CHAPTER III

RESEARCH METHOD

This chapter presents the research method used in conducting the study. It covers research design, population, sampling and sample, variable data, data source, data collecting method and instruments and technique of data analysis.

A. Research Design

In this study, the writer used an experimental research design with two groups used a quantitative approach. The experimental research design is a research design intended to test the effectiveness of something. According to Creswell (2008:3) research is a process of steps used to collect and analyze information to increase our understanding of a topic or issues". Thus, the research method is an important part of a research to find out the scientific truth. The experimental method is a systematic and scientific approach to research in which the researcher manipulates one or more variables, and controls and measures any change in other variables. This is an experiment where the researcher manipulates one variable, and control/randomizes the rest of the variables. It has a control group, the subjects have been randomly assigned between the groups, and the researcher only tests one effect at a time. The researcher wants to know the usage of Krebs Media in improving student's speaking skill.

In this research, the researcher used a Quasi-Experimental research design with quantitative approach. This design-focused on treatment and

outcome. The data was collected from pretest and post-test in order to know whether the usage of Krebs Media in improving student's speaking skill. The table below shows the design of the research.

Table 3.1:

The Design of the Reasearch

Group	Y1	X	Y2
Experimental Group (8.7)	Pre-test	Krebs Media Technique	Student's Speaking Ability
Control Group (8.9)	Pre-test	Conventional Technique	Student's Speaking Ability

Based on the Table 3.1, the procedures of using two groups pre-test post-test design were:

1. Administering a pre-test to both classes (8.7 and 8.9) to measure the score of speaking mastery of the students at second grade in MTsN 1 Tulungagung.

2. Applying the experimental treatment in 8.7 class to teach speaking by using Krebs Media and applying control treatment in 8.9 to teach speaking by using conventional method to the student of second grade in MTsN 1 Tulngagung.

3. Administering a post-test in both classess (8.7 and 8.9) to measure the score of vocabulary mastery of the students at second grade in MTsN 1 Tulungagung.

B. Population, Sampling and Sample

1. Population

Population is all subjects being studied. Parahoo (1997:218) defines the population as "the total number of units from which data can be collected", such as individuals, artifacts, events or organizations. Burns and Grove (2003:213) describe population as all the elements that meet the criteria for inclusion in a study. While sample is part of the population that is being studied.

The illustration of population and sample
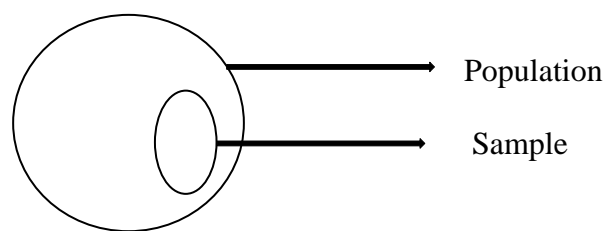
The population of this research was students in the seventh grade of Islamic Junior High School of Tulungagung 1 which has around 800 students. The school has nine second grade classes with 200 students. They were 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9 in academic years of 2019/2020 each class consists different number of students.

2. Sampling

Sampling is a technique to take sample from group of population. Sampling an important characteristic of inferential, and statistics is the process of going from the part to whole (Ary et al, 2010:148). In this study, the researcher did not use all populations to be

sampled. Therefore, purposive sampling used to take samples. Purposive sampling technique is a type of non probability sampling where the researcher consciously selects subjects for addition in a study so as to make sure that the elements will have certain characteristics pertinent to the study. In purposive sampling, which also referred to as judgment sampling, sample elements judge to a typical or representative are chosen from the population (Ary, 2010:156). It was technique to determined sample with a particular consideration. Purposive sampling used based on a certain consideration and the main consideration was the chosen classess had homogeneous speaking mastery. In other words, the students in those classess had average proficiency in speaking.

3. Sample

It was impossible to use all the population as the sample due to some considerations. The researcher took two classes as the sample of the research. The research sample was selected by the cluster random technique. It is the sample selection in which all members of the population are naturally grouped in units (Wiersma and Jurs, 2009: 355).

Based on the description above, the researcher defined that sample was a small group taken from population with a certain media that was researched by researcher.

For the study was quasi experimental, the sample was choosen by applying purposive sampling. Applying this method, two classes were choose by using a certain criterion in which the choose classes must be

normal or in average. Based on the criterion the sample of this research as the students' of 8.9 as a control group and 8.7 as a experimental group at MTsN Tulungagung 1 in which total of 8.9 class was 39 students' and 8.7 was 39 students'. So, the total sample was 78 students'.

Table 3.2: The Research Sample by Class

No.	Class	The Number of Students
1	8.7	39
2	8.9	39

The sample above was divided into two groups. The first group was class 8.7 as the experimental group and the second group was class 8.9 as the control group. The experimental group was given the Krebs Media in their speaking activities. The researcher made lesson plans and some activities for every meeting based on the school-based curriculum. The control group was given the same materials but without using Krebs Media in their speaking activities. The students listening to the audio and the teacher gave short explanation about the rules of using Krebs Media. Then, the students made a conversation with their peer to present it in front of the class. The table below shows the distribution of treatment in the research.

Table 3.3: The Distribution of the Treatment

Group	Class	Treatment	Number of Students
Experimental	8.7	Krebs Media	39
Control	8.9	Conventional	39

C. Variable

Variable is anything that will be researched by the researcher. According to Frankel and Wallen (2006:40) variable is a concept a noun that stands for variation within a class of subject such as gender, colour, motivation, chair, eye, achievement, or running speed. Based on the title of this research, there were two variables:

1. Independent Variable

   Independent Variable is the one affecting another variable. In this research Krebs Media was an independent variable because it can affect the student's speaking ability.

2. Dependent Variable

   Dependent Variable is the one affected by another variable. In this research student's speaking ability was a dependent variable.

D. Research Instrument

An instrument is very important in every research. The function of the instrument is as a tool to get the data. The instrument which was used by the researcher was practice. The aim of the practice was to know the student's speaking ability before and after treated by using Krebs Media.

Gay, 1992: 154 stated that a test is not necessarily a written set of question to which an individual respond in order to determine whether he or she passes. It means that after the students are being tested, students can measure their ability whether they are passed the test or not. When students get high score means that they can pass the exam. In opposite, when the student's

score is low, means that they can't pass the exam or their ability still low. But the important measurement of this research is how far the student's speaking ability will change.

Since the study was quasi-experimental, the instruments to collect the data were a pre-test and a post-test. The collected data were the scores obtained from the pre-test and the post-test of both the control group and the experimental group. The scores from the pre-test were used to see the speaking ability of both classes before the treatment. On the other hand, the scores from the post-test were used to measure whether the implemented method affected the experimental group or not.

In this research study, the speaking test served as the research instrument. The speaking test was held twice, in the pre-test and the post-test. It was used to reveal the significant difference in the speaking ability between the eighth grade students of junior high school who were taught by using Krebs Media and those who were not.

In formulating the test instrument, the points to be considered are the relevance of the test instruments to the purpose of the study, and the relevance of the test instruments to the curriculum. The test was intended to measure students' speaking abilitybefore and after the treatment. The speaking tests were made based on the school-based curriculum for the seventh grade of the junior high school in the second semester. The syllabus for the seventh grade junior high school students in the speaking aspect was taken as the

considerations in formulating test instruments. In making the test, the researcher took some sources from some English books and the internet.

The instrument was a group of oral assessment in the form of simple transactional and interpersonal conversation/dialogues. The test was given twice. The first test was used to gain the pre-test score before the treatment and the second was used to gain the post-test score after the treatment. The score of the test was based on criteria on the speaking test rubric adapted from Hughes (1989). They are content, fluency, vocabulary, and grammar

E. Validty and Reliability Testing

Reliability and validity are interrelated and rely on many aspects. In a broad sense, Henning (1987, p. 198) defines validity as the extent to which a test measures the ability or knowledge that it is purported to measure.

1. Validity

As like what Underhill (1987, p. 9) says that Validity is concerned with whether a test measures what it is supposed to. Many important aspects of tests have a bearing on validity and reliability, and some worth mentioning here include backwash effects, face validity, content validity and construct validity. Hughes (1989, p. 1) states the effect of testing on teaching and learning is known as backwash.

This research was using face validity as the subjective measurement, construct validity as the criteria of a person who full filled

the success speaking ability, and content validity as a non-empirical expert judgment of the extent to which the content of a test is comprehensive and representative of the content domain purported to be measured by the test. And as the result, the result of this measurement on the speaking test could present an out of high value.

a. Face Validity

Henning (1987, p. 192) defines face validity as a subjective impression, usually on the part of examinees, of the extent to which the test and its format fulfills the intended purpose of measurement.

In this research, the researcher was using face validity by consulting to the expert, that was the English Teacher of MTsN 1 Tulungagung, as the validator. And the result between the expert and researcher could be seen on the following aspects of consideration in making the good test based on face validity:

1. The instrucion must be clear for he students.

2. The question must not be ambiguity to make students are able to answer it.

3. The time acllocation must be appropriate for each student who are tested.

b. Construct Validity

Henning (1987, p. 190) defines construct validity as the validity of the constructs measured by a test. Construct validity is related to content validity, in that it is concerned with the contents of the test and their wider context. Construct validity thus refers to

whether the test shares the same philosophy of the teaching program of which it is a part, and can be measured by both statistical and intuitive methods, according to Underhill (1987, p. 106) who adds that: construct validity is not an easy idea to work with...to reduce it to its simplest statement it says: does the test match your views on language learning? In practice, there may be little difference between construct and content validity.

The scoring rubric used in this study was adapted and modified from Hughes (1989).

Table 3.4: Cunstruct Validity

Criteria	Weight	Exemplary	Accomplished	Developing	Begining
Content	35%	The content is clear, related to dialy expresion and speak with high confidence.	The content is fairly clear, elated to dialy expresion and speak with high confidence.	The content is fairly clear, related to dialy expresion and the aspect which must be informed can't be delivered completely and looks uncomfortable.	The content is unclear can't tell the appropriate aspect, and looks nervous.
Fluency	30%	Speech on all professional	Speech is effortless and	Speech is frequently	Speech is very slow and eneven

		and general topics as effortless and smooth as a native speaker's	smooth but perceptibly non-native in spech and evenness.	hesitant, sentence may be left incomplate.	except for short sentences.
Vocabulary	20%	Effective choice of words and use of idoms and word forms.	Adequate choice of words but some misuse of vocabulary, idioms, and word forms.	Limited range, confused use of words, idioms, and word forms.	Very limited range, very poor knowledge of word, idioms, and word forms.
Grammar	15%	No more than three errors during dialogue activity.	Few errors with no patterns of failure	Frequent errors showing some major patterns uncontroled and causing occasional misunderstanding.	Gramar almost entirely inaccurate phrases.

c. Content Validity

Face validity is closely associated with content validity, defined by Henning (1987, p. 190) as usually a non-empirical expert judgment of the extent to which the content of a test is comprehensive and representative of the content domain purported to be measured by the test. Face and content value, therefore, refer

to the extent to which the test is recognizable as a fair test by learners, who thereby perform to their ability as a result. Tests that lack face and content cause negative backwash effects and result in student underperformance, as well as the results being contested by both teachers and learners.

The speaking validity test employed content validity. According to Wiersma and Jurs (2009: 355), content validity is the process of how the test establishes the representativeness of the items in a certain domain of skills, tasks, knowledge, and other aspects that are being measured. It means the test was developed in reference to the Standard of Competence and Basic Competence of junior high school year VIII of the second semester of English subject.

Table 3.5 Matrix of Content Validity

Syllabus	Indicator	Technique	Test Item
Basic Competence: Berkomunikasi secara interpersonal, transaksional dan fungsional tentang diri sendiri, keluarga, serta orang, binatang, dan benda, kongkret dan imajinatif yang terdekat dengan kehidupan dan kegiatan siswa seharihari di rumah,	- Siswa dapat mengungkapkan kalimat interpersonal - Siswa dapat mengungkapkan kalimat transaksional	Speaking test	-Pretest -Posttest

sekolah, dan masyarakat, serta terkait dengan mata pelajaran lain dan dunia kerja			

Based on the Table 3.4 above, the instrument of the test could be said have the content validity because the test has equal purpose with the core competence and basic competence in syllabus of Curriculum of 2013, which was testing the students' ability in vocabulary with the correct structures.

2. Reliability

Henning (1987, p. 198) defines reliability as the consistency of the scores obtainable from a test. It is usually an estimate on a scale of zero to one of the likelihood that the test would rank testees in the same order from one administration to another proximate one.

In this research, the writer used inter-rater reliability where the two scorers did the scoring and the two sets of scores gotten from the two scorers were collected to get the correlation coefficient. The two scorers were the researcher himself and his partner who was from the same major.the researcher chooses the rater because he can understand every point in the scoring rubrics.

In this research, the researcher conducted the reliability test with try out (post-test) by using pearson Product-Moment in SPSS 23.0 for getting correlation coefficient. The criteria of reliability instrument in pearson product-moment divided into 5 according to Hatch and Farhady (1982:247). They are:

a. A very low reliability ranges from 0.00 to 0.19
b. A low reliability ranges from 0.20 to 0.39
c. An average reliability ranges from 0.40 to 0.59
d. A high reliability ranges from 0.60 to 0.79
e. A very high reliability ranges from 0.80 to 1.00

The researcher only took 10 students to be a samples in conducting try out. The table of data on the table below was the score abtained from the try out. Rater 1 was taken from the researcher and the rater 2 was taken from his partner in conducting this research. After obtaining the two scores, the researcher did a reliability testing and got the result as presented on table below.

Table 3.6 The Scores Obtained from the Try Out (Post - Pre)

Subject	Rater 1	Rater 2
A	75	76
B	75	75
C	83	85
D	77	75
E	78	75
F	80	80

G	85	83
H	83	85
I	75	73
J	80	83

The result of reliability testing can be seen from table:

Table 3.7 The Result of Reiability

Correlations

		Rater_1	Rater_2
Rater_1	Pearson Correlation	1	,904[**]
	Sig. (2-tailed)		,000
	N	10	10
Rater_2	Pearson Correlation	,904[**]	1
	Sig. (2-tailed)	,000	
	N	10	10

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

The result of calculation showed that reliability coefficient was 0.904 and the ideal reliability coefficient was 1. In this research, the calculation was comparable to 1, it means the instruments of this research was very reliable because 0.953 closer the reliability coefficient to 1.

F. Normality and Homogenety Testing

Before analyzing the significant difference score in speaking of the students' taught by using Krebs Media and those taught by using conventional method, the data should be normal distribution and homogenous. Normality and Homogeneity is one of the pre-requisite tests

used to analyze data at independent sample t test. The purpose was to find out whether the data has been distributed normally and is there any difference variance in the two groups above. In addition, to measure the data computation were normal distribution and homogenous, the researcher conducted normality testing and homogenity testing. The definition as follow:

1. **Normality Testing**

    Normality tests are used to determine whether a data set is well-modeled by a normal distribution or not, or to compute how likely an underlying random variable is to be normally distributed. The purpose of using normality To know the normality, the researcher used *kolmogorv-smirnove test* with SPSS.23.0.*Kolmogorov-smirnove* is a test for normality for large samples.

2. **Homogeneity Testing**

    Homogeneity testing is intended to make sure that the collected manipulation data in analysis truly taken from population which is too different each other. To know the homogenity, the researcher used T test of Homogeneity of Variances with SPSS.23.0.

G. **Procedures of the Research**

    In this study, the treatment administered four meeting completed with pre-test and post-test since the researcher has no authority to conduct more than it moreover the class did not belong to the researcher herself. The

treatment was given after conducted the pre-test and before the post-test. To know the schedule of the research, it can be seen in Table 3.7 below:

Table 3.8 The Schedule of the Research

No	Group	Meeting	Date	Activity	Time
1.	Conrol (8.9)	I	Thursday, April 2nd 2019	Pretest	3-4
2.	Experimental (8.7)		Wednesday, April 3rd 2019	Pretest	3-4
3.	Conrol (8.9)	II	Saturday, April 6th 2019	Treatment 1 conventional	5-6
4.	Experimental (8.7)		Saturday, April 6th 2019	Treatment 1 by Krebs Media	7-8
5.	Conrol (8.9)	III	Thusday, April 9th 2019	Treatment 2 conventional	3-4
6.	Experimental (8.7)		Wednesday, April 10th 2019	Treatment 2 by Krebs Media	3-4
7.	Conrol (8.9)	IV	Saturday, April 13th 2019	posttest	5-6
8.	Experimental (8.7)		Saturday, April 13th 2019	posttest	7-8

The test was given by conducted pre-test and post-test which consisted with speaking test. Those test was done on fisrt meeting and the last meeting. While, the treatment was given after pre-test and before the post-test. In this study, the group was got the treatment by using Krebs Media is experimental

group only. Thus, the researcher would explain more about those treatment. The procedures of treatment can be seen as follow:

1. First treatment was conducted on April 6th 2019

    Before beginning applied the Krebs Media, the researcher introduced the media especially for teach speaking by listening to the audio of krebs media. Then, the researcher asked students' to prepare their own smartphone and would be received the audio of krebs media. After delivering the audio of krebs media to all sudents, then the researcher explained about the rules of using the audio for their lerning.

2. Second treatment was conducted on April, 9th and 10th 2019

    In the second meeting, the treatment was same with the first meeting. The students' were still listened the audio of krebs media and the researcher explained interpersonal and transactional dialogue.

H. Data Collecting Method

Data collecting method is the way to collect the data. The researcher used two kinds of tests, they were pretest and posttest.

1. Pre-Test

    Pre-test was given to the students before the researcher taught by Krebs Media. Pre-test is needed to know how far the students'speaking ability in English subject without Krebs Media. The form of pre-test is speaking test which consist 4 items about interpersonal dialogue and transactional dialogue. The pre-test given to know the basic competence

for students' and to know them earlier knowledge before they get treatment. The students' answered the task of pre-test is 60 minutes.

2. Post-Test

After the treatment, the post test was given to the students'. The test item in the post-test was same. This test is to measure students' speaking skill after treatment applied. The form post-test was also speaking test which consist 4 items about interpersonal dialogue and transactional dialogue. It was given to know the final score and the students' difference achievement before and after they get treatment. Time allocation to answer the task is 60 minutes.

I. Data Analysis

The collected data were analyzed to know the effectiveness of using Krebs Media toward students' speaking skill. The researcher divided the test result into two groups, they were the test result from the experimental group and the test result of the control group. Data obtained from the posttest from both of Experiment class and Control class would be analyzed statistically using *Independent-Sample T-Test* through SPSS 23.0 for windows. The researcher used t-test to know the significant value was higher or smaller than 0.05. The technique of data analysis used by the researcher belonged to quantitative data analysis.