

CHAPTER III

METHODS

This chapter presents methods used in conducting the research. It consists of research design, research variable, population, sample, sampling technique of the study, research instrument, validity and reliability testing, normality and homogeneity of testing, data collecting method, procedures of the study and data analysis.

A. Research Design

Qualitative and quantitative are types of approaches that are very popular in research. Qualitative produces data in the form of description, while quantitative produces numeric data. Brown and Rodgers (2002) stated that the qualitative research approach is typically the label for non-numerical research and quantitative research approach is a numerical research, a data conversion is needed for this approach. In another analysis, Grotjahn (1987) he stated that the experimental or non-experimental data collection method, qualitative or quantitative is the type of data the resulted, and statistical or interpretative is the type of analysis conducted on the data.

In this study, the researcher used quantitative research approach with the design was experimental research. According to Ary et al. (2010) Experimental research involves a study of the effects of systematic manipulation of one variable (s) on another variable. This design is important because it determines the study's internal validity, the ability to reach valid conclusions about the effect of the experimental treatment on the dependent

variable. To know the effectiveness of Flipped Classroom on students' reading recount text achievement, the researcher used quasi-experimental design with nonrandomized control group, pretest-posttest design.

Experimental Nonrandomized control group, pretest-posttest design was conducted with two groups, they were experimental group and control group. The experimental group was taught by using Flipped Classroom and control group was taught by using a conventional strategy. The two groups were measured or observed not only after being exposed to a treatment of some sorts but also before. So, the researcher compared the two groups by giving a pre-test and post-test to measure the different attained mean scores in reading recount text achievement before and after each group was given the pre-determined treatment. The design is illustrated as follows:

Table 3.1 Nonrandomized Control Group, Pretest-Posttest Design adapted from Ary et al. (2010)

Group	Pre-test	Independent Variable	Post-test
E	Y_1	X	Y_2
D	Y_1	-	Y_2

Notes:

E : Experimental group

D : Control group

Y_1 : Pre-test in experimental group before treatment

Y_1 : Pre-test in control group

Y_2 : Post-test in experimental group after treatment

Y_2 : Post-test in control group

X : Treatment using Flipped Classroom for experimental group

- : Treatment using a conventional strategy for control group

Based on the design above, the procedures of experimental research using nonrandomized control group, pretest-posttest design were:

1. Selecting two classes as the samples of the study; E and D classes of the eighth grade students at MTsN 4 Tulungagung.
2. Administering a pre-test with a purpose of measuring reading recount text achievement of the eighth grade students at MTsN 4 Tulungagung before being taught by using Flipped Classroom for the experimental group and by using a conventional strategy for control group.
3. Giving the treatment applying Flipped Classroom to the experimental group and using a conventional strategy to the control group of the eighth grade students at MTsN 4 Tulungagung.
4. Administering a post-test to the two groups to measure their reading recount text achievement after each group was given different treatment.

B. Population, Sampling, and Sample

1. Population

According to Ary at al. (2010: 148) a population is defined as all members of any well-defined class of people, events, or objects. So, the population includes all of the subject or object where the data will be gathered. In this research, the population was all of the eighth grade students of MTsN 4 Tulungagung which consisted of 393 students.

Those were divided into nine classrooms. Class A, B, C, D, E, F, G, H, I,

J. As the Table 3.2 below:

Table 3.2 Population of Research

No	Class	Gender	
		Male	Female
1	VIII A	14 students	16 students
2	VIII B	12 students	22 students
3	VIII C	18 students	19 students
4	VIII D	8 students	26 students
5	VIII E	8 students	27 students
6	VIII F	16 students	21 students
7	VIII G	17 students	22 students
8	VIII H	18 students	18 students
9	VIII I	16 students	20 students
10	VIII J	11 students	20 students
Total students		349 Students	

2. Sampling Technique

Sampling is a technique for taking samples from populations. This sampling is representative of the population. Non-probability sampling in the form of purposive sampling was used in this study. According to Cohen et al. (2007) in the purposive sampling technique, the sample is satisfactory to specific needs. As its name suggests, the sample was chosen for a specific purpose. The chosen classes were the classes that had average ability in reading. Hence, these classes were identified as normal classes, meaning that the selected classes would tend to develop when they were given a stimulation or a treatment by using Flipped Classroom and a conventional strategy.

3. Sample

Sample is a part of population which is used to generalize the finding due to the researchers' limitation of time, capability, to take all members of population. Ary et al. (2010: 148) stated that the sample is the small group that is observed. Sample is part of population that is being studied. It means that a good sample must represent the entire populations. Based on the sample selection criteria above, the researcher selected two classes; VIII-D and VIII-E classes as samples. Then, VIII-D as a control group was taught by using a conventional strategy consisted of 34 students, and VIII-E as an experimental group that was taught by using Flipped Classroom consisted of 35 students.

C. Research Variable

According to Ary et al. (2010: 37) variable is a construct or a characteristic that can take on different values or scores. Or to say otherwise, the variable is characteristic of research subject which might change or be different from one individual to another or from time to time. There were two kinds of variables used in this research, they were:

1. Independent Variable

According to Ary et al. (2010: 266) independent variable is manipulated (changed) by the experimenter. It means that the independent variable could have affected to dependent variable. Independent variable in this research was Flipped Classroom and it was symbolized in this research as "X".

2. Dependent Variable

Ary et al. (2010: 266) stated that the variable on which the effects of the changes are observed is called the dependent variable, which is observed but not manipulated by the experimenter. So, dependent variable was affected from independent variable. Dependent variable in this research was achievement of reading recount text, and it was symbolized in this research as “Y”.

D. Research Instrument

Research instrument is a tool used to collect the data. This research used instrument in the form of test. The test had important in this research. Developing a test was one of steps in conducting this research. Ary et al. (2010: 201) stated that test is a set of stimuli presented to an individual in order to elicit responses on the basis of which a numerical score can be assigned. This score based on the representative sample of the individual's behavior which is an indicator of the extent to which the subject has the characteristic being measured. The test used of this study was Reading test. The test format was objective test in the form of multiple choice. The researcher gave 25 questions of multiple choice. They were two kinds of test, the test were pre-test and post-test. The test was administered twice, they were pre-test and post-test.

The pre-test was administered before giving treatment and it was administered to experimental and control groups. After getting the result of pre-test from experimental group and control group, the researcher gave

treatment to teach reading recount text for experimental group by using Flipped Classroom. Meanwhile, the researcher gave a conventional strategy to teach reading recount text for control group. The post-test was used to know the students' reading recount text achievement after being taught by using Flipped Classroom and those taught by using a conventional strategy.

E. Validity and Reliability Testing

A research is always dependent upon measurement. There are two important characteristics that every measuring instrument should go through a process of validity and reliability. This argument is supported by Ary et al. (2010:224) who said that there are two concepts that must be understood by the researcher when they used instruments, they were validity and reliability. Therefore, the instrument used in this study must be valid and reliable before using it to collect the data.

1. Validity Testing

Validity is the most important consideration in developing and evaluating measuring instruments (Ary et.al, 2010: 225). There are four types of validity; content validity, criterion validity, construct validity, face validity. In this research, the researcher analyzed the test from content validity, construct validity, and face validity.

a. Content Validity

According to Haynes, Ricard, and Kubany (1995) content validity is the degree to which elements of an instrument are relevant to and representative of the targeted construct for a particular

assessment purpose. So, the instrument can be said to have content validity if the content of the test match or appropriate with sample of language skill. Ary et al. (2010: 226) stated that to have a content validity, the instrument is representative of some defined universe or domain of content. In this study, the test had content validity because the test was based on the core competence and basic competence in the syllabus of the eighth grade students at Junior High School which focus on recount text.

Table 3.3 Matrix of Core Competence and Basic Competence

Core Competence	Basic Competence	Competence Indicator	Test Item
KI 4: Trying, processing, and serving in a concrete realm (using, parsing, composing, modifying, and creating) and abstract realm (writing, reading, computing, drawing and composing) with those studied in schools and other sources in the same angle view/theory.	4.7.1 Applying the structure of the text and linguistic elements to carry out the social function of stating and asking for actions/events carried out/happened in the past according to the context of its use.	1. Students are able to determine the main idea of the text. 2. Students are able to categorize the text based on the kinds of the text. 3. Students are able to determine the purpose of the text. 4. Students are able to classify each paragraph based on the generic structure of the text. 5. Students are able to know how to use the	1-25

		process of reading to determine the elements of the text. 6. Students are able to determine the specific information of the text.	
--	--	---	--

The matrix above showed that recount text is one of the texts that must be mastered by the eighth grade students of Junior High School in K13 curriculum. That is why teaching and testing students' ability in recount text was done in this study.

b. Construct Validity

The construct validity of test is test which is capable of measuring certain specific characteristics in accordance with a theory of language behavior and learning (Heaton, 1975: 159). Construct validity is validity which shows how far the tests are suitable with the theory that becomes a foundation on composing those tests. Construct validity refers to the composing of instrument. This type of validity is judgment based on the accumulation of evidence from numerous studies using a specific measuring instrument. It means that construct validity is one kind of validity that is measures the skill which is supposed to measure.

Based on explanation above, in the test, the researcher asked the students to answer the multiple choice based on recount

text to measure the students' achievement in reading and this fulfill the construct of reading text and therefore valid in term of construct validity. See the construct of validation in the term of blue print.

c. Face Validity

A test is said to have face validity if it measures what is intended to be measured. Face validity is hardly a scientific concept that is very important. Face validity refers to the extent to which examiners believe the instrument is measuring what it is supposed to measure (Ary et al, 2010: 228). A test which does not have face validity may not be accepted by test takers, teachers, educations, authorities or employers. In this test, there are some aspects to be considered from this test to make a good test based on the validity.

- 1) In this test, the instruction was clear for the students, so the students were able to understand what they should to do in the test.
- 2) In this study, the students of the eighth grade were instructed to work on multiple choice about recount text. Thus, the researcher gave questions that were suitable with syllabus and their level.
- 3) The consideration of time allocation was suitable. So, the students could finish their questions given to them well. In this test, the researcher gave the time allotment about 40 minutes.

2. Reliability Testing

One of the indicators to determine that the test use in this study was good was by determining its reliability. Reliability refers to the consistency of the scores resulted from the instrument (test). Brown (2002: 250) stated that reliability is concerned with the effect of such random errors of measurement on the consistency of scores. A test is said to be reliable if the subject give the same test to the same subject or matched subject on two different occasions, the test itself should yield similar result. Meanwhile, Ary et al. (2010: 236) said that reliability of a measuring instrument is the degree of consistency with which it measures whatever it is measuring.

Checking the reliability of the test result in this study was by conducting the try out. Before doing the tryout, the researcher consulted the instrument with two experts. They were the researcher's adviser and the English teacher of MTsN 4 Tulungagung. This try out was done to determine whether the test would be used in the post-test yield consistent score. There must be two row scores to determine the reliability. In this study, to get two row scores the test-retest was done. The first test was administered on Monday, February 3rd 2020 and the second test was administered on Wednesday, February 5th 2020. Tryout was implemented to 37 students of VIII-F at MTsN 4 Tulungagung. The gotten two row scores then were computed by using Pearson Product Moment with SPSS 16.0 for windows. To determine how reliable the

result of the test, the result was matched with the following classifications of the Cronbach's Alpha according to Riduwan (2004) as follows:

1. If the *alpha cronbach* score 0.00 – 0.20: less reliable
2. If the *alpha cronbach* score 0.21 – 0.40: rather reliable
3. If the *alpha cronbach* score 0.41 – 0.60: enough reliable
4. If the *alpha cronbach* score 0.61 – 0.80: reliable
5. If the *alpha cronbach* score 0.81 – 1.00: very reliable

The result of reliability testing by using Pearson Product Moment can be seen from the Table 3.4, as bellow:

**Table 3.4 The Result of Reliability
Correlations**

		Test1	Test2
Test1	Pearson Correlation	1	.918**
	Sig. (2-tailed)		.000
	N	37	37
Test2	Pearson Correlation	.918**	1
	Sig. (2-tailed)	.000	
	N	37	37

**. Correlation is significant at the 0.01 level (2-tailed).

Table 3.4 above showed that the result of reliability computation was 0.918. Based on the Cronbach's Alpha interpretation by Riduwan (2004), the value was 0.918 lies on level 0.81-1.00 that indicates very

reliable. Thus, the instrument (test) in this study was reliable and it was good used to collect data in the post-test.

F. Normality and Homogeneity Testing

1. Normality Testing

Normality distribution test is a test to measure whether our data has a normal distribution or not. That is important to get the normality data because showed the sample data represent to population when it is come from a normal distribution. To know the normality, the researcher used *Kolmogorov-Smirnov* test with SPSS 16.0. *Kolmogorov-Smirnov D* test is a test of normality for large samples. The normality of data can be seen based on the significant value (α) = 0.05. Testing of data normality is conducted by the rules as follows:

- a. If the value of significant level is bigger, so the distribution of the data is normal.
- b. If the value of significant level is smaller, so the distribution of the data is not normal.

While the criteria of acceptance or rejection of normality test were as follows:

- a. H_0 was accepted if $\text{sig} > \alpha = 0.05$
- b. H_a was accepted if $\text{sig} < \alpha = 0.05$

The result of the normality test computed by SPSS 16.0 was as follow:

- a. Normality Testing of Experimental Group

Table 3.5
Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest	.118	35	.200*	.963	35	.285
Posttest	.147	35	.052	.940	35	.057

a. Lilliefors Significance Correction

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

Based on the table above it was known that the significance value from the pre-test was 0.200 and from the post-test was 0.052. The sig/p value on pre-test was 0.200 and it was bigger than 0.05 ($0.200 > 0.05$) means that the data was in normal distribution. Then, for the post-test score the value of sig/p was 0.052 and that was bigger than 0.05 ($0.052 > 0.05$) meant that the data was in normal distribution. It also meant that H_0 was accepted and H_a was rejected. So, it can be interpreted that both of data (pre-test and post-test score) were in normal distribution.

b. Normality Testing of Control Group

Table 3.6
Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pretest	.146	34	.064	.925	34	.022
Posttest	.108	34	.200*	.975	34	.619

a. Lilliefors Significance Correction

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

Based on the table above it was known that the significance value from the pre-test was 0.064 and from the post-test was 0.200. The sig/p

value on pre-test was 0.064 and that was bigger than 0.05 ($0.064 > 0.05$) meant that the data was in normal distribution. Then, for the post-test score the value of sig/p was 0.200 and that was bigger than 0.05 ($0.200 > 0.05$) meant that the data was in normal distribution. It also meant that H_0 was accepted and H_a was rejected. So, it can be interpreted that both of data (pre-test and post-test) were in normal distribution.

2. Homogeneity Testing

Homogeneity testing is used to determine whether the data gained has a homogeneous variance or not. To know the homogeneity, the researcher used One Way Anova formula by using SPSS program 16.0 by the value of significance (α) = 0.05. Homogeneity testing was done after doing distribution score of group involved. Before doing homogeneity testing, the researcher decided hypothesis in this homogeneity as follows:

- a. H_0 : 1 variance (Experimental group and control group) was homogenous.
- b. H_a : 1 variance (Experimental group and control group) was not homogenous.

There is certainty in taking decision of homogeneity testing, as follow:

- a. If the value of significance > 0.05 , H_0 is accepted
- b. If the value of significance < 0.05 , H_a is rejected

The result of the homogeneity test computed by SPSS 16.0 was as follow:

Table 3.7
Test of Homogeneity of Variances

Result

Levene Statistic	df1	df2	Sig.
.029	1	67	.866

Based on the table above it was known that the significance value was 0.866, it means that the significant value was more than 0.05 ($0.866 > 0.05$). It means that H_0 was accepted and H_a was rejected. So, it can be interpreted that the homogeneity testing of variance in both groups in this research showed that the data had homogeneous variance, so it was qualified to be analyzed.

G. Data Collecting Method

Collecting data means identifying and selecting individuals for a study, obtaining their permission on the study and gathering information by asking people questions or observing their behavior (Creswell, 2012: 9). Therefore, the data collection was carried out by the researcher to obtain data or gather information needed to measure the effectiveness of a learning strategy called Flipped Classroom on students' reading recount text achievement. The researcher used test instruments. The instruments were pre-test and post-test. The pre-test was used to know the students' mean score in reading recount text achievement before being taught by using Flipped Classroom and those taught by using a conventional strategy. Meanwhile, the post-test was used to know the students' mean score in reading recount text

achievement after being taught by using Flipped Classroom and those taught by using a conventional strategy. The techniques of collecting data were as follows:

1. Pre-test

The test was administered before the treatment was given. After getting two classes, one class was as an experimental group that used Flipped Classroom in teaching reading recount text and another class as control group that used a conventional strategy in teaching reading recount text. The pre-test was aimed to know the basic competence and their previous knowledge before they get the treatment. It was done on Wednesday, February 5th 2020 (control group) and on Thursday, February 6th 2020 (experimental group). The pre-test were asked the students to read recount text and answered a multiple choice consisted of 25 items. After finishing the test or student' works, the researcher computed the score of pre-test. The aim of computation was to know the result of pre-test before the treatment.

2. Post-test

After the treatment process, the researcher continued to administered a post-test to the experimental group and the control group. A post-test was a measure on some attributes or characteristics that is assessed for participants after a treatment (Creswell, 2008: 201). Post-test was given in the last meeting of teaching learning process. It was used to measure students' reading recount text achievement after

being given treatment Flipped Classroom (experimental group) and those taught by using a conventional strategy (control group). Similarly to pre-test, the post-test also consisted of 25 items of Reading test in the form of multiple choice. The two scores of post-test taken from the groups were analyzed to find out whether there was any significant difference mean score in reading recount text between the control group and the experimental group. The post-test was done on Wednesday, February 19th 2020 (control group) and on Thursday, February 20th 2020 (experimental group).

H. Procedures of the Study

The treatment was given after administered the pre-test and before the post-test for control group and experimental group. The treatment here means that the researcher used Flipped Classroom in teaching process. Flipped Classroom was given to experimental group and the control group was taught by using a conventional strategy. During the treatment, the experimental group was given video learning through online classroom through online media such as WhatsApp group class and the control group was taught by using a conventional strategy. The treatment was done in three times. The steps in teaching reading recount text were adapted from Jannah (2007):

a. Experimental Class

- Outside the classroom

- 1) The teacher shared a video teaching about recount text explanation through WhatsApp group that the teacher has been made. The students were able to download the video through their laptop or smartphone.
- 2) The students watched the video at home or wherever place that the students comfort to study in. They can also study individually or in groups.
- 3) During watching the video, the students took some notes from the video and wrote a problem that they faced when they learn from the video.
- 4) The students were given some exercise through the video.

Note: To make sure that the students watched the video by themselves, the teacher gave students exercise through the video that could be fulfilled after watching the video that has been shared. When in the class, the teacher asked the students to work on some questions about the video.

- Inside the classroom

- 1) The teacher asked the students about what they had learned from the video and corrected the answer that the student had to answer in the video with the class.
- 2) The teacher asked the students whether there was something that they did not understand by the video.
- 3) The teacher gave the students a recount text.

- 4) The teacher divided the students into 4 groups, each group consisted of 4 or 5 students.
- 5) In group, the students identified and explored the recount text in each paragraph.
- 6) The students used their own words to write important information from the text to answer questions.
- 7) The students continued this activity until the end of paragraph from a recount text that was given by the teacher.
- 8) The teacher and the students discussed the content of the text. The teacher controlled and provided the right answer while checking the answer.
- 9) The teacher evaluated the learning process.

b. Control Class

The control class did not receive treatment as it was given to the experimental group and the learning process was done using a conventional learning style. The learning plans for control group were:

- 1) The teacher explained the lesson in front the class in form of conventional strategy.
- 2) The teacher gave a recount text to students.
- 3) The teacher asked the students to read the learning material using silent reading and pointed some important information from the text.
- 4) The students collected the work.

I. Data Analysis

Data analysis is a time consuming and difficult process, because typically the researcher faces massive amounts of field notes, interview transcripts, reflections, and information from documents to examine and interpret (Ary et al, 2014). The data obtained from results of students' test were analyzed quantitatively. Quantitative analysis was done using statistic is called as statistical analysis or inferential statistic. The quantitative data of this research in analyzed using statistical computation.

The statistical computation was used using Independent Sample T-test through the application of SPSS statistic 16.0 for windows. Practically, the result of this statistic was used to test hypothesis. To know any significant difference students' mean score of reading recount text achievement between both of class and to know the effectiveness of Flipped Classroom on students' reading recount text achievement. If the result of t-test was bigger than at the level of significance 0.05, the null hypothesis (H_0) was accepted, and alternative hypothesis (H_a) was rejected. It means that there was no significant difference mean score of reading recount text achievement between the students taught by using Flipped Classroom and those taught by using a conventional strategy. In contrast, if the result of t-test was lower than at the level of significance 0.05, the null hypothesis (H_0) was rejected, and alternative hypothesis (H_a) was accepted. It means that there was significant difference mean score of reading recount text

achievement between the students taught by using Flipped Classroom and those taught by using a conventional strategy.