CHAPTER III

RESEARCH METHOD

This chapter presents about research method. It consists of research design, research variable, population, sample, sampling technique, research instrument, validity and reliability testing, data collecting method, treatment, procedure of using peer feedback, normality and homogeneity testing, and also data analysis.

A. Research Design

In this research, the researcher used experimental design. According to Ary (2010: 301), an experimental design is the general plan for carrying out a study with an active independent variable. The design is important because it determines the study's internal validity, which is the ability to reach valid conclusions about the effect of the experimental treatment on the dependent variable. It could be concluded that there was a treatment for experimental class to know whether the treatment was effective for dependent variable or not. In this design, quasi-experimental design was used by the researcher to find out the significant difference on students' writing recount text ability assessed with and without using peer feedback.

Quasi-experimental designs are similar to randomized experimental designs in that they involve manipulation of an independent variable but differ in that subjects are not randomly assigned to treatment group (Ary: 2010). We can apply the pre and post-test design approach to quasi-

experimental design (Creswell: 2012). It meant that there were two classes that selected. They were experimental class and control class. Both of classes got pre-test and post-test. The pre-test was conducted before giving treatment (peer feedback). The treatment was only conducted for experimental class, and control class got treatment but didn't get peer feedback treatment as experimental class. After treatment was conducted, post-test was given for both of classes. The result of post-test was used to determine the significant difference on students' writing recount text ability assessed with and without using peer feedback. The research design could be presented as follows:

Table 3.1 Research Design

Group	Pretest	Independent Variable	Posttest
E	Y_1	X	Y_2
C	Y_1	-	Y_2

E : Experimental group
C : Control group
Y_1 : Pre-test for experimental and control group
X : Treatment by using peer feedback for experimental group
Y_2 : Post-test for experimental and control group

B. Research Variable

According to Ary (2010: 37), variable is a construct or a characteristic that can take on different values or score. According to

Creswell (2008: 112), a variable is characteristic or attribute of an individual or an organization that (a) researcher can measure or observe and (b) varies among individuals or organization studied. It meant that variable was characteristic of subject that studied in which differable or changeable. There were two variables which used in this research. They were independent variable and dependent variable. For the detail explanation, could be seen as follows:

1. Independent Variable

   An independent variable is an attribute or characteristic that influences or affects an outcome or dependent variable (Creswell, 2008: 116). In this research, an independent variable was the effectiveness of peer feedback.

2. Dependent Variable

   According to Creswell (2008:115), a dependent variable is an attribute or characteristic that is dependent on or influenced by independent variable. In this research, a dependent variable was students' writing recount text ability.

C. Population, Sample, and Sampling Technique

1. Population

According to Ary (2010: 148), population is defined as all members of any well-defined class of people, events, or objects. It is the larger group about what which the generalization is made.

According to Creswell (2012: 142), population is a group of individuals who have the some characteristics. From previous explanations, population was all subjects that will be studied. They were not only individual, events, or object, but they also had the some characteristics. In this research, the population was the students of second grade at SMP Terpadu Al Anwar Durenan Trenggalek in academic year 2019/2020. The second grade of SMP Terpadu Al Anwar Durenan Trenggalek divided into 5 classes. The total number of student was 166 students.

Table 3.2 The Number of the Second Grade Students at SMP Terpadu Al Anwar Durenan Trengalek in Academic Year 2019/2020

Class	VIII-A	VIII-B	VIII-C	VIII-D	VIII-E
The Number of Student	30	33	31	34	38
Total	166				

2. Sample

According to Ary (2010: 148), a sample is a portion of population. It is smaller than population. According to Creswell (2012: 142), a sample is a subgroup of the target population that the researcher plans to study for generalizing about the target population. It meant that sample was part of population that will be studied. There were five classes of the second grade students at SMP Terpadu Al Anwar Durenan Trengalek, but only two classes were taken as

samples. One class was experimental class and the other was control class. The class that became an experimental class was VIII-D that consisted of 34 students, and the control class was VIII-E that consisted of 34 students. The research sample could be seen in table 3.3 as follows:

Table 3.3 Research Sample

Class	The number of Student	Total
VIII-D (Experimental Class)	34	72
VIII-E (Control Class)	38	

To make sure two classes were equal or not, the researcher calculated the data of pre-test scores both of classes. In this case, the researcher used Mann-Whitney U Test in SPSS 16.0 version for windows because the result of normality testing of pre-test scores was not normally distributed (see normality testing result of pre-test in page 57). The result showed that sig (2-tailed) or *p*-value was 0.000. It meant that the result was less than significance level 0.05 (0.000<0.05). Although this research used one right tailed, but to know that two classes were equal or not, the result of sig (2-tailed) was not divided into two. So, it indicated that two samples or two classes were not equal. The result of calculation could be seen in table 3.4 as follows:

Table 3.4 Mann-Whitney U Test for Pre-Test

Test Statistics[a]

	Scores
Mann-Whitney U	176.000
Wilcoxon W	842.000
Z	-5.157
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: classes

3. Sampling Technique

Sampling technique was technique that used to determine the sample. In this research, researcher used purposive sampling that included as non-probability sampling. According to Creswell (2012: 145), in non-probability sampling, the researcher selects individuals because they are available, convenient, and represent some characteristic the investigator seeks to study. The researcher chose the samples because they were appropriate with the purpose of this research.

D. Research Instrument

Instrument had important part in this research. According to Creswell (2008: 151), an instrument is a tool for measuring, observing, or documenting quantitative data. In making research instrument, the reseacher

used the basic competence of English lesson in writing skill for the second grade students at SMP Terpadu Al Anwar Durenan Trenggalek. This basic competence was “arranging an oral and written of recount texts, short and simple about activities, events, phenomenon with due regard to social functions, structures of text, linguistic elements in a context correctly”.

Identified before the researcher collecting the data, the instrument might be a test, questionnaire, tally sheet, log, observational checklist, inventory, or assessment instrument (Creswell, 2008: 151). In this research, the researcher used test as instrument to collect data. According to Ary (2010: 201), a 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. The researcher conducted some steps to develop the research instrument, starting from reviewing material in syllabus of the second grade students of Junior High School until writing final instrument. For detail explanation, could be seen in the following explanation.

1. Reviewing Syllabus

   The first step that conducted by the researcher to develop research instrument was reviewing syllabus of the second grade students of Junior High School. The researcher looked for the material related with writing ability. But, the researcher chose recount text as the type of text that used to measure students’ writing ability.

2. Making Test Specification

After reviewing the syllabus, the next step was making test specification. The researcher made test specification related with the content of material. In test specification, the researcher also gave the indicator of the test (see appendix 1).

3. Drafting Instrument

In this stage, the researcher made test item of instrument. There were two tests that used (pre-test and post-test) and each of test consisted of one item. The test was about writing essay. So, the students were asked to write recount text minimum in 10 lines. The purpose was to know students' writing ability in recount text.

4. Validation

After making test item, the next step was validation. Validation was used to make good test. So, before applying test for the students, the researcher asked the expert to make validation the item of instrument and test specification. The researcher asked one of lecturer and English teacher at SMP Terpadu Al Anwar Durenan Trenggalek as the expert of validation.

5. Revising

In this step, the researcher revised what should be revised. It related with the comment or suggestions of the experts. In validation sheet of instrument, there were no suggestions from the experts. So, the researcher didn't revise it.

6. Tryout

After revising test item, the researcher conducted tryout before the test was applied in pre-test for two classes. The class that used to conduct tryout was different from research class (experimental and control class). The researcher used VIII-B as the tryout class. They were given 30 minutes to do the test. They asked to make recount text which the topics were determined by the researcher.

7. Writing Final Instrument

It was the last step in developing research instrument. In this step, the researcher made complete instrument that used to collect data in two classes, experimental class and control class.

This research, the researcher used two tests to collect the data. One test was pre-test and the other was post-test. The post-test were used to measure the significant difference on students' writing recount text ability assessed with and without using peer feedback. Both of tests had one item. The pre-test was given for both of classes (experimental class and control class) with the same topics. The pre-test was given before the treatment (using peer feedback), while the post-test was given for both of classes after given the treatment. In post-test, an experimental class and control class had the same topics. Both of tests were developed based on students' writing skill, which referred to the basic competence in syllabus of second grade students at SMP Terpadu Al Anwar Durenan Trengalek. Before the

researcher implementing the instrument to the samples, the reliability and validity of the pre-test and post-test instrument was calculated.

E. Validity and Reliability Testing

1. Validity

In this research, the researcher used instrument to collect the data. The instrument was tests. According to Ary (2010: 224) there two concepts that must be understood by the researchers when they uses instrument are validity and reliability. Therefore, validity and reliability were very important in using instrument. They could make a good test. A valid instrument meant the extent to which an instrument measured what was supposed to measure. There were three validities that used in this research. They were content validity, construct validity, and face validity.

a. Content Validity

According to Creswell (2008: 618), content validity is the extent to which the questions on the instrument and the scores from these questions are representative of all the possible questions that could be asked about the content or skill. In content validity, the test must be appropriate with the purpose of that test. In this research, the writing test must be appropriate to measure writing ability in recount text. So, the test was developed based on the basic competence of the second grade student syllabus of Junior High School. For the basic competence and indicator could be seen in the table 3.5. Table 3.5 could be seen in the next page.

Table 3.5 Basic Competence and Indicator of the Second Grade Students of Junior High School

Basic Competence	Indicator
4.15 Arranging oral and written of recount texts, short and simple about activities, events, phenomenon with due regard to social functions, structures of text, linguistic elements in a context correctly.	1. The students are able to make simple text about the activities, events, and phenomenon by using the structure and the language features of recount text correctly.

b. Construct Validity

Construct validity showed how far the test items are relevant with the theory. A construct is any theory, hypothesis, or model that attempts to explain observed phenomena in our universe of perceptions (Brown, 2003: 25). According to Brown (2001: 335), writing as a written product of thinking, drafting, and revising that requires specialized skill on how to generate ideas, how to recognize them coherently, how to use discourse makers and rhetorical conventional to put them cohesively into a written text, how to revise a text for clear meanings, how to produce a final product. So, in this research, the test was appropriate with the theory because the students were asked to make a product. It was recount text writing. The researcher constructed the instrument based on the test specification of the writing recount text consisting specific indicator to fulfill construct validity. To assess students' writing work, the researcher used scoring rubric adapted from Jacob et al. It consisted of five aspects,

namely content, organization, language use, vocabulary, and also mechanics. The detail scoring rubric could be seen in appendix 4.

c. Face Validity

Face validity means that the students perceive the test to be valid (Brown, 2003: 27). Face validity is a term sometimes used in connection with the test's content (Ary et al, 2010: 228). In this research, the researcher made a test specification as the concept to make a test. Then, the researcher consulted and asked the expert to give the comments and suggestions for the test. There were some aspects that considered making a good test based on the validity.

a) The instructions of the test must be clear. Thus, the students could understand and know what they should do in the test.
b) The time allocation must be considered. The time allocation should be enough for the students. It was not too long and not too fast. The researcher gave thirty minutes for the students to write recount text.
c) Based on the syllabus of the second grade students of Junior High School, the students asked to write a simple recount text about activities, events, and phenomenon. So, the degree of difficulty of the test must be suitable with their level.

2. Reliability

Reliability was the second aspect to measure test instrument. According to Brown (2003: 20), a reliable test is consistent and dependable. Reliability meant that scores from an instrument are stable

and consistent. It meant that when the researcher gave the same test to the same students or match students on two different occasions, the test should yield similar result. Scores should be nearly the same when the researchers administered the instrument multiple times at different times. If scores were not reliable, they were not valid. Scores needed to be stable and consistent first before they could be meaningful (Creswell, 2008: 159).

To measure the reliability of this instrument, the researcher conducted tryout to see if the result of test was reliable or not. The try out was conducted to make sure that the test was clear and it was not too difficult or not too easy for the students. It conducted in VIII-B class that consisted of 33 students (but 2 students were absent). To see the reliability of the instrument of writing ability, the researcher used inter-rater reliability.

According to Creswell (2008: 161), inter-rater reliability involves observations made by two or more individuals of an individual's or several individuals' behavior. It meant that two scorer did the scoring. After that, the scores were calculated to see if the scores were similar or different. The result of the scores was calculated by using *Cronbach's Alpha* from SPSS Statistics 16.0 version for windows. It was represented in the following table 3.6 (see in the next page).

Table 3.6 The Result of Reliability Testing

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.873	.881	2

Based on the result of inter-rater test in table 3.6, it could be concluded that the test was reliable. It known from the value of *Cronbach's Alpha* was 0.873. The level of reliability was very high. According to Suharto as cited in Risdiani (2018), the rate of value in the instrument was presented in the following table:

Table 3.7 The Value of Reliability Coefficient

Reliability Coefficient	Reliability Category
0.800 up to 1.000	Very High
0.600 up to 0.799	High
0.400 up to 0.599	Fair
0.200 up to 0.399	Low
0.000 up to 0.199	Very Low

F. Data Collecting Method

In this research, the researcher used pre-test and post-test to collect the data. For detail explanation could be explained as follows:

1. Pre-test

The pre-test was given to the students before given the treatment. It was given for both of class, namely experimental and control class. VIII-D class became an experimental class, while VIII-E class became control class. The researcher gave the pre-test in the same day but different time, because both of class had different schedule for English lesson. The first pre-test is given for VIII-D class. It was held on Thursday, February 6th 2020 on 07.10-07.55 a.m. Then, in the next time, the researcher gave the pre-test for VIII-E class. It was also held on Thursday, February 6th 2020 but on 11.15-12.00 a.m. The researcher gave the same topics for them. The topics were An Experience in the Beach and Sad Experience. The students could choose one of topics. They were asked to write minimum in 10 lines and without any maximum lines. The researcher gave 30 minutes for the students to write recount text related with the topic chosen. So, both of classes (control and experimental class) had the same topics and time allocation. The purpose of this pre-test was to know their ability and knowledge in writing recount text before they were given the treatment. After that, their work was assessed by using scoring rubric that adapted from Jacobs et al (in Weigle, 2002).

2. Post-test

Posttest was given after the treatments were done. The post-test was also given for both of classes (experimental and control class). The researcher gave the same topics about “The Last Holiday in Your

Elementary School and An Embarrassing Experience". They could choose one of topics that they wanted. The researcher gave 30 minutes for the students to do the test. The first post-test was given for VIII-E (control class). It was held on Wednesday, February 12th 2020. The second post-test was for VIII-D (experimental class). It was held on Thursday, February 13th 2020. Finally, their writing work was assessed by using scoring rubric adapted from Jacobs et al (in Weigle, 2002).

The purpose of post-test was to measure their improvement, ability and knowledge of recount text after they learnt writing recount text by using peer feedback in experimental class and without using peer feedback in control class. The result of post-test score in experimental class was compared with post-test score of control class to know the significant difference of students' writing recount text assessed with and without using peer feedback. The schedules of collecting data and treatment were presented in the following table 3.8.

Table 3.8 The Schedule of Collecting Data and Treatment

Time	Material	Activities	Goal
Thursday, February 06th 2020 On 07.10-07.55 a.m.	Recount Text	**Pre-test of Experimental Class** **Topic**: An Experience in the Beach and Sad Experience	To measure students' writing ability before the treatment
Thursday, February 06th 2020 On 11.15-12.00 a.m.	Recount Text	**Pre-test of Control Class** **Topic**: An Experience in the Beach and Sad Experience	To measure students' writing ability before the treatment

Continued

Continuation

Saturday, February 08th 2020	Recount Text	**1st Treatment Topic:** Free topic (students' own experience)	The students wrote and implementing peer feedback
Tuesday, February 11th 2020	Recount Text	**2nd Treatment Topic:** Free topic (students' own experience)	The students wrote and implementing peer feedback
Wednesday, February 12th 2020	Recount Text	**Post-test of Control Class Topic**: The Last Holiday in Your Elementary School and An Embarrassing Experience	To measure students' writing ability after the treatment
Thursday February 13th 2020	Recount Text	**Post-test of Experimental Class Topic**: The Last Holiday in Your Elementary School and An Embarrassing Experience	To measure students' writing ability after the treatment

G. Treatment

The researcher gave the treatment for experimental class only after the researcher giving the pre-test for both of classes (control class and experimental class). The treatment was given twice for experimental class (VIII-D) by using peer feedback. It was held on Saturday, February 8th 2020 and on Tuesday, February 11th 2020. In contrary, there was no special treatment for control class (VIII-E). They were treated without peer feedback. It was held on Saturday, February 8th 2020 and on Monday, February 10th 2020. The procedures that conducted by the researcher to

conduct the first and the second treatment in experimental class were the same, but it was only different in time allocation. The time allocation to conduct the first treatment was 45 minutes and the second was 90 minutes. The procedures of treatment could be seen in the following table 3.9.

.

Table 3.9 The Procedures of Conducting Treatment

NO	STAGES	DESCRIPTION OF ACTIVITIES
1	Pre-Activity	• Greeting. • Praying together before the lesson started. • The researcher checked the students' attendance. • The researcher asked some things related with the topic.
2	Main Activity	• Observing - The students listened and pay attention in class presentation or the explanation of the researcher about recount text. • Asking - The researcher gave the time to the students to ask some questions that they understand yet about the topic given. • Collecting Information (20 menit) - The researcher asked the students to make recount text about their own experiences. - The students made recount text with pay attention the structures and language features of recount text. They made recount text minimum in 10 lines and without any maximum lines. • Associating - After they finished, the researcher asked the students to exchange their writing work to their peer. - The students red their peer's writing work and gave the correction of the mistakes in writing work. They also gave the scores based on scoring rubric given. - The students filled the peer feedback checklist.

Continued

Continuation

		• Communicating - The student who pointed by the researcher conveyed the mistakes of peer's writing work to the researcher. - The students returned the recount text writing and peer feedback checklist to the peer in order to the student could know and repaired the mistakes.
3	Post-Activity	• The researcher gave feedback on the result of the study. • The researcher gave the information about what they should do in the next meeting. • The researcher gave motivation to the students. • Praying together. • Greeting

H. Procedures of Using Peer Feedback

In using peer feedback in experimental class, there are 10 procedures that conducted by the researcher to the students.

1) The first, the researcher explained about what peer feedback was and also explained what they should do in this treatment. The researcher explained that they had to make a recount text minimum in 10 lines. But, the topic was related their own experiences. Besides, the researcher also explained that they had to correct and give comments of their peer's writing based on criteria given and fill peer feedback checklist.
2) After they understood, the researcher asked the students to make a group consisted of two students. They could choose their seatmate. Because the students in experimental class consisted of 38 students, so each student had peer.

3) After they had a peer, the researcher distributed question sheet, answer sheet, scoring rubric, and also peer feedback checklist to the students.
4) Next, the researcher opened the question section if they didn't understand yet. They could ask if they didn't understand about what the researcher had explained in the beginning.
5) After they understood what they should do, the researcher asked them to start their writing. The researcher gave 30 minutes for the students to write.
6) The researcher stopped them when the time was up. Then, the researcher asked them to exchange their writing work to their peer and vice versa.
7) Next, the students red, correct or give the feedback, and also gave the scores of their peer's writing based on some criteria that given by the researcher in scoring rubric.
8) After they had already finished, they filled the peer feedback checklist instrument by giving tick (✓) and cross (X) depended on criteria determined.
9) Then, the students returned his/her work and vice versa in order to their peer know their mistakes and they would be revised in their next writing.
10) The last, they returned their writing work and peer feedback checklist to the researcher to be given the feedback if they had already done to see their mistakes in their writing.

I. Normality and Homogeneity Testing

1. Test of Normality

Normality testing was used in most of inferential statistic. Test of normality was used to know whether the data was normally distributed or not. There were two ways that could be used to test normality of the data, namely Shapiro-Wilk and Kolmogorov-Smirnov. But, the researcher used Kolmogorov-Smirnov because the sample that taken by the researcher was 72 students. According to Dahlan (2010), if the research sample are more than 50, the normality test uses Kolmogorov-Smirnov and while the research sample are less than 50, it can be used Shapiro-Wilk. In this research, the researcher calculated normality testing by using SPSS 16.0 version for windows. It was gained from post-test and pre-test score of the control and experiment class that had already conducted by the researcher. The hypothesis for normality testing was:

H_0 : If the value significance > 0.05, data is in normal distribution

H_a : If the value significance < 0.05, data is not in normal distribution

The result of normality testing of post-test and pre-test could be seen in the following table 3.10 and table 3.11.

Table 3.10The Result of Normality Testing of Post-Test

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	CLASS	Statistic	Df	Sig.	Statistic	df	Sig.
SCORES	Post-test_CONTROL	.161	36	.019	.940	36	.051
	Post-test_EXPERIMENT	.229	33	.000	.851	33	.000

a. Lilliefors Significance Correction

Based on the result of normality testing of post-test in table 3.10, it found that the data was not normally distributed. Because the samples were more than 50 students, so it used Kolmogrov-Smirnov. It could be seen in Kolmogorov-Smirnov column was found that the *p*-value (sig) in post-test of control class was 0.019 and *p*-value (sig) in experimental class was 0.000. It indicated that the result was less than significance level 0.05 (0.019<0.05) and (0.000<0.05). It indicated that the data of post-test score in both of classes was not normally distributed.

Table 3.11 The Result of Normality Testing of Pre-Test

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	CLASS	Statistic	df	Sig.	Statistic	df	Sig.
SCORES	Pre-test_CONTROL	.239	36	.000	.859	36	.000
	Pre-test_Experiment	.240	33	.000	.883	33	.002

a. Lilliefors Significance Correction

Based on the result of normality testing of pre-test in table 3.11, it found that the data was not normally distributed. Because the samples were more than 50 students, so it used Kolmogrov-Smirnov. It could be seen in Kolmogorov-Smirnov column was found that the *p*-value (sig) in pre-test of control class was 0.000 and *p*-value (sig) in experimental class was 0.000. It indicated that the result was less than significance level 0.05 (0.000<0.05) and (0.000<0.05). It indicated that the data of pre-test scores in both of classes was not normally distributed.

2. Test of Homogeneity

According to Arikunto (2010: 98), homogeneity is a measurement which can be used to determine data variation. It was used to analyze whether the sample variance was homogeneous or not. Parsetyowati (2016: 94) states that the considerations of testing homogeneity are:

a. If the result is higher than significance level 0.05, the data is homogeneous.
b. If the result is lower than significance level 0.05, the data is not homogeneous.

The result of homogeneity testing could be seen in table 3.12.

Table 3.12 The Result of Homogenity Testing

Test of Homogeneity of Variances

SCORES

Levene Statistic	df1	df2	Sig.
13.894	1	67	.000

The result of homogeneity testing by using SPSS 16.0 for windows in table 3.10 showed that the level of significance was 0.000. It indicated that the result was less than significance level 0.05 (0.000<0.05). Thus, it could be concluded that the data was not homogeneous. It meant that two classes that used did not have the variance which same relative or it was not homogeneous.

J. Data Analysis

After collecting data from the students, the next step was analyzing the data gotten. According to Richmond (2006: 13), data analysis can refer to a variety of specific procedures and methods. It involves goals; relationships; making decision; and ideas, in addition to working with the actual data itself. To know the significant difference from experimental and control class, the researcher calculated the N-Gain score of pre-test and post-test both of classes by using Mann-Whitney. This case calculated by using N-Gain Score because those classes were not equal. Their mean in pre-test in experimental class was very different with the mean of pre-test in control class. To analyze those data, the researcher used SPSS 16.0 version for windows. Qomari (2009) states that SPSS (Statistical Programs for Social Sciencies) helps some researchers to analyze the data in their research such as descriptive statistic, correlative and comparative. In this research, the researcher used Mann-Whitney U Test because the result of normality testing of N-Gain score data showed that the data was not normally distributed.