80 with the percent 24.0%. That table was a tool of presentation as a media to show the consisting of data frequency, the cumulative percent, and the distributed frequency in percentage.

Then, the diagram from the data of Table 4.9 is presented as following picture:



Picture 4.3

The Diagram of Pre-test Scores of Control Class

The score that have the greatest frequency was 60 with the number of score was 7 students. Then, the score that have the most little frequency was 40 with the number of the score was 1 student. The conclusion was the student with the median score was more dominant than the highest and the lowest score.

b.) Post-test Scores of Control Class

The class with conventional method (lecture method) was control class. The researcher conducted post-test for this class after the treatment. The control class was 8F that consisted of 25 students. There was an absent student, so there were 24 students took the post-test.

Furthermore, the table 4.10 showed the test result which analyzed the scores of **control class**'s students in post-test as follow:

Table 4.10: Post-test Scores of Control Class

Students (Y)	Post-test
AIDI	52
ALY	60
ANG	76
ANI	64
CI	68
DHE	84
DI	68
FIL	44
GLA	40
IRM	52
IRT	68
ITN	80
LUT	68
MEL	40
NABQ	64
NAB	60
NAD	68
NAJ	64
NEZ	84
NUR	S
PUT	56
RIK	56
UNI	56
ZF	80
ZQ	64

In Table 4.10 showed the post-test scores from 24 students of control class. The minimum score was 40 and the maximum score was 84. Then, the result of post-test from VIII F class was presented in following table, Table 4.11:

Table 4.11: Descriptive Statistics Post – test of Control Class

Statistics

post_test_co

N	Valid	24
	Missing	0
Mean		63,17
Std. Error of Mean		2,548
Median		64,00
Mode		68
Std. Deviation		12,482
Variance		155,797
Range		44
Minimum		40
Maximum		84
Sum		1516

In Table 4.11 showed the data from control class's post-test score which was from 24 students. Mean score was the total scores which divided by number of data and mean of the data was 63.17. Next, median score was defined as the value of numbers divided higher half of the population, sample of data, or possibility distribution of lower half. The median score was 64.00. Then, there was mode as the most frequently appeared element in a given. The number of mode was 68. Next, there were set of scores formed from population that was called as variance. The number of variance was 155.797. Besides that, there was deviation standard which was a deviated measurement in a group as a whole that indicated the extent of quantity. The deviation standard that can be seen from the table was 12.482. Then, the distance between the highest value and the lowest one, called as range which was 44. The maximum score was 84 and the minimum score was 40. The total score of the data was 1516.

According to Table 4.11, the data was also presented in table of frequency distribution which showed in Table 4.12 below:

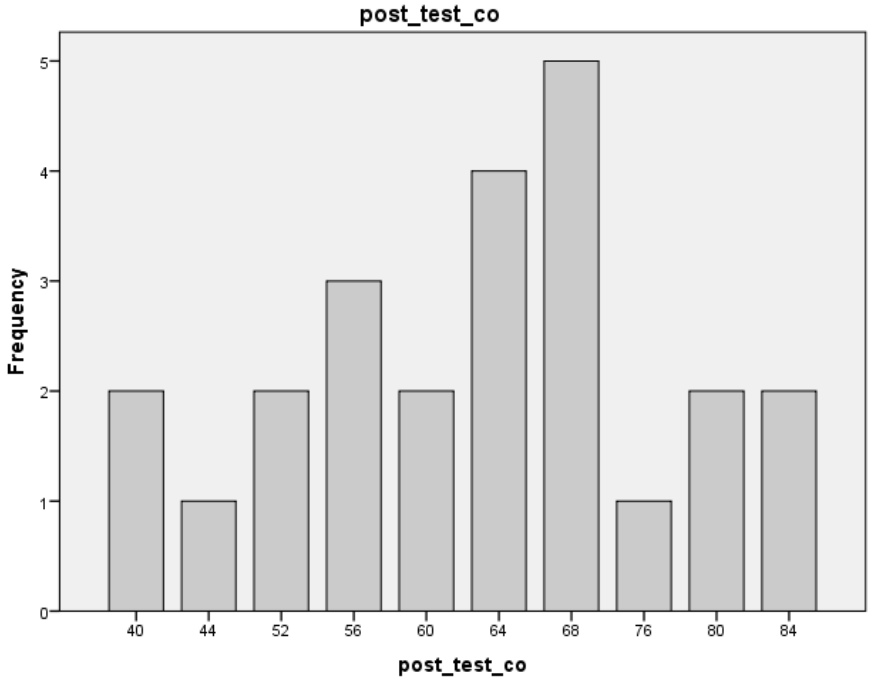
Table 4.12: Frequency Distribution Post – test of Control Class

post_test_co

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	40	2	8,3	8,3	8,3
	44	1	4,2	4,2	12,5
	52	2	8,3	8,3	20,8
	56	3	12,5	12,5	33,3
	60	2	8,3	8,3	41,7
	64	4	16,7	16,7	58,3
	68	5	20,8	20,8	79,2
	76	1	4,2	4,2	83,3
	80	2	8,3	8,3	91,7
	84	2	8,3	8,3	100,0
	Total	24	100,0	100,0	

From the Table 4.12, it showed the frequency of the scores. Apparently, two students have the score of 40 with the percent 8.3%, a student got the score of 44 with the percent 4.2%, 2 students got the score of 52 with the percent 8.3%, 3 students got the score of 56 with the percent 12.5%, 2 students got the score of 60 with the percent 8.3%, 4 students got the score of 64 with the percent 16.7%, 5 students got the score of 68 with the percent 20.8%, a student got the score of 76 with the percent 4.2%, 2 students got the score of 80 with the percent 8.3%, and 2 students got the score of 84 with the percent 8.3%. That table was a tool of presentation as a media to show the consisting of data frequency, the cumulative percent, and the distributed frequency in percentage.

Based on the frequency distribution of post-test data, the diagram from Table 4.12 was presented as following picture:



Picture 4.4

The Diagram of Post – test Scores of Control Class

According to the Picture 4.4, the score that have the greatest frequency was 60 with the number of score was 7 students. Then, the score that have the most little frequency was 40 with the number of the score was 1 student. So, the student who got the median score was more dominant than the highest and the lowest score.

From the descriptive statistic (data) of pre-test and post-test both of classes, this study tried to compare their mean score which the experimental was lower than the control in pre-test. It was the same condition with the mean score of experimental in post-test which lower than the control class. But, between both

of classes**,** only the experimental class that had an increasing in the mean score before and after getting the treatment.

Beside that, from the pre-test and post-test, the researcher knew that the control class got the score higher than the experiment class from each test. It was known from the sum and the mean score both of classes. In pre-test, the experimental class got the sum 1290 with the mean score 53.75 and the control class got the sum 1610 with the mean score 64.40. While from post-test, the experimental class got the sum 1356 with the mean score 56.50 and the control class got the sum 1516 with the mean score 63.17. The highest score in control class reached score of 84, while in experimental class only reached score of 76 for post-test. It showed that the ability both of classes were not equal. The ability of control class (8F) was better than the ability of experimental class (8G) because the achievement of control class was higher. So, there was difference from the achievement both of classes.

B. Analysis of the Data

In previous sub-chapter, the researcher knew the central tendency and the variability from the results of pre-test and post-test both of classes. Based on the descriptive statistics, the researcher determined the mode, median, mean, standard deviation, and range of the scores. From the mode score, median score, and mean score, the researcher determined the normally distribution of the data. It was seen from the frequency or the picture of bar diagram. From central tendency of the data can be known that only experimental class's post-test score was normally distributed. But in this sub-chapter, the researcher conducted analysis in

inferential statistics which was fulfillment test consists of normality testing and homogeneity testing.

1.) The Results of Normality Testing

The researcher conducted normality testing by *Kolmogorov-Smirnov* technique in *IBM SPSS 22.0 versions for Windows* with the criterion $\rho > 0.05$. This test was conducted to find the sample of the data whether obtained from the population that distributed normally or not. Normally distribution was requisite for parametric research before testing homogeneity and t-test. This test was taken from the data of pre-test score which both of sample classes. The result of normality testing was shown in Table 4.13:

Table 4.13:

Normality Test Result from Pre-test Scores

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Group	Statistic	Df	Sig.	Statistic	Df	Sig.
Pre_test	kelas 8G_ex	,186	24	,031	,869	24	,005
	kelas 8F_co	,164	25	,082	,904	25	,022

a. Lilliefors Significance Correction

In Table 4.13 showed the significance value of normality testing from pre-test score both of classes. Remembering the requirement before, if the significance scores of *(Asyim. Sig.)* > 0.05, so the data was normally distributed (of population). On contrary if *(Asyim. Sig.)*< 0.05, it was not normally distributed (of population). The significance of normality of control class was 0.082. It was more than significance level 0.082 > 0.05. Then, the

significance of normality of experimental class was 0.031. It was less than significance level 0.031 < 0.05. So, the data was not normally distributed.

Table 4.14:

Normality Test Result from Post-test Scores

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
post_test_ex	,108	24	,200*	,973	24	,747
post_test_co	,141	24	,200*	,956	24	,362

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

In Table 4.14showed the significance value of normality testing from post-test score both of classes. Remembering the requirement before, if the significance scores of *(Asyim. Sig.)* > 0.05, it meant the data was normally distributed (of population). On contrary if *(Asyim. Sig.)* < 0.05, it was not normally distributed (of population). The significance of normality of control class was 0.200. It was more than significance level 0.200> 0.05. Then, the significance of normality of experimental was 0.200. It was also higher than significance level 0.031 < 0.05. So, the data of post-test was normally distributed.

Although, the significance of normality test from control's pre-test score was normal distribution, but the data from experimental pre-test score was not from normal distribution. While, the result of normality testing both of classes in post-test score were normally distributed. So, the data of this research was not normally distributed.

1.) The Results of Homogeneity Testing

After normality testing, then the researcher conducted homogeneity testing by using *IBM SPSS 22.0 version for Windows*. Homogeneity test was applied to analyze the sample variance of the data whether come from homogeneous population. Homogeneity test was used to find out whether the data was homogeneous or not. The results of homogeneity test from post test scores were presented in Table 4.15:

Table 4.15:

Homogeneity Test Result from Post-Test Scores

Test of Homogeneity of Variances

Reading Comprehension

Levene Statistic	df1	df2	Sig.
1,627	1	46	,209

In Table 4.8 showed the significance value of homogeneity testing from post-test result both of classes, it was 0.209. If the value of (Sig.) $\rho > 0.05$ it showed that the population of data has the same variant (homogeneous). Then, if the value of (Sig.) $\rho < 0.05$, it showed that the population of data has not the same variant population (not homogeneous). Meanwhile, the significance indicated that $0.209 > 0.05$. It meant that the data was come from homogeneous population.

From the inferential statistics (both of tests), the conclusion was the data from scores of both classes were not normally distributed and homogeneous. It was shown from the significance value of pre-post-tests result in normality test were higher than 0.05. But, only experimental pre-test

score was lower than 0.05. Then, the result in homogeneity testing was higher than 0.05. The researcher had to ensure that the data was normally distributed and homogeneous, because to do hypothesis testing was based on the testing result of normality and homogeneity. If all the data was normally distributed and homogeneous, then the researcher can do the hypothesis test by parametric test. Whereas, if one of the data was not from normal population and heterogeneous or showed one of them, the researcher could do the hypothesis test by non-parametric test. Because the data was homogeneous but not normally distribution, the researcher conducted non-parametric test to do hypothesis testing.

2.) N-Gain Score of Pre-test Scores

After getting statistical analysis of the data both of classes, the researcher found the pre-test's mean score which experimental was lower than control class. It was showed from Table 4.2 and Table 4.8 where the experimental had mean 53.75 and the control class had mean 64.40. Then, the result of Mann-Whitney U-test of pre-test both of classes was not equal in their reading comprehension ability (Table 3.5). So, the researcher should calculate N-Gain score of pre-test and post-test both of classes by using *IBM SPSS v.22 for Windows*.

The purpose of N-Gain score was to see the effectiveness using a method, strategy, or technique in this research. N-Gain score gained from students' scores of all tests included experimental and control class. From N-Gain score, the researcher knew whether the treatment was effective or not.

Based on the Category of N-Gain Score, if the ρ N-gain > 0.70, it meant the treatment was High or effective to use. Then, if the ρ N-gain < 0.30, it meant the treatment was low or less effective to use. According to Sundayana (2015: 151) the category of N-Gain score as follows:

Table 4.16: N-Gain Score Category

N-Gain Score	Category
N-Gain > 0.70	High
0.30 ≤ N-Gain ≥0.70	Average
N-Gain < 0.30	Low

Furthermore, the result of N-Gain Score of experimental and control class was displayed in Table 4.17:

Table 4.17: The Result of N-Gain Score

No	Students' Name	Experiment Class	No	Students' Name	Control Class
		N-Gain Score (%)			N-Gain Score (%)
1	ANJ	12	1	AIDI	4
2	CYN	0	2	ALY	33,33
3	DEV	20	3	ANG	52
4	DYA	-20	4	ANI	-80
5	ENG	-30	5	CI	-60
6	FAI	12	6	DHE	20
7	FEL	26,67	7	DI	36
8	HEL	-100	8	FIL	-86,67
9	INT	33,33	9	GLA	-20
10	JES	-33,33	10	IRM	-140
11	KHO	20	11	IRT	20
12	NAB	-86,67	12	ITN	50
13	NAD	46,67	13	LUT	20
14	NATH	-20	14	MEL	-50
15	NID	33,33	15	NABQ	-20
16	RE	-20	16	NAB	-33,33
17	REY	-20	17	NAD	-6,67

18	RIZ	10	18	NAJ	-20
19	SEM	4	19	NEZ	20
20	SIL	12	20	NUR	-150
21	USA	13,33	21	PUT	12
22	VAN	-20	22	RIK	-10
23	WAH	25,71	23	UNI	-10
24	WIJ	46,67	24	ZF	0
Mean		-1,4286	25	ZQ	-20
Min.		-100,00	Mean		-17,5733
Max.		46,67	Min.		-150,00
			Max.		52,00

Table 4.17 showed the mean of experimental class was -1.4286or equal to -1.43. It meant the N-Gain score of experimental class was lower than 0.30 (-1.43<0.30). Then, the mean of control class was -17.5733or equal to -17.5. It meant the N-Gain score of control class was less than 0.30 (-17.5 <0.30). It could be summed up that the treatments (CSR method and conventional method) were less effective to be applied in both of classes.

Furthermore, to see the significant difference and how high the differences of students achievement in reading comprehension in recount text which taught by different method, the researcher conducted normality testing of N-Gain score. Normality of N-Gain score had to be calculated to know whether the data of N-Gain score was normally distribution or not. It had to applied because between both of classes were not equal. The normality result of N-Gain score can be seen in Table 4.18, as follows:

Table 4.18: The Normality Testing of N-Gain Score

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Group	Statistic	Df	Sig.	Statistic	Df	Sig.

NGain_score	1_ex	,164	24	,094	,887	24	,012
	2_co	,202	25	,010	,902	25	,020

a. Lilliefors Significance Correction

Based on Table 4.18, it could be known in *Kolmogorov-Smirnov* column at (Sig.) value of normality testing that the experimental had 0.094 and the control class had 0.01. It was indicated, based on the significance level (0.05) the experimental class was higher that 0.05 (0.0 94>0.05) and the control class was lower than 0.05 (0.01<0.05). So, the data of N-Gain was not normally distributed.

3.) Results of Hypothesis Testing

The next step was hypothesis testing. It was the most important calculation to resolve the research problem. Hypothesis testing purposed to see whether there was difference achievement or not between the students which taught by Collaborative Strategic Reading method and conventional method. Previously, the data was not from normal population, so the researcher did not need to do homogeneity test from students' N-Gain score. The researcher used *Mann-Whitney U-Test* in *IBM SPSS 22.0 versions for Windows* to find the-$_{value}$ in this research. The criteria for hypothesis testing are presented as follows:

- If significance of ρ-$_{value}$> 0.05 = the H_0 was accepted.
- If significance of ρ-$_{value}$ < 0.05 = theH_0 was rejected / the $H_1$was accepted.

H_0 : There is no significant difference in the achievement of students in reading comprehension taught with and without the Collaborative Strategic Reading (CSR) for second graders at MTs Ma'arif NU Kota Blitar.

H_1 : There is significant difference in the achievement of students in reading comprehension taught with and without the Collaborative Strategic Reading (CSR) for second graders at MTs Ma'arif NU Kota Blitar.

Before determining the hypothesis testing, the researcher concluded from the tests before, that N-gain score was not normally distributed (see Table 4.12). So, this research used non-parametric test to do hypothesis testing. The researcher used *Mann-Whitney U-test* to know the significant difference in achievement of students in reading comprehension taught with and without the Collaborative Strategic Reading (CSR). According to Chojimah (2019: 58), Mann-Whitney is one of non-parametric test functioning to compare two means of independent samples. As such, Mann-Whitney was an alternative to conduct t-test in independent samples. Therefore, it was used when the samples with ordinal or interval scale was not normally distributed. The result of hypothesis testing showed in Table 4.19, as follows:

Table 4.19: The Result of Hypothesis Testing of N-Gain Score

Test Statistics[a]

	NGain_score
Mann-Whitney U	255,500

Wilcoxon W	580,500
Z	-,894
Asymp. Sig. (2-tailed)	,371

a. Grouping Variable: group

From the Table 4.19, the researcher made the interpretation for hypothesis testing that ρ -value *Asyim. Sig. (2-tailed)* was 0.371. Because this research used one-right-tailed, the value was divided into two = 0.1585. It meant that ρ-value was higher than the significance level 0.05 (0.1585>0.05). In consequence, the result was rejecting the alternative hypothesis (H_1) and accepting the null hypothesis (H_0). It meant there was no significant difference on students' achievement of reading comprehension taught with and without using Collaborative Strategic Reading (CSR). In contrary, the alternative hypothesis that stated there was significant difference on students' achievement of reading comprehension taught with and without using CSR for second graders at MTs Ma'arif NU Kota Blitar was rejected.

C. Discussion of the Results

The last section in this chapter was discussion or interpretation of the result in this research. The discussion was started from test as instrument of this research to get the data. The researcher conducted the test before and after giving the treatment for experimental class and conventional method for control class. The test was in multiple choice forms which appropriate with the lesson plan and the syllabus.

The researcher compared that the experimental class's mean score was lower than control class in pre-test. It was same condition with the mean scores in

post-test, where the experimental class was lower than the control class. whereas only the experimental class that had an increasing in the mean score before getting the treatment and the after. Besides that, the researcher knew that the scores of control class were higher than the experiment class on each test. It was known from the sum and the mean both of classes, which in pre-test, the sum of experimental class was 1290 with the mean score 53.75 and the control class got the sum 1610 with the mean score 64.40. While in post-test, the sum of experimental class was 1356 with the mean score 56.50 and the control class got the sum 1516 with the mean score 63.17. The highest score of control class reached score of 84, while the experimental class only reached score of 76 in the post-test score. It indicated that students' ability both of classes were not equal. The ability of control class (8F) was competent than the ability of experimental class (8G).

According to the analysis of the data in hypothesis testing with N-Gain scores, the finding was no significant difference on students' achievement of reading comprehension in experimental class taught by CSR method and control class by lecture method (conventional method). It could be seen from the significance ρ -value was higher than 0.05 (0.1585> 0.05). It rejected H_1 and accepted H_0. The researcher used N-Gain score to conducted hypothesis testing because between control class and experimental class were not equal. It was also known by the pre-test's mean score of experimental class (53.75) and control class (64.40).

Furthermore, from this research could be known that applying CSR method in teaching students' reading comprehension gave an influence on students' achievement especially for recount text. Although based on the N-Gain score, Collaborative Strategic Reading (CSR) was less effective to be applied for second grade students at MTs Ma'arif NU Kota Blitar. When the treatment was done, the researcher found some obstacles because students got difficult to understand each strategy. It caused by many strategies were done directly and a short time. The treatment was conducted in two days. Apparently, it made students cannot catch the material clearly. So, the students' mastery on the material was still minimum. Although, the student have been actively asked. The researcher needed more time to convince every strategy. But, the students said from CSR, they found something new in learning.

While, the results of previous study said that Collaborative Strategic Reading method contributed in improving the achievement of the students in reading comprehension after getting the treatment process. Like the previous studies of this research, Mohd. Rafi Riyawi (2018) that he found there is a significant rise on achievement of students after teaching by Collaborative Strategic Reading (CSR) method. The previous studies showed the outcome of this research that CSR method that was positively effetivefor teaching students' reading comprehension. But, this research verified the theory that Collaborative Strategic Reading (CSR) was less effective to teach students' reading comprehension in recount text for second grader at MTs Ma'arif NU Kota Blitar. It did not too help in increasing the students' reading achievement.

Besides that, the finding another study that also conducted by Indah Fadhilah Rahman (2015) that she focused on intended to enhance students' motivation and improve their students' proficiency in reading comprehension was considered necessary. The researcher modified the strategy for learning by using Collaborative Strategic Reading in reading skill especially for reading comprehension and also beginner. It was appropriate with the second grader of junior high school as beginner in reading comprehension, to make them interested. Collaborative Strategic Reading has four varied steps to do in reading class. Each step has different concentration and goals to understand the content of a reading text. It was appropriate with the necessary in indicator of basic competence. Hence, the result of this research was positively responded by the experimental class. It meant realization of CSR showed change in students' reading comprehension achievement. The previous explanation was appropriate with the advantages of Collaborative Strategic Reading (CSR) that increasing achievement of student and academic, according to M. J. Z. Abidin (2012: 65).