

CHAPTER III

RESEARCH METHOD

The chapter is about methodology of the research to determine case study. This chapter presents of explanation about: Research Design; Data and Data Source; Population, Sample, and Sampling; Research Variable; Research Instrument; Validity and Reliability Testing; Data Collection Method; Treatment; and Data Analysis.

A. Research Design

This research used quantitative research design and quantitative approach to resolve the research problem. For analyzing the data, the writer chose a quasi-experimental. The quasi-experimental was an experiment which applied in two classes the first class was the subject to the control and the other class as an experimental class. Quasi-experimental was similar to randomized experiment that it involved an independent variable manipulation, but different with subjects were not assigned randomly for treatment group. Both of groups asked to conduct pre-test and post-test, but the treatment was only given to experimental class. The researcher chose quasi-experimental research design to see significant different of students' reading comprehension that taught with and without Collaborative Strategic Reading (CSR) and to know effectiveness of the treatment for second graders of MTs Ma'arif NU Kota Blitar.

Furthermore, this Quasi-experimental was conducted by nonrandomized control group or without randomization and both groups were balance. The researcher could not randomize both classes to reorganize in accommodation of a research. Hence, the researcher used two groups that already organized in classes. Quasi-experimental research design was also needed two groups which was become control group and experimental group. As Arikunto (2006: 311) assumed that, if conducting experiment research, the researcher entails two groups to carry out the experiment, the first is group which does not get a treatment or called by control group and group which gets a treatment or experimental group. The researcher asked English teacher for suggestions to determine which from both of classes becoming the experimental class and the control class. Based on the English teacher's advice, the class that was less competent better gets the treatment in this research (as experimental class). While the class that was excellent better become control class.

Group	Pre-test	Independent variable	Post-test
G	Y_1	X	Y_2
F	Y_1	-	Y_2

B. Data and Data Source

In this study, the researcher used data from second graders of junior high school level and conducted by quasi-experimental as the manner of the research to see that applying Collaborative Strategic Reading (CSR) was effective or not for teaching students' reading comprehension especially in recount text. The data were collected from second graders at Islamic Junior High School (MTs) of

Ma'arif NU Blitar which is located in Jl. Ciliwung No. 56, Bendo, Kec. Kepanjenkidul, Kota Blitar. The researcher has conducted the treatment of the research for two weeks and four times for meeting. It began on February 24th 2020 until March 3rd 2020.

The object of the variable was the second grader or eighth grade which were seven classes. The researcher selected two classes to become sample of this research (two groups). One of the groups which conducted by a treatment, was named as experimental group. Then, another group which not conducted a treatment was named as control group. This research obtained the data from the score of the test result, which there were test before treatment (pre-test) and after treatment (post-test). It purposes to see the students' significant differences of achievement on reading comprehension which applied by treatment and without treatment. The treatment class was used Collaborative Strategic Reading (CSR) method and lecture method -the teacher as centered in learning- (conventional method) was for non-treatment class. To know the research sketch, the researcher designed a form of Quasi-experimental research design. The following Table 3.1 showed the research design:

Table 3.1: Quasi-experimental research design

Group	Y¹	X	Y²
Experimental Class (8G)	Pre-test	Collaborative Strategic Reading	Post-test
Control Class (8F)	Pre-test	Conventional method	Post-test

C. Population, Sample, and Sampling

The researcher carried out the study on Islamic Junior High School (MTs) Ma'arif NU Kota Blitar. This school has 653 students which divided into 23 classes. Then, there were 9 classes of ninth grade, 7 classes of eighth grade, and 7 classes of seventh grade. Furthermore, the researcher required the population of this research, the sampling techniques, and the research sample.

- The total student in second grade of MTs Ma'arif NU for academic year 2019/2020 was the population of this research. The total number of research population was 185 students that were divided into 7 classes. The researcher chose student level in second grade of junior high school because the researcher wanted to apply Collaborative Strategic Reading (CSR) for junior high school level. Besides that, reading comprehension was appropriate with the purpose of this research where it was basic competence from syllabus on second grade of junior high school. Ary (2006: 167) argued a way that selecting number of individuals to become sample in the research which represents the population was meaning of sampling technique.

Table 3.2: The Total Number of Population

No.	Class	Number of Student
1.	8A	30
2.	8B	30
3.	8C	27

4.	8D	18
5.	8E	31
6.	8F	25
7.	8G	24
Total		185

- The sampling technique of this research was based on non-probability sampling because the school permit the researcher use of the certain classes. Therefore the researcher used Purposive Sampling, which includes as non-probability sampling. The researcher selected the sample because it was available for conducted research, convenient, and represents the characteristic of the purpose in this study. Moreover, the samples were appropriate with the purpose of this research. Based on Ary (2010: 156) for purposive sampling, researcher used representative or typical judgement to choose the sample elements from the population (referred as judgement sampling)
- Sample was a part of population that was being studied, observed, and experimented. According to Chojimah (2019: 3) sample was selected object to represent in quantitative study for approaching very large population. Besides that, based on Khotari (2004: 14), sample was items which selected constitute technically. There were seven classes of the second graders of MTs Ma'arif NU Kota Blitar. This research chose two classes, 8F and 8G as the sample. In 8F class consisted of 25 students, than 8G class was 24 students. The number of the research sample was 49

students. The researcher determined that the control class was 8F and the experimental class was 8G (according to pre-test result). It was also according to English teacher's advice.

Then, the research sample and the distribution of the treatment are presented in the following table, Table 3.3 and Table 3.4:

Table 3.3: the Research Sample based on Class

No.	Class	Number of Students	Total
1.	8F	25	49
2.	8G	24	

Table 3.4: The Treatment Distribution

Class	Treatment	Number of Students	Group
8G	Collaborative Strategic Reading	24	Experimental
8F	Conventional method	25	Control

Both of the classes were almost balanced and equitable as excellent class in MTs Ma'arif NU, but there was a slight difference. The sample of this research was equal assumed by English teacher, but it was not clearly equal. Making sure that two classes were equal or not, the researcher tried to process the data of pre-test score. It could be proven from the result of t-test.

The researcher conducted *Mann-Whitney U-test* in *IBM SPSS v.22 for Windows* to know both of classes were equal or not. It was caused the data from pre-test score was not normally distributed. The normality test both of classes showed 8F p -value was 0.082 and for 8G p -value was

0.031. It meant 8F was higher than 0.05 ($0.082 > 0.05$) and 8G was lower than 0.05 ($0.031 < 0.05$). Then, the result of Mann-Whitney showed 0.008. It meant that the result was less than significance level 0.05 ($0.008 < 0.05$). It indicated that there was variance of the data and the two classes were not equal. T-test result of pre-test scores was showed on the Table 3.5, as follow:

Table 3.5: T-Test of the Pre-test scores using Mann-Whitney U-test

Test Statistics ^a	
	Pre_test
Mann-Whitney U	171,000
Wilcoxon W	471,000
Z	-2,639
Asymp. Sig. (2-tailed)	,008

a. Grouping Variable: group

D. Research Variable

Variable was an aspect that supporting a study in which the change was a subject of the research. The variable was regarded the other variables that change as a result of the research. According to Ary (2006: 37), a characteristic or a construct which can be taken on different scores or values was called variable. Then, based on Khotari (2004: 33), variable was a frame of research which can make different values quantitatively. Hence, variable basically was characteristics which might change or different from type to type or individual to individual. Based on Ary also, there were two variable types: a.) Categorical variables/dichotomous variable, where we can't make range and nothing gradation, such as gender and religious; b.) Continuous variable, where variable

resulting in classification having gradation, such as achievement. This research was conducting by continuous variable, because using score achievement.

Besides that, based on the function in the research, Khotari (2004: 34), variable has divided into two parts:

a. Independent variable

The variable affected to dependent variable was termed as an independent variable. In this research, the independent variable was teaching students' reading comprehension using Collaborative Strategic Reading (CSR). This variable affected to students' achievement in reading comprehension. Characteristic or attribute which affects and influences an outcome (dependent variable) was termed independent variable, based on Creswell (2008: 116).

b. Dependent variable

Then, the variable which depends on the other variable or gets an impact from the other variable, it was called by dependent variable. In this research, the dependent variable was students' achievement of reading comprehension, because after implementing the treatment, the result could be changed. As well as according to Creswell (2008: 115), an attribute that is affected by independent variable termed of dependent variable.

E. Research Instrument

The research needed an instrument for collecting the data. The researcher used test about English reading text as instrument which consisted of 4 passages

and 25 questions of multiple choice form that have been created by the researcher. The 25 items of the test were tested before (try out the test) in third grade of junior high school students. After applying the try out, the researcher measured the validity and reliability of the test by using *IBM SPSS 22.0 version for Windows*. The researcher chose test in obtaining the data to see the achievement of students in reading comprehension mastery because more appropriate by using test. Ary (2010: 10) stated that a set of stimuli served to get a response of individual by numerical score that can be assigned was termed by test.

The test form which used to get impact after using treatment was called as post-test. In brief, post-test was conducted after giving treatment to the sample. The post-test purposed to measure and determine the result of change after treatment based on the students' score in reading comprehension. Not only treatment class, but also control class which used conventional method was held post-test.

For this objective, the researcher chose recount text which appropriate with the syllabus of English lesson for second grader of junior high school at second semester to conduct this research instrument. The material also was based on the syllabus which about teaching recount text section. The instrument (post-test) was recount text in multiple choice tests. In multiple choices, students have to choose one of the best answer each item. The format of post-test was recount text which there was twenty-five items. Therefore, the test purposed to find the difference son students' achievement with a clearly significant based on both of classes.

Before using instrument for collecting the data, firstly, the researcher should try out for the test. The tryout purposed to get the reliability score and the validity level of the instrument. The tryout was conducted and trialed in other different subject from the research sample and has learned about the topic. The instrument was tried out at 20 students of third grade of junior high school in an English course class. The researcher chose the third graders of junior high school based on their ability in recount text, because they ever learned recount text and almost near with the research sample.

F. Validity and Reliability Testing

1. Validity Test

Validity testing indicated and made certain about the instrument that it was appropriate to measure in a research. Based on Ary (2010: 225) defined validity was a standard of instrument measure what claimed to be measured. For this research, the researcher used construct validity as the criteria in comparing items text with the theory of the skill that there was accordance with the purpose or not. Then, as a non-empirical, some expert judgments defined that content validity as the extent the content of a test which was appropriate with the purpose of the test material. Validity testing was important to generate good and appropriate test as treatment in a research. According to Khotari (2004: 74) content validity is good when the instrument contains a representative sample of the universe.

a. Construct validity

In analyzing a test, there was construct validity that measure whether test has represented consistently with the theoretical propositions, exactly showed the material or not. It meant the test as the instrument was formed to concern aspects that were appropriated with the theory. This test instrument was consulted to English teacher of the school and the adviser.

Construct validity showed how far the test items that relevant with the theory. Construct validation was based on the compatibility any theory or hypothesis that try to explain observed phenomena. This study dealt with reading skill. It conducted reading skill with testing techniques for reading skill. Related to reading skill, according to Brown (2004: 190), he mentioned some techniques for testing-reading skill. Those are multiple-choice, picture-cued items, reading aloud, matching test, Gap-Filling test, written response, cloze test, C-test, short-answer test, summarizing test, and ordering test. From the techniques, the researcher determined that reading achievement was appropriate if conducting test by using multiple choice. The students were given some passages of text. Then, the students had chosen the one of correct answer from four choices. The choices were appropriate with understanding the meaning of passage or text. If student got many

correct answers, it meant student has good comprehension in this reading test.

Reading process was also done by readers to get the message or meaning which was delivered by authors word by word that could be known by reader. It meant, reading was an activity of getting meaning. Likewise comprehending a reading text, the reader must get the meaning to interpret the information which was conveyed by the writer. So, in this research, the test was appropriate with the theory because the students were asked to understand the meaning or information in it.

The researcher constructed the instrument appropriate with the test specification of reading comprehension which consisting basic competence, specific indicator, test type, number and total item test, and time. The specific indicator was relegated from basic competence to fulfill construct validity. The researcher chose recount text to become the object of instrument.

This instrument has been constructed according to the recount theory, such as the generic structure, the features, and the purposes. Construct validity of this instrument used the expert opinion/validation, especially from the English teacher that knows what the students' habit when learning English clearly. Besides that, the adviser has given corrections based on the test

specification and the basic competence. The experts who checked and gave opinions on the instrument were Ani Mar'atus Sholihah, S.Hum., as English teacher of MTs Ma'arifNU and Dr. Hj. Ida Isnawati, M.Pd., as the adviser and lecturer in IAIN Tulungagung. In addition, after consulting the instrument, the researcher was also given suggestion to try out the test before in which consisted of minimum 10 students.

b. Content validity

A certainty of the extent that indicate the content of a test was appropriate with the purpose of the test material, termed by content validity. Furthermore, a test or an instrument has to have compatibility and accuracy on the material or to the purpose. A content validity was indicated with the relevant objectives of the content and the test type of the test items. Based on Ary (2010: 225) assumed that validity was also done by teacher to check the test and assess whether content of the test as sample was appropriate with the aims to be stepped. The instrument has adjusted on the basic competency and competency Standerton the school. It based on syllabus and curriculum of second grader of English junior high school.

Content validity of the instrument purposed to make accuracy in collecting data. The instrument departed from blue

print, where the blue print made up the formation of the instrument. According to Ary also (2010: 227) stated to assure content validity in a test, teacher is better to prepare a “blueprint” which display the content domain included and emphasize the relative priority conveyed to every aspect of the domain. Before making a blue print, firstly was mentioning definition of pedagogical content knowledge. The second was establishing the aspects that want to be researched. The third was making questions that were tested based on the purpose. From that, blue print for the instrument can be formed.

An instrument was said appropriate in content validity, when the content of instrument indicated to represent sample of the indicators in syllabus. Therefore, the content validity of the instrument was appropriate and fulfilled the indicators in the syllabus. Before fixing the instrument, the researcher made a trial or try out the blue print (questions) to 20 students of third grade students’ of junior high school. The result of this try out became consideration in fixing the instrument later. Then, not only for fixing but also to establish the reliability and the validity of the instrument. The instrument was valid in term of content validity, as shown at Table 3.6:

Table 3.6: Table of Content validity

Main Competence	1. Mengolah, menyaji, dan menalar dalam ranah konkret (menggunakan, mengurai, merangkai, memodifikasi, dan membuat), dan ranah abstrak (menulis, membaca, menghitung, menggambar, dan mengarang) sesuai dengan yang dipelajari di sekolah dan sumber lain yang sama dalam sudut pandang/teori.
Basic Competence	4.14 Menangkap makna teks <i>recount</i> lisan dan tulis, pendek dan sederhana, tentang kegiatan, kejadian, peristiwa.
Indicators	1. Siswa mampu memahami isi cerita dalam teks <i>recount</i> dengan membaca dan menanyakan tentang cerita pendek dan sederhana, tentang kegiatan, kejadian, peristiwa yang pernah dialami, dengan struktur teks yang tepat.
Technique	Achievement test (multiple choice)
Instrument	Pre-test and Post-test

According to Table 3.6, test as the instrument was said appropriate with the purpose of basic competence and indicators in Curriculum 2013 syllabus. It meant the test has content validity which appropriate with the syllabus. The syllabus showed the student targets for their abilities of reading comprehension and the correct structures in recount text. The ‘blueprint’ was appropriate with the indicator of the basic competence. It was made appropriate with the operational of the indicator. The test items gave four choices and students chose the correct one by comprehending the content of the text before.

2. Reliability Testing

Reliability indicated that the instrument was reliable and reasonable for collecting the data. Reliability indicated that the test instrument can be trusted in collecting the data process and got the approved data in this research. So, the instrument had to be a very good quality in collecting data sample. As said by Arikunto (2010: 221) reliability proves that the instrument can be trusted to collect and get data because the instrument has a very good quality in data collection. Then, according to Ary (2010: 236), reliability testing was needed to measure the instrument, because it was the consistency degrees of measuring no matter what was measured. From reliability, an instrument could be said reliable if the test provided consistent results.

Not far from validity, a valid instrument was almost reliable instrument. Reliability was easier to be assessed than validity, so it was not as valuable and complicated as validity. Such as according to Khotari (2004: 74) reliability is not needed to be a valid instrument. In contrary, reliability in instrument is contributed to validity. A test was also reliable if it was dependable. It meant that reliable allow near or even same result in any occasion of the test, whenever and whoever the test was administered.

For reliability testing, the researcher used *IBM SPSS 22.0 version* to analyze the test instrument whether reliable or not. The criteria of reliability's rating score can be seen from the following table, while the instrument of the test can be seen in appendix 2.

Table 3.7:
Table Classification of Reliability Test based on Cronbach Alpha

Reliability Test Coefficient	Classification
0.00 – 0.20	Poor / less reliable
0.20 – 0.40	Fair / rather reliable
0.40 – 0.60	Moderate / quite reliable
0.60 – 0.80	Good / reliable
0.80 – 1.00	Very Good / very reliable

Then, the reliability test result of try out score showed in Table 3.8:

Table 3.8: Result of Reliability Testing

Case Processing Summary			
		N	%
Cases	Valid	20	100,0
	Excluded ^a	0	,0
	Total	20	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics	
Cronbach's Alpha	N of Items
,428	25

If the researcher wanted to see the instrument was reliable or not, it could be known by analyzing from Cronbach Alpha's classification about reliability testing. Based on Cronbach Alpha, if the result was under 0.60,

it meant not reliable, but quite reliable. Then, if the result was upper 0.60, the instrument was reliable. Based on Cronbach Alpha's Classification, the reliability score of the instrument was 0.428, so the test instrument was quite reliable.

G. Data Collection Method

Data collection method was techniques to get data in a research. It purposed to get data that needed for conducting scientific research. For this research, the researcher conducted pre-test and post-test to both of control class and experimental class. Firstly, pre-test was administered by the researcher for the students. From pre-test, the researcher wanted to know the basic ability of students (in achievement) about reading comprehension before applying Collaborative Strategic Reading (CSR) method. According to pre-test score, the researcher also knew both of classes were equal or not. Then, the next was given post-test which conducted to see the effectiveness using CSR method on students 'achievement of reading comprehension during certain period.

In collecting data, the researcher needed 4 meetings to conduct Collaborative Strategic Reading (CSR) method for each class, as follow:

- Day 1 (February, 24th 2020) : Pre-test and introduction to the material (treatment 1)
- Day 2 (February, 25th 2020) : Implement the CSR method in experimental class (treatment 2), and conventional method in control class
- Day 3 (March, 2nd 2020) : Implement the CSR method in experimental class (treatment 3), and conventional method in control class

- Day 4 (March, 3rd 2020) : Conducting post-test and sharing feedback

Pre-test : test which given before the researcher taught by the CSR method (before treatment). In pre-test, the researcher applied multiple choice forms about reading comprehension in descriptive text which consisted of 10 items. The researcher chose descriptive text because this material was learnt by the students. Pre-test was carried out to measure how far the students mastered the reading comprehension in English text and to know both of classes were equal or not. Pre-test was conducted in 25 minutes.

Post-test : test which given after the researcher taught by the CSR method (after giving the treatment). In post-test, the form of test was multiple choices about reading comprehension in recount text and consisted 25 items. The researcher chose recount text because this material was appropriate with the syllabus of second grade of junior high school. Conducting post-test was to know how far students mastered reading comprehension aspect after applying the treatment. It was done to calculate the final score measuring the differences achievement before and after giving treatment. Post-test was held for 60 minutes. The first post-test was given for VIII-F (control class). It was held on Tuesday, March 3rd 2020 at 08.20-09.20 a.m. The second post-test was for VIII-G (experimental class). It was also held on Tuesday, March 3rd 2020 at 10.00-11.00 a.m.

H. Treatment

The treatment of this research was Collaborative Strategic Reading (CSR). It only gave for experimental class (8G). The researcher not only explained the treatment but also showed how to use Collaborative Strategic Reading (CSR). In

using CSR method, there were some procedures that conducted by the researcher, as follow:

Table 3.9: The Procedures of CSR

Before Reading	Preview	<ol style="list-style-type: none"> 1. The researcher explained about recount text. The points that explained were the purpose, the features and convention, and the structures of recount text. 2. The researcher explained about CSR and what should the students do in this treatment. 3. The researcher asked the students to create groups which consisted of 5-6 students. 4. The researcher showed a simple text in recount text about an unforgettable experience. 5. The researcher conducted brainstorming to the students which tried to know about the topic would be learned. 6. The researcher gave instruction to predict what the students would be read about the topic of the text.
During Reading	Click and Clunk	<ol style="list-style-type: none"> 7. The researcher tried to lead the students reading the text. The text consisted of between 4-5 short paragraphs. 8. The researcher asked the students to find the difficult word if any parts of the text that were hard to understand. The difficult word called as clunk word. 9. The researcher asked the students to open the clunk cards. The clunk cards were about the steps what the students should do. There were 3 clunk cards that used, as follow: <ol style="list-style-type: none"> a. "Reread the sentence without the word. Think about what would make sense in the sentence". The students were asked to guess the key ideas which helping the students to understand the word. b. "Reread the sentence with the clunk word, and then read the sentences before and after the clunk word and look for clues". The students were asked read the sentences before and after the clunk word to make a sense in the word or tried to look for the clues about the clunk word. c. "Look for a prefix or suffix in the word that might help, especially in verbs". The students were asked to break the word from the prefix/suffix to know the basic of the word. 10. The researcher instructed the students to write the clunk

		word and the meaning in the clunk sheet. 11. Then, the students tried to understand the content of the text based on the finding of the clunk words.
	Get the Gist	12. The researcher led the students to determine 5Ws+1H; what was the text about?, where did the experience take place?, who were in the text?, when the experience happened?, why the experience happened?, and how the experience happened? 13. The researcher asked the students to arrange the answers of 5Ws+1H and determine the main idea and the structures of each paragraphs. The students wrote in the gist sheet.
After Reading	Wrap Up	14. The researcher followed up the students whether they have understood and got the information from the text or not. 15. The researcher asked students to make a summarization from the text based on their own comprehension. 16. Last, the students were asked to present their summaries, it can be retelling the text.

I. Data Analysis

A systematically process in applying logical or statistical techniques to describe, and evaluate the data was termed by “data analysis”. Data analysis process used analytical and logical reasoning to obtain the data information. The data analysis aimed to find out meaning of data so the knowledge that established can be used to give clues for a decision. For this data analysis, the researcher conducted *t-test* to see the significant difference of students’ scores. It was obtained from students’ pre-test and post-test score. Before revealing the hypothesis result, firstly the researcher explained the data in descriptive statistics. Then, secondly the researcher measured the data in inferential statistics, that is normality and homogeneity test, because the data should be normally distribution

and homogenous. Thus, the researcher conducted the inferential statistics and *t-test* used *IMB SPSS v.22* software. Before conducting *t-test*, the researcher should determine to use parametric test or non-parametric test, which depending on the result of normality test and homogeneity test.

1. Normality Test

After getting the data, the researcher has to do normality test before determining the hypothesis. Normality test was committed towards the scores of experimental class and control class. The researcher wanted to see the sample data (two groups) came from normal distributed population or not by normality test. Conducting normality test was helped by *IBM SPSS 22.0 version for Windows* with requirement: the data was come from the normal population, if the significance value of normality test was (*Asym Sig 2 tailed*) > 0.05 . While, if the significance value was (*Asym Sig 2 tailed*) < 0.05 , so the data was not come from normal population.

2. Homogeneity Test

After getting significance value of normality, if it indicated that the data was normal distribution or came from normal population, then the researcher had to do homogeneity test. If the data did not distribute normally or not come from normal population, so the researcher did not need conducting homogeneity test. The homogeneity test was conducted for knowing the similarity conditions or population both of groups. The researcher also used *IBM SPSS 22.0 version of Windows* to get the result

of homogeneity test by *Analyze-Compare Means-One Way ANOVA*. The data was homogeneous if the significance score of (*Asym Sig 2 tailed*) > 0.05 , but if the significance score (*Asym Sig 2 tailed*) < 0.05 , the data was not homogeneous or heterogeneous. If one of the data was not homogeneous, so the researcher could not use parametric test, but should use non-parametric test.

3. Hypothesis Testing

The next step was hypothesis testing. Some supposition or assumption that proved or disproved was termed hypothesis, as conveyed by Khotari (2004: 184). Meanwhile, hypothesis testing was a strategy to determine whether the data samples offer supporting for general hypothesis that can be made. In hypothesis test, the researcher used *Independent Sample t-test* (homogeneous and normally distributed) to find out the result which showed there are differences between two variables or not. This research used *IBM SPSS 22.0 version* program to conduct the hypothesis testing or obtain the result ($p\text{-value}$) in this study. After getting the $p\text{-value}$, the researcher compared $p\text{-value}$ and $\alpha = 0.05$.

The criteria of hypothesis testing were based on the significance degree 0.05. The conclusion was obtained according to:

- If significance of $p\text{-value} > 0.05$ = the H_0 was accepted.
- If significance of $p\text{-value} < 0.05$ = the H_0 was rejected / the H_1 was accepted

H₀ : 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.

H₁ : 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.

In data analysis, if the researcher wanted to see whether there was significant difference both of classes or not, the researcher calculated the N-Gain score of pre-test and post-test both of classes. The significant difference was based on achievement of students in reading comprehension taught with and without using Collaborative Strategic Reading (CSR). Then, the researcher conducted *Mann-Whitney U-test for N-Gain score* by using *IBM SPSS 22version for PC*. This study conducted N-Gain score because those classes were not equal. It was shown from mean score of experimental and control class in pre-test. If the normality test of N-Gain score showed that the data was normally distributed, the researcher could use *Independent Sample T-test for N-Gain score*. But, if the normality test of N-Gain score showed that the data was not normally distribution, the researcher could use *Mann-Whitney U-test*.