

## **CHAPTER III**

### **RESEARCH METHOD**

This chapter provides the research method. The discussion includes research design, variable of the research, population, sample and sampling technique, method of data collection and instrument, technique of data analysis, validity, and reliability.

#### **A. Research Design**

Research design is a way of thinking and planning for the processing and use of data to archive the research target. In conducting this study, the researcher uses the quantitative approach. Creswell, (2016: 5) found that quantitative research is an approach to test empirical hypotheses by analyzing the relationship between variables. These variables can be calculated, typically on instruments, so that the numbered data can be analyzed using statistical methods. Then, the quantitative approach is a technique which concerned with statistics and everything that can be calculated in a structured way to analyze phenomena and their connection.

Concerning with the quantitative approach, this study uses the correlational research design. According to Creswell, (2012: 338), correlational design offers an opportunity to forecast scores and describe the relationship between variables. In correlational study designs, researchers use the correlation statistical test to define and calculate the degree of association (or relationship)

between two or more variables or sets of scores. It can be inferred that correlational research design includes to quantitative approach by which investigators calculate the level of the relationship between two or more variables using the statistical method analysis of the correlation. This degree of association, interpreted as the number, indicates whether the two variables are related or whether one can predict others.

In correlational research design, there are several results based on expertise. According to Nunen (1992: 39) there are three potential outcomes of a correlation study: positive correlation, negative correlation, and no correlation. The correlation coefficient is a measure of the intensity of the correlation and can range from -1.00 to +1.00. Perfect positive correlation will result in a +1 score. Perfect negative correlation will result in a -1 ratio.

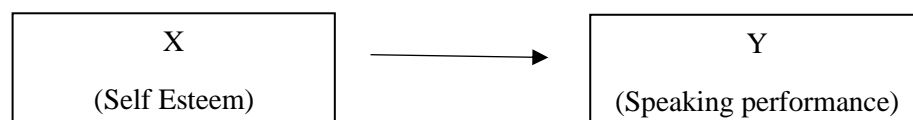
1. Positive correlation: The two variables increase or decrease at the same time. A correlation coefficient close +1.00 indicated a significant positive correlation
2. Negative correlations: indicates that, as the sum of one variable increases, the other decreases. A correlation coefficient near -1.00 indicated a significant negative correlation.
3. No Correlation: indicates that there is no relationship between the two variables. A correlation coefficient of 0 indicates that there is no correlation.

## **B. Variable of Research**

Variables basically refer to a person, place, object, or phenomenon that are supposed to assess in some context. Creswell (2012: 112) noted that a variable is a function or trait of a person or organization that can be calculated or observed by researchers and differs between individuals or organizations studied. In correlational research design, there are two kinds of variable include predictor variable and criterion variable. Predictor variables are variables (probably) causes or predict to the outcome or other variable. While criterion variables are the variables act as the outcome of the effect of the predictor variables. In this study, the researcher examine two kinds of variable to know the correlation between self esteem and speaking performance, those are :

- a. Predictor variable (X), that is students' self-esteem
- b. Criterion variable (Y), that is students' speaking performance

**Figure 3.1. The correlation between variables**



### **C. Population, Sample and Sampling Technique**

#### **1. Population**

Population is a group of individual which has a certain quality. In this term, Ary Donald et al. (2010: 54) describes population as a group of individuals who the researcher wants to study further. Meanwhile, Cresswell (2014: 142) comments that population is a group of individuals who have the same characteristic. It can inferred that population refers to a group of individuals or individuals who represent

one more of the characteristics from which data can be obtained and analyzed. In this study, the populations is all of the second grade students of MTs Ma'arif Udanawu Blitar in academic years 2020-2021. There are nine classes with the total number of students are 373 students. The following table shows the distribution of the total number of students of each class.

**Table 3.1 Population of the research**

<b>No.</b>	<b>Class</b>	<b>Number of students</b>
<b>1.</b>	VIII A	34
<b>2.</b>	VIII B	35
<b>3.</b>	VIII C	36
<b>4.</b>	VIII D	36
<b>5.</b>	VIII E	38
<b>6.</b>	VIII F	38
<b>7.</b>	VIII G	38
<b>8.</b>	VIII H	38
<b>9.</b>	VIII I (Excellent program)	39
<b>10.</b>	VIII J (Excellent program)	40
<b>Total</b>		<b>372</b>

## 2. Sample

Sample is used to get the data which will be analyzed. According to Creswell (2012: 14), sample is a particular group of the population being studied aims to analyze for the generalization of the target population. In the other words, sample is the part population where the data are collected and become the focus of the study. In this study, the reasearcher use one of the excellent classes of the second grade (VIII J) as the sample.

## 3. Sampling technique

In this research,, the researcher use purposive sampling technique. Ary Donald et al. (2010: 156) explains that purposive sampling technique takes sample elements that are regarded to be typical or representative from the population. In conclusion, purposive sampling technique use to get that data with several consideration and belongs to non-probability sample.

As the discussion before, this study use one of the excellent classes of the second grade (VIII J) as the sample. VIII J class was used as the sample because of the English teacher recomendation. That class has the better respect in terms of accepting and doing the school tasks than the other classes. That is why the sample of this study is belongs to the purposive sampling technique.

#### **D. Method of Data Collection**

In collecting the data, the researcher use questionnaires and test. The questionnaires use to measure the students' self esteem and the test use to know the students' speaking performance.

##### **a. Questionnaire**

Questionnaire is used to get the data of the student's self-esteem. Sugiyono (2008: 142) clarified that a data collection process which is carried out by giving respondents some questions or statements is called questionnaire. It can be concluded that questionnaire is one of research instrument contains some question or statements which should be

answered by the research sample. In this research, the questionnaire was given to the students to find numerical data of students' self esteem in the form of online questionnaire via Google Form because of pandemic situation. There was a very minimum possibility to distribute the questionnaire sheets because the school where the data were collected using online media in teaching learning activity.

In making and developing the question or statement of questionnaire, the researcher adapted a theory of self esteem indicators from Rosenberg (1965) with likert scale. The self-esteem scale include 10 items scale that measures global self-worth by measuring both positive and negative feelings dealing with the students' speaking performance. The ten scales of self -esteem are :

1. On the whole, I am satisfied with myself.
2. At times I think I am no good at all.
3. I feel that I have a number of good qualities.
4. I am able to do things as well as most other people.
5. I feel I do not have much to be proud of
6. I certainly feel useless at times.
7. I feel that I'm a person of worth, at least on an equal plane with others.
8. I wish I could have more respect for myself.
9. All in all, I am inclined to feel that I am a failure.
10. I take a positive attitude toward myself.

Then, those ten scales of self-esteem by Rosenberg (1965) the statements in the questionnaire had been modified and developed by the researcher in terms of wording. The questionnaire was made in English and Indonesian in order to avoid misunderstanding the meaning.

The choices of self-esteem questionnaire answer include strongly agree, agree, disagree and strongly disagree. For scoring system as follow:

**Table 3.2 Self-Esteem Questionnaire Scoring System**

<b>Negative statement</b>	<b>Categories</b>	<b>Positive statement</b>
1	SA (Strongly Agree)	4
2	A (Agree)	3
3	DA (Disagree)	2
4	SDA (Strongly Disagree)	1

Furthermore, the following table is used to classify the score of self esteem questionnaire

**Table 3.3 Classification of Self-Esteem Score**

<b>Score</b>	<b>Classification</b>
76 – 100	Very high
51 – 75	High
26 – 50	Low
1 – 25	Very low

b. Test

In getting the students' score of speaking performance, the researcher used test. Arikunto (2010: 266) describing test as measuring students' ability in terms of the presence or absence and amount the object. Test is to measure the presence or absence and the amount of the object under study capability. In this study, the researcher used students are

tested to get the speaking scores data and descriptive text theme used to test the speaking performance as the following table explanation

**Table 3.4 Speaking Test Material**

<b>Indicator</b>	<b>Sub Indicator</b>
4.11. Menangkap makna dalam teks deskriptif lisan dan tulis pendek dan sederhana	Providing information in descriptive text about person, animal and things orally
4.12. Menyusun teks deskriptif lisan dan tulis, pendek dan sederhana, tentang orang, binatang, dan benda, dengan memperhatikan fungsi sosial, struktur teks, dan unsur kebahasaan yang benar dan sesuai konteks.	

For try out, the students were asked to describe one of their classmate in a video include her or his name, place and date of birth, address, physical appearance and hobby. The students also can add more information about one of their friend who they are described. For research, the students were asked to describe one of Indonesian actor/actrees which are provided by the researcher in a video include her or his name and his/her physical appearances. The students also can add more information about one of the actor/actrees who they are described. Then, send it to the Whatsapp number of the researcher.

In scoring the students' speaking performance, the scoring rubric for speaking from O'Malley and Pierce (1996: 144) used to score the speaking test of the student. The blueprint of speaking test (try out and research) located in the appendixes. The following is the scoring rubric for speaking.

**Table 3.5**  
**Scoring Rubric for Speaking by O'Malley and Pierce (1996: 144)**

<b>Categorized</b>	<b>Score</b>	<b>Description</b>
<b>Pronunciation</b>	<b>5</b>	➤ Have few traces of foreign language accent



	4	➤ Always intellegible, though one is conscious of a definite accent
	3	➤ Pronunciation problem necessitate concentrated listening and occasionally lead to misunderstanding
	2	➤ Very hard to understand because of pronunciation problems, must frequently be asked to repeat
	1	➤ Pronunciation problems to severe as to make speech virtually unintelligible
<b>Vocabulary</b>	5	➤ Use of vocabullary is virtually that of a native speaker
	4	➤ Sometimes use inappropriate terms and or must rephrase the idea because of lexical inadequate
	3	➤ Frequently uses the wrong words. Coversation somewhat limited because inadequate vocabulary
	2	➤ Miss use of word and very limited vocabulary make comprehension quite difficult
	1	➤ Vocabulary limitations so extreme as to make conversation virtually impossible
<b>Grammar/Structure</b>	5	➤ Makes view noticable errors of grammar and word order
	4	➤ Occasionally makes grammatical and/ or word order error which do not, however obscure the meaning
	3	➤ Makes frequent errors of grammar and word order which occasionally obscure meaning
	2	➤ Grammar and word order comprehension difficult. Must often rephrase sentencesand or restrict him to basic patterns
	1	➤ Errors in grammar and word order so, severe as to make speech virtually unintelligible
<b>Fluency</b>	5	➤ Speed as fluent and effortless as that of a native speaker
	4	➤ Speed of the speech seems to be slightly affected by language problem
	3	➤ Speed and fluency are rather strongly affected by language problem
	2	➤ Usually hesistant often forced into silent by language limitations
	1	➤ Speech is as halting and fragmentary as to make conversation
<b>Comprehension</b>	5	➤ Appears to understand everything without difficulty
	4	➤ Understand nearly everything at normal speed, although occasionally repetition maybe necessary

	<b>3</b>	➤ Understand most of what is said at slower than normal speed with repetition
	<b>2</b>	➤ Has great difficulty following what is said
	<b>1</b>	➤ Can not be said understand even simple conversation virtually impossible

After that, in finding the predicates of students' achievement in speaking, the students are scored by the following classifications

**Table 3.6 Classifications of Speaking Score**

No	Score	Level of ability
1	21– 25	Excellent
2	16 – 20	Good
3	11 – 15	Average
4	5 – 10	Poor

### **E. Technique of Data Analysis**

After collecting the data, the next step was analyzing them to know whether there is any positive correlation between students' self esteem and students' speaking performance.

#### **a. Normality Testing**

Normality testing is used to check whether the data of the research are normally distributed or not. The researcher use SPSS 24. 0 Program in measuring the normality testing. Normality testing was done through *Kolmogorov Smirnov* because the total number of samples are 40 students (under 50). The significance value of the test is ( $\alpha = 0, 05$ ). Basic decisions making in normality testing as follows

- 1) If significance value  $> 0,05$  means that the data has normal distribution

- 2) If significance value  $< 0,05$ , means that the data does not has normal distribution

#### **b. Linearity Testing**

The purpose of linearity testing is to know the linear relationship between both variables (independent and independent variable) in which was done by correlation statistical analysis through SPSS 24.0 Program. The significance value of the test is ( $\alpha = 0, 05$ ). Basic decisions making in normality testing as follows

- 1) If significance value  $> 0,05$  means that the distribution of the data is linear
- 2) If significance value  $< 0,05$ , means that the the distribution of the data does not linear

#### **c. Correlation between Students' Self-Esteem and Students' Speaking Performance**

In finding wether there is a significant correlation between students' self-esteem and students' speaking performance, the data was analyzed by using statistical formula and the researcher used Spearman Rho correlation coefficient through SPPS 24.0 Program. The score of questionnaire (variable X) and the score of speaking test (variable Y) were used to test wether there is an correlation between both of variables or not. Basic decision of Spearman Rho correlation coefficient as follow

- 1) If significance value  $> 0,05$  means that there is no significant correlation between variable X and variable Y

- 2) If significance value  $< 0,05$ , means that there is a significant correlation between variable X and variable Y

Then, the ranging of correlation coefficient is as follow:

- 1) 0, 00 - 0, 199 means very low correlation
- 2) 0, 20 - 0, 399 means low correlation
- 3) 0, 40 - 0, 599 means fair correlation
- 4) 0, 60 - 0, 799 means high correlation
- 5) 0, 80 - 1, 000 means very high correlation

## **F. Validity**

Validity is one of an important key to effective research. Brown (2004: 24) wrote that a test is shown to have content validity if its content includes a representative sample of the language skills, systems, or etcetera being tested. Validity is the amount to which an instrument measured what it supposed to measure and it contributes to the intepretation and meaning of the scores produced from it (Ary Donald et al. 2010: 225). Then, the instrument of reasearch fulfill the validity requirements if it contains a sufficient sample of the structure or material that is applicable to the intent of the test. To measure whether the test has good validity or not, the instruments were analyzed in term of content validity and construct validity.

### **a. Content Validity**

One of the important thing in designing the test is it should fulfill the requirement of content validity. Brown (2004: 24) claims that a test is

shown to have content validity, if its content symbolizes a representative sample of the language skills, constructs, etc. being studied. If it contains a sufficient sample of the structure or material that is applicable to the intent of the test, the test would have content validity. Moreover, Brown (2002: 23) added his explanation that a test is claimed to be accurate in terms of content validity if its content includes the test goals which are in turn members of the syllabus. Then, content validity can be described as a reflection of the substance of the test which will be tested or investigated. That is why a detailed and systematic review of the content validity is required. In order to prove the content validity, the researcher has tried out both instruments used in this study (questionnaire and speaking test) as follows :

#### 1) Questionnaire

The self-esteem questionnaire was used to fulfill the requirement of content validity because it measures global self-worth by measuring both positive and negative feelings dealing with the students' speaking performance. For example ; *I feel that I am good at describing person/things in English* or *I certainly feel awkward and shy when I have to speak in English in front of my classmate at times*. Those examples of questionnaire statements are appropriate with the English speaking topic for the test.

For statistical analysis, 30 items of questionnaire were tried out to 39 total number of students at VIII I (excellent class). Then, SPSS 24.0

Program was used in analyzing the result of self-esteem questionnaire. R observed to R table at significant level 5% is 0,344 with (df = N - 2 = 28). To be considered a valid question, the R observed of each item should be greater than the R table. If the r on the study of less than R table is less than R, it can be inferred that these items are not substantially associated with the total score (declared invalid) and should be omitted or corrected.

The analysis result showed that there were 25 items valid and the rest were invalid. The invalid items were statements of number 2, 4, 16, 25 and 27. It means those 5 items were invalid because the scores was not fulfill 0,3440. Then, those 25 items of questionnaire are decided as the valid instrument and it ready to use to collect the data. The result of validity testing for self-esteem questionnaire can be seen on this following table

**Table 3.7**  
**Validity Testing Result of Self-Esteem Questionnaire (Try Out)**

<b>Item No</b>	<b>R Observed</b>	<b>R Table</b>	<b>Status</b>
Question 1	0, 383	0, 3440	Valid
Question 2	0, 238	0, 3440	<b>Invalid</b>
Question 3	0, 534	0, 3440	Valid
Question 4	0, 146	0, 3440	<b>Invalid</b>
Question 5	0, 385	0, 3440	Valid
Question 6	0, 565	0, 3440	Valid
Question 7	0, 498	0, 3440	Valid
Question 8	0, 597	0, 3440	Valid
Question 9	0, 662	0, 3440	Valid
Question 10	0, 611	0, 3440	Valid
Question 11	0, 385	0, 3440	Valid
Question 12	0, 435	0, 3440	Valid
Question 13	0, 529	0, 3440	Valid
Question 14	0, 633	0, 3440	Valid
Question 15	0, 582	0, 3440	Valid

Question 16	0, 213	0, 3440	<b>Invalid</b>
Question 17	0, 394	0, 3440	Valid
Question 18	0, 501	0, 3440	Valid
Question 19	0, 569	0, 3440	Valid
Question 20	0, 623	0, 3440	Valid
Question 21	0, 617	0, 3440	Valid
Question 22	0, 603	0, 3440	Valid
Question 23	0, 608	0, 3440	Valid
Question 24	0, 413	0, 3440	Valid
Question 25	0, 077	0, 3440	<b>Invalid</b>
Question 26	0, 452	0,3440	Valid
Question 27	0, 237	0,3440	<b>Invalid</b>
Question 28	0, 535	0,3440	Valid
Question 29	0, 629	0,3440	Valid
Question 30	0, 495	0,3440	Valid

## 2) Speaking test

In deciding the subject and the content of the speaking test, the researcher choose was on the chapter two in basic competence number 4.12, descriptive text material which focus on speaking performance in providing information (describing person orally). The material chosen had been learned by the students of the second grade students at MTs Ma'arif Udanawu Blitar. The subject of the speaking test was presented from the material which selected from the syllabus, 13' curriculum and the year program of the second grade students at MTs Ma'arif Udanawu Blitar.

### b. Construct Validity

In designing instrument of the test, it should fulfill the requirement of construct validity. Brown (2004: 25) comments if a test can be shown that it measures only the ability to measure, a test is said to have

construct validity. The word 'construct' refers to any fundamental ability that is speculated in a language ability theory. He added that any theory, hypothesis, or model attempting to explain observable facts is a construct. In line with Brown, James Dean (2000: 9) writes that construct validity has historically been defined as the experimental demonstration that a test measures the construct it intended to measure. From those explanation, it can be conclude that construct validity can be define as how the instrument of research represent the pattern of apect that being tested. In order to prove the construct validity, the researcher has tried out both instruments used in this study (questionnaire and speaking test) as follow:

1) Questionnaire

The self-esteem questionnaire fulfilled the construct validity because it the form is questionnaire form with 30 items (for try out) and 25 items (for research) which measures positive and negative feelings about the self and also problem in speaking which usually faced by the students.

2) Speaking test

The speaking test has fulfilled the construct validity because it test the students' speaking performance through providing information about something to measure their speaking performance. The students were asked to speak about providing information about describing one



of her or his classmate (for try out) and describing one of Indonesian actors or actresses (for research) in a video.

### **G. Reliability**

Reliability is how far the the consistent the result of the test. Creswell (2008: 168) stated that if the degree of instrument's scores are stable and consistent, it means that the test had been fulfilled the requirements of reliability. Meanwhile, Ary Donald et al. (2010: 236) wrote that if instrument measures what it supposed to measure and it shows the degree of consistency, the instrument is reliable. It can be implied that reliability is the consistent and the dependable of the result of the test. Then, the following classification from Harris (1977: 14) was used in measuring the level of reliability.

**Table 3.8 Level or Reliability**

<b>Score</b>	<b>Level</b>
0.81 – 1.00	Very Good
0.61 – 0.80	Good
0.41 – 0.60	Fair
0.21 – 0.40	Low
0.00 - 0.20	Empty

In proving the reliability, the researcher has tried out both instruments used in this study (questionnaire and speaking test) as follow:

#### 1) Questionnaire

In obtaining the reliability of self-esteem questionnaire, SPSS 24.0 Program was used. The following table was the result of reliability test of self-esteem questionnaire.

**Table 3.9 Reliability Statistics Result of Self-esteem Questionnaire**

<b>Reliability Statistics</b>	
<b>Cronbach's Alpha</b>	<b>N of items</b>
.883	30

If Cronbach's Alpha  $> 0,60$  the questionnaire is declared reliable. And if Cronbach's Alpha  $< 0,60$  the questionnaire is declared not reliable. As shown in the table above, the value of Cronbach's alpha was 0.883. It means that the value of self-esteem questionnaire was higher than 0.60, which was the standard Cronbach's alpha. It can be inferred that the self-esteem questionnaire was fulfill the requirement of reliability in very good level.

## 2) Speaking test

The speaking test also fulfill the requirement of reliability. In deciding the quality of reliability, Creswell (2008: 168) wrote there are five categories of reliability, those are test-retest reliability, alternative forms reliability, alternate forms and test-retest reliability, inter-rater reliability, and internal consistency reliability. In this study, the researcher used inter-rater reliability in explaining the reliability of speaking test when it used two raters to score and rank the students' speaking skill. Gay at al. (2000: 169) observed that reliability of inter-rater refers to the accuracy of two or more independent ratings, raters or observers. So, there were two correctors in scoring student's speaking test to produce two raw scores. The two correctors were the researcher

and the English teacher of MTs Ma'arif Udanawu Blitar who guides the reasearcher during conducts this study.

After that, the speaking scores form rater 1 correlated with the speaking scores from rater 2. Then, the result of each student's score was calculated with the following formula :

**Figure 3.2 Speaking Final Score Formula from Rater 1 and Rater 2**

$$RS = \frac{\text{rater 1} + \text{rater 2}}{2}$$

To measure the reliability of speaking test from both of raters, Cohen's Kappa Coefficient was used. Try out for speaking test was conducted in 8I class (excellent class) via online using video and Whats App. From 39 students, there are 9 students who do not submitted the speaking video. Then, there are 30 students who submitted the speaking video. Those number of students are enough to make the data of speaking test can be analyzed. In proving the inter-rater reliability, the score of score of speaking test was tested with Cohen's Kappa Coefficient. Fleis (1981: 10) interpreted the Kappa value as the following table :

**Table 3.10 Kappa Interpretation**

<b>Kappa Index</b>	<b>Agreement</b>
< 0,40	Bad
0,40 – 0,60	Fair
0,60 – 0,75	Good
> 0, 75	Excellent

The data was analyzed to measure Cohen's Kappa Coefficient using SPSS 24.0 Program and the result as the following :

**Table 3.11 Reliability Testing Result of Speaking Test**

		Symmetric Measures			
		Asymptotic		Approximate	
		Value	Standard Error <sup>a</sup>	T <sup>b</sup>	Significance
<b>Measure of Agreement</b>	<b>Kappa</b>	,750	,085	8,350	,000
<b>N of Valid Cases</b>		40			

The table above shows the result of reliability speaking test. The value of Kappa agreement was 0,750. It means that the speaking test fulfilled the requirement of reliability at good level of agreement.