# CHAPTER III

RESEARCH METHODOLOGY

3.1. Research Design

Research design is commonly defined as the way how the researcher gets the valid data of the research. This research is quantitative research, it is focused on the product (result of the test) not the process of teaching learning and the objective is to find out the correlation between students schemata and their reading comprehension.

According to Ary (2010:22) quantitative research uses objective measurement to gather numeric data that are used to answer questions or test predetermined hypotheses. It generally requires a well-controlled setting. In addition, Creswell (2012:619) correlational research are quantitative designs in which investigators use a correlation statistical technique to describe and measure the degree of association between two or more variables or sets of scores.

This kind of correlation is *Pearson Correlation* because the kind of data correlation is ordinal and interval. There were two variables in this research, independent and dependent variables. Hartono said that independent variable is a variable that gives influence, and dependent variable is the one that is affected by independent variable. The independent variable is students' content schemata (X) and dependent variable is reading comprehension (Y).

STUDENTS' CONTENT SCHEMATA (X)	⟷	READING COMPREHENSION ON EXPOSITORY TEXT (Y)
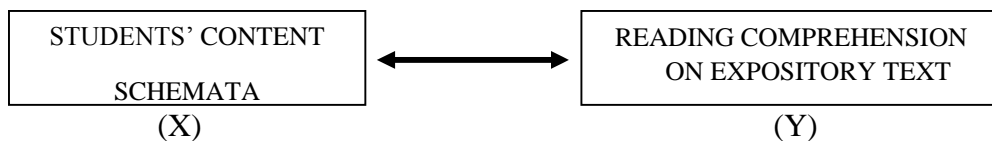
3.2. Population and Sample of the Research

Arikunto (2010: 72) said the overall population is the subject of research. Agreeing with this notion, Sugiyono (2010:117) states that population is a generalization region involves an object or a subject that has quality and special characteristics in which chosen by the researcher to investigate and make conclusion. It means that population refers to large group. The population of this research was the second year students of MAN 2 Jombang with total number 104 students. They were divided into 4 classes, XI IPA1, XI IPA 2, XI IPS1, XI IPS 2.

Slovin formula was used to choose as the participant. Slovin formula is a formula or formula to calculate the minimum sample size if the behavior of a population is not known with certainty. This formula was first introduced by Slovin in 1960. This Slovin formula is commonly used in survey research where the number of samples is usually very large, so a formula is needed to get a small sample but can represent the entire population. One of the methods used to determine the number of samples is using the Slovin formula, the Slovin formula is used if the total population is known.

$n = N/(1+Ne^2)$

n : number of samples

N : population

e : fault tolerance limit (Error tolerance)

Based on the notation of the minimum research sample size formula by Slovin above, if we have 104 people in a population, we can determine the minimum sample to be studied. The set margin of error is 5% or 0.05.

$n = N/(1+Ne^2)$

$n = 104 / (1 + (104 \times 0,05^2)$

$n = 104 / (1 + (104 \times 0,0025))$

$n = 104 / (1 + 0,26)$

$n = 104 / 1,26$

$n = 82.5396$

So If rounded off, the minimum sample size of 104 populations at a margin of error of 5% is 83. The amount of the sample is shown in the table as follows:

Table III.1

Population and Sample of the Research

No.	Class	Population
1	XI IPA 1	20
2	XI IPA 2	21
3	XI IPS 1	21
4	XI IPS 2	21
TOTAL		**83**

3.3. Technique of the Data Collection

After the design of the research was designed, the next step was to draw up the instrument. Research instrument occupies the most important position on the field to obtain the data. Research instrument was defined as tool to measure the nature or social phenomena being observed (Sugiyono, 2014:148). Here, it is purpose as a tool to get the data. There are two instruments used to conduct this research. They are:

3.3.1. Questionnaire

To determine the students' content schemata, the instrument used questionnaire. The items of the questionnaire were constructed based on the indicators. The questions consisted of 20 items adopted from Al-Jahwari, & Al-Humaidi (2015). The questionnaire represented the students' answers.

According to Rensis Likert, Likert scale is the most widely used scale in survey research and certainly the one that has found its way into popular culture. The classic use of the Likert scale was to pose questions or items to participants and have them respond using an agreement scale by selecting a number that best represented their response. Therefore, to determine students' content schemata can be scaled into five categories as follows:

Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	2	3	4	5

3.3.2. Multiple Choice Test

The test was used to measure students' reading comprehension on expository text. In teaching reading in our curriculum (K13), if the students are able to achieve goal, this means that assessment of reading ability needs to be correlated with purposes of reading.

According to Hughes, there are many techniques that can assess the students' comprehension but the writer used multiple choices technique. Multiple choices technique is a technique that will be designed by using four choices and the respondent chooses one based on the question. This technique can assess the student's reading comprehension. In this research, the writer gave twenty five

questions for the respondent. They were based on the indicators of reading comprehension in operational concept. There were five indicators in reading comprehension and for each indicator the writer made 5 questions. Due to the Covid-19 virus, the expository text test was carried out by the teacher concerned and monitored by the researcher.

3.4. Validity and Reliability

3.4.1. Validity of the Test

Every test, whether it is a short, informal classroom test, or a public examination should be as valid as the test constructor can make it. The instrument of the test must aim at providing a true measure. The instrument of the test is valid if the instrument used can measure the thing that will be measured.

The researcher administered try out twice, the researcher carried out in other class that was not included in sample. The purpose of try out was to obtain validity and reliability to the test. It was determined by finding the difficulty level of each item.

$$r_{xy} = \frac{\sum xy}{\sum x^2 \; \sum y^2}$$

Where:

r_{xy} = Correlation product moment x and y

$\sum xy$ = Total x and y

X^2 = X quadrant

Y^2 = Y quadrant

In the previous study the validity test of this instrument is in 0.768, it means that the validity is *Good*.

According to Arikunto the ranges of validity are:

Table III.2

The Classification of Validity

No	Classification	Score
1	Excellent	0,800-1,00
2	Good	0,600-0,800
3	Fair	0,400-0,600
4	Poor	0,200-0,400
5	Very Poor	0,00-0,200

(*Arikunto,2010*)

A test can be said valid if the test measures the object which can be measured and it is suitable for the criteria. Because of the questionnaire has been tested previously, the researcher tested a new set of questionnaire about content schemata test using SPSS software version 16.0 for windows programme. The items of the questionnaire is valid if the value of the coefficient of the validity is higher than r table or lower than significant tolerance 5% or 0.05. This validity test used 83 students as a population.

3.4.2. Reliability

A test must first be reliable as measuring instrument. Reliability is a necessary characteristic of good test. According to Brown, a reliable test is consistent and dependable. Reliability has to do with accuracy of measurement. The test was reliable when an examinee's result was consistent on repeated measurement.

To obtain the reliability of the test given, the researcher used Kuder Richardson 20 (K-R 20) formula to calculate the reliability of the test. The formula is as follows:

$$r_{11} = \frac{n}{n-1} \quad \frac{S^2 - \sum pq}{S^2}$$

Where:

r_{11} = Reliability

P = Proportion the correct scores

q = Proportion the incorrect scores

$\sum pq$ = Total of p times q

n = Total items

S = Variance total of the test

To know the test is reliable or not, the value of r11 must be compared with r product moment. The value of r11 must be higher than rtable. From the previous study, the value of r11 is 0.977. Then the rtable at 5% grade of significance is 0.381. While rtable at 1% grade significance is 0.487. So, it can be concluded that **0.444<0.977>0.561**. In other words, the instrument is reliable because the value of r11 is higher than rtable.

According to (Ary et al, 2010) said that the reliability of measuring instrument is the degree of consistency with measures whatever it is measuring. To make sure that the questionnaire is reliable, the writer uses *Cronbach's Alpha Coefficient* between 0 and 1. The result of the Content Schemata questionnaire

showed that the alpha is 0.922 ($\alpha > 0.7$). It can be concluded that the Content

Schemata questionnaire was reliable.

Table III.3

Reliability of Content Schemata

Reliability Statistics

Cronbach's Alpha	N of Items
.922	20

3.4.3. Normality

Because of the sample taken are 83 respondents, so the normality testing

uses manual computation by using One-Sample Kolmogorov-Smirnov test that

calculate by SPSS 16.0 for windows.

Table III.4

One-Sample Kolmogorov-Smirnov Test

		Unstandardized Residual
N		83
Normal Parameters[a]	Mean	.0000000
	Std. Deviation	8.99636132
Most Extreme Differences	Absolute	.145
	Positive	.069
	Negative	-.145
Kolmogorov-Smirnov Z		1.324
Asymp. Sig. (2-tailed)		.060

a. Test distribution is Normal.

From the data computation above, can be interpret that if the significance score is > 0.05 it means that the normality data is normal distribution. If the significance score is < 0.05 it means that the normality data is not normal distribution. From the data computation above, we can see that the significance score of One-Sample Kolmogorov-Smirnov test is 1.324, it means that the normality data is normal distribution.

3.5. Technique of the Data Analysis

For the technique of data analysis, the researcher applied a quantitative analysis. According to David Nunan(2002), quantitative research describes a research problem thought a description of trends or a need for an explanation of the relationship among variables by collecting numeric data from a large number of people using instruments with present question and responses. Considering both variables as ordinal and interval data, thus, the writer used *Pearson Correlation* formula in analyzing and calculating the data. The collective data were analyzed by using the following formula:

$$r_{ser} = \frac{\sum O_r - O_t \ M}{\sum \frac{O_r - O_t{}^2}{p}}$$

Where :

r_{ser} : Serial Correlation Coefficient

SD_{tot} : Standard Deviation

O_r : The Lower Ordinate

O_t : The Higher Ordinate

M   : Mean Score

p   : Individual

Furthermore, to categorize how big strength of correlation between content schemata and reading comprehension on expository text of the second year students at MAN 2 Jombang, the researcher used categories in Hartono's book. They are:

Table III.5

Interpretation of Correlation Coefficient

Coefficient Interval	Level of Correlation
0.00-0.200	Very Low
0.200-0.400	Low
0.400-0.700	Middle
0.700-0.900	Strong
0.900-1.00	Very Strong

(*Hartono,2008*)

The data of students' content schemata should be calculated by formula below:

$$P = \frac{F}{N} \times 100\%$$

Where :

P   : Percentage

F   : Frequency

N   : Number of items

Furthermore, the data of reading comprehension should be scored byformula below:

$$\underline{S} = \frac{R}{N} \times 100\%$$

Where:

    S   : Individual Score

    R   : Number of Correct Answer

    N   : Number of Items

Table III.6

The Classification of Students' Score

The Score Level	Category	Level
80-100	Very Good	High
66-79	Good	Middle
56-65	Enough	Low
40-55	Less	Very Low
30-39	Fail	Bad

(Arikunto,2010)

Meanwhile, in order to get easy in analyzing the data, the researcher used SPSS 17.0 program to Windows. The product moment correlation coefficient was obtained by considering the degree of freedom (df) = N-nr; (N= number of sample, nr = number of variable)

Statistically, the Hypotheses are:

Ha: > r table

Ho: ro< r table

Ha is accepted if *ro>* r *table* or there is a significant correlation between the students' content schemata and reading comprehension on expository text.

Ho is accepted if *ro<*r *table* or there is a significant correlation between the students' content schemata and reading comprehension on expository text.

46