CHAPTER III

RESEARCH METHOD

In this chapter, the researcher presents research design, population and sample, research instrument, validity and reliability testing, normality and homogeneity testing, data collecting method, and data analysis.

A. Research Design

The research design of this research is pre-experimental with one group pretest-posttest design because it has little or no control extraneous variable. Ary, et al (2010: 302) states that pre-experimental design do not have random assignment of subjects to groups or other strategies to control extraneous variables. In one-group pretest-posttest design, the result of pretest and posttest will be compared to know the significant difference before and after applying treatment.

Pre-experimental design involve two variables, they are independent and dependent variable. According to Cohen, et al (2007: 272) experiment involves making a change of one variable (independent variable) and observing the effect of that change on another variable (dependent variable). Creswell (2012: 295) adds that "in an experiment, you test an idea (or practice or procedure) to determine whether it influences on outcome or dependent variable". The independent variable of this research is Think-Pair-Share strategy. While, the dependent variable is students' speaking ability.

The procedures of pre-experimental with one group pretest-posttest design in this research are:

1. Administering pretest to measure students' speaking ability before giving treatment.
2. Applying the experimental treatment of teaching speaking by using Think Pair Share strategy.
3. Administering posttest to measure students' speaking ability after giving treatment.

The design of this research can be seen at the table below:

Table 3.1 The Design of One Group Pretest-posttest

Pretest	Independent	Posttest
Y1	X	Y2

(Adapted from Ari et al, 2010: 304)

Based on the table above, experimental design used pre-experimental research with one group pretest-posttest consist of pretest (Y1), treatment (X), and posttest (Y2). This research intended to investigate the effectiveness of using Think Pair Share strategy on students' speaking ability of the eighth grade at MTs Darul Hikmah Tulungagung.

B. Population and Sample

1. Population

The first step in selecting sample is determine the population. Before selecting the sample, the researcher need to determine the population. Creswell (2008: 151) defines that "population is a group of individuals who have the same characteristics". Population is consists of

every member of groups that have been determined as the subject of study by the researcher (Irianto, 1989: 18).

The population of this research is all of the students of the eight grade at MTs Darul Hikmah Tulungagung which consists of four classes (A, B, C, and D) which consist of 137 students.

2. Sample and sampling

After knowing the population, the next step was selecting sample of the research. The researcher took samples as the representative of the population. According to Hadi (1990: 9) Sample is portion or representative of the population that will be observed. Creswell (2012: 142) adds that "sample is subgroup of the target population that the researcher plans to study for the purpose of making generalization about the target population". In this research, the researcher took one class of the eight grade at MTs Darul Hikmah Tulungagung. That is the students of class 8B which consists of 32 students.

The sample of this research is taken by using purposive sampling because the English teacher recommend to conduct the research in class 8B based on consideration that the students in that class have average ability in speaking. The meaning of average ability in this research is the students' speaking ability is not too good and too bad. Cohen (2007: 144) state that "in purposive sampling, the researcher handpick the cases to be included in the sample on the basis of their judgment of their typicality or possession on the particular characteristics being sought".

C. Research Instrument

In collecting the data, the researcher needed to use the research instrument. According to Arikunto (2002: 136) research instrument is a device used by the researcher while collecting data to make his work become easier and get a better result complete and systematic in order to make the data easy to be processed. Creswell (2012: 14) adds that "an instrument is a tool for measuring, observing, or documenting quantitative data. It contain specific questions and response possibilities that you establish or developed in advance of the study".

In this research, the researcher used test as instrument in collecting the data. The test used to collect information about the students' speaking ability before and after the students were being taught by using Think Pair Share strategy. Before conducted test, the instrument was validated by advisor and English teacher. Then, the researcher tried out the instrument to twenty students and analyzed the result of try out by using Pearson Product Moment with SPSS 16.0 to know whether the instrument is reliable or not reliable. After the instrument is reliable, the researcher conducted test to all students of class 8B at MTs Darul Hikmah Tulungagung.

There were two test in this research, they were pre-test and post-test. The test that was used is speaking test. The format of test is presentation. Pre-test was given before the students were taught by using Think-Pair-Share strategy to know the students' speaking ability before they got the treatment. The students were asked to present their idea about the effect of smoking in

front of the class. While, post-test was given after the students were taught by using think pair share strategy to know the students' ability after they got the treatment. The students were asked to present the effect of using media social. In assessing the students' speaking ability, the researcher used the scoring rubrics which include the criteria such as accent, grammar, vocabulary, fluency, and comprehension. (See appendix 4)

D. Validity and Reliability Testing

This study used a test as the research instrument to measure students' speaking ability. As the instrument, the test should has the characteristics to get the data is good. That characteristics are validity and reliability. Ary et al (2002: 213) support that there are two important characteristic that every measuring instrument should process: validity and reliability. The validity and reliability of this research will be explained as follows:

1. Validity

Validity is the most complex criterion of an effective test and the most important principle of language testing (Isnawati, 2014: 27). Cohen (2007: 133) states that "validity should be seen as a matter of degree rather than as an absolute state. To measure whether the test has good validity, the researcher analyzed that test from content validity, face validity and construct validity.

a. Content validity

The test is said to have content validity if the content of test represent the purpose of the test. The test to measure students'

speaking ability is speaking test. The form of this test presentation activity. Isnawati (2014: 27) state that "the test will have content validity if it includes a proper sample of the structure or content which is relevant with the purpose of the test".

This research has content validity because the test is designed refer to the syllabus of the eight grade that is used in MTs Darul Hikmah Tulungagung. Based on the syllabus Kurikulum 2013, the students of eight grade of Junior High School are expected able to understand and arrange oral and written text to state and ask the cause effect of action or event around them and opposite relationship. It is included on basic competence 3.8 and 4.9. This test focus on cause effect material and oral text in basic competence 4.9. The students was given speaking test to present the idea about cause and effect of the topic that given by the research. The content validity of this research can be seen at the table below:

Table 3.2 Content validity of test

Basic Competence	Indicator	Instrument	Item Test
4.9 Arrange oral and written text to explain and ask about cause effect and opposite relationship by concernin	4.9.1 Understanding social function, structure text, and language element about cause effect relationship. 4.9.2 Arrange and present oral text of cause effect relationship about topic that was given by	Speaking Test	**Pre-test** : Present the effect of smoking. You have 3 minutes to present that topic. **Post-test** : Present the effect of using media social. You have 3 minutes to present that topic.

g social function, structure text, and language element correctly and based on context.	concerning social function, structure text and language element.		

Based on table above showed that the instrument of the test was valid based on the core curriculum, based competence, and indicator which mention in syllabus.

b. Face validity

A test is said to have face validity if it looks as if it measures what it is supposed to measure. Face validity is hardly a scientific concept, yet it is very important. A test which does not have face validity may not be accepted by test-takers, teachers, education authorities or employers (Isnawati. 2014: 29). The researcher used face validity by consulting with advisor and English teacher of the eight grade at MTs Darul Hikmah Tulungagung.

c. Construct validity

Construct validity is one kind of validity that is measure the ability which is supposed to measure. The word ‘construct’ refers to any underlying ability which is hypothesized in the theory of language ability (Isnawati, 2014: 29). Based on theory above, in the test the researcher asked the students to present their idea about topic that was given. The students used speaking scoring rubric which involves accent, grammar, vocabulary, fluency, and comprehension

to assess the students' speaking ability. This is fulfill the construct of speaking test. So, this instrument is valid in term of construct validity.

Table 3.3 Speaking Scoring Rubric

Aspect	Score	Proficiency Description
Accent	1	Pronunciation frequently unintelligible
	2	Frequent gross errors and a very heavy accent make understanding difficult, require frequent repetition.
	3	"Foreign accent" requires concentrated listening, and mispronunciations lead to occasional misunderstanding and apparent errors in grammar or vocabulary.
	4	Marked "foreign accent" and occasional mispronunciations which do not interfere with understanding.
	5	No conspicuous mispronunciations, but would not be taken for a native speaker.
	6	Native pronunciation, with no trace of "foreign accent".
Grammar	1	Grammar almost entirely inaccurate except in stock phrases.
	2	Constant errors showing control of very few major patterns and frequently preventing communication.
	3	Frequent errors showing some major patterns uncontrolled and causing occasional irritation and misunderstanding.
	4	Occasional errors showing imperfect control of some patterns but no weakness that cause misunderstanding.
	5	Few errors, with no patterns of failure.
	6	No more than two errors during interview.
Vocabulary	1	Vocabulary inadequate for even the simplest conversation.
	2	Vocabulary limited to basic personal and survival areas (time, food, transportation, family, etc.).
	3	Choice of words sometimes inaccurate, limitations of vocabulary prevent discussion of some common professional and social topics.
	4	Professional vocabulary adequate to discuss special interest; general vocabulary permits discussion of any non-technical subject with

		some circumlocutions.
	5	Professional vocabulary broad and precise; general vocabulary adequate to cope with complex practical problems and varied social situations.
	6	Vocabulary apparently as accurate and extensive as that of an educated native speaker
Fluency	1	Speech is so halting and fragmentary that conversation is virtually impossible.
	2	Speech is very slow and uneven except for short or routine sentences.
	3	Speech is frequently hesitant and jerky; sentences may be left uncompleted.
	4	Speech is occasionally hesitant, with some unevenness caused by rephrasing and groping for words.
	5	Speech is effortless and smooth, but perceptively non native in speed and evenness.
	6	Speech on all professional and general topics as effortless and smooth as a native speaker's.
Comprehension	1	Understands too little for the simplest type of conversation.
	2	Understand only slow, very simple speech on common social touristic topics; requires constants repetition and rephrasing.
	3	Understands careful, somewhat simplified speech when engaged in a dialogue, but may require considerable repetition and rephrasing.
	4	Understands quite well normal educated speech when engaged in a dialogue, but requires occasional repetition or rephrasing.
	5	Understands everything in normal educated conversation except for very colloquial or low frequency items, one exceptionally rapid or slurred speech.
	6	Understands everything in both formal and colloquial speech to be expected of an educated native speaker.

(Adapted from Hughes**,** 2003: 113)

2. Reliability

Reliability is used to know whether the test is consistent and reliable. Creswell (2012: 159) explains that score from an instrument are

stable and consistent. Score should be nearly the same when researchers administer the instrument multiple times at different times.

The researcher used inter rater reliability in which the researcher used two raters to test speaking ability. Inter-rater reliability is achieved when two scorers or two raters do the scoring. Then, the two sets of scores gotten from the two raters are calculated to get the correlation coefficient (Isnawati, 2014: 23). The researcher calculated two sets of score in try out pre-test and try out post-test to get the correlation between them by using *Pearson Product Moment formula* with SPSS 16.0. The result of reliability testing can be seen in the table below:

Table 3.4 Correlation of pre-test score (try out)

Correlations

		Rater_1	Rater_2
Rater_1	Pearson Correlation	1	.763**
	Sig. (2-tailed)		.000
	N	20	20
Rater_2	Pearson Correlation	.763**	1
	Sig. (2-tailed)	.000	
	N	20	20

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

Table 3.4 show that Pearson Correlation is 0.763 and numeral significance is 0.000. The result of Pearson correlation (0.763) is closer 1 and the numeral significant is lower than (0.000 < 0.05). It means that the test was reliable.

Table 3.5 Correlation of post-test (try out)

Correlations

		Rater_1	Rater_2
Rater_1	Pearson Correlation	1	.912**
	Sig. (2-tailed)		.000
	N	20	20
Rater_2	Pearson Correlation	.912**	1
	Sig. (2-tailed)	.000	
	N	20	20

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

Table 3.5 show that Pearson Correlation is 0.912 and numeral significance is 0.000. This result of Pearson correlation (0.912) is closer 1 and the numeral significant is lower than (0.000 < 0.05). It means that the test was reliable.

E. Normality and Homogeneity Testing

1. Normality Testing

Normality testing is needed to find out whether the data is in normal distribution or not. It is intended to show that the sample data come from a normality distributed population. To know the normality, the researcher used *One-Sample Kolmogorov-Smirnov test* in SPSS 16.0 with significance value (α) = 0.05. The normality testing was done towards the pretest and posttest score in tryout. The hypothesis for testing normality as follow:

a. H_0 : If the value of significance > 0.05, means data is in normal distribution.

b. H_1 : If the value of significance < 0.05, means data is not in normal distribution.

The result of normality testing with *One Sample Kolmogorov-Smirnov Test* can be seen in table 3.6 below:

Table 3.6 The Result of Normality Testing

One-Sample Kolmogorov-Smirnov Test

		Pretest	Posttest	Unstandardized Residual
N		20	20	20
Normal Parameters[a]	Mean	50.10	74.00	.0000000
	Std. Deviation	5.684	7.377	5.18610780
Most Extreme Differences	Absolute	.195	.207	.108
	Positive	.165	.111	.108
	Negative	-.195	-.207	-.105
Kolmogorov-Smirnov Z		.872	.925	.484
Asymp. Sig. (2-tailed)		.432	.359	.973

a. Test distribution is Normal.

Based on the table 3.6 above, the significance value of pre-test is 0.432 and the significant value of post-test is 0.359. Both significance value from pre-test and post-test are bigger than 0.05. The significance value of pre-test is bigger than 0.05 (0.435 > 0.05) and significance value of post-test is bigger than 0.05 (0.359 > 0.05). So, it is concluded that test is in normal distribution and residual has normal distribution.

2. Homogeneity Testing

Homogeneity testing is intended to show that two or more groups of data samples come from populations having the same variance. To know the homogeneity, the researcher used *Test of Homogeneity of Variances* in SPSS 16.0 with significant value (α) = 0.05. The

Homogeneity Testing was done towards the pretest and posttest score in tryout. The hypothesis of testing homogeneity as follow:

a. H_0 : If the value of significance > 0.05, means data is homogeny

b. H_1 : If the value of significance < 0.05, means data is not homogeny

The result of homogeneity testing with *Test of Homogeneity of Variance* can be seen in table 3.7 below:

Table 3.7 The Result of Homogeneity Test

Test of Homogeneity of Variances

Pretest

Levene Statistic	df1	df2	Sig.
2.225	5	11	.125

Based on the table above, the significant value is 0.125. This value is bigger than 0.05 (0.125 > 0.05). It means that H_0 is accepted and H_1 is rejected. So, it is concluded that the test is homogeny.

F. Data Collecting Method

Data collecting method is process to collect the data in the research. To get the data, the researcher used method of data collecting as follow:

a. Pre-test

Pre-test was conducted at 4th February 2017. Pre-test was given before the students were taught by using Think-Pair-Share strategy to know the students' speaking ability before give the treatment. In pre-test, the students were given the topic and they were given 3 minutes to present their idea about the topic in front of the class. The topic of pre-

test is the effect of smoking. The students asked present the effect of smoking.

b. Post-test

Post-test was conducted at 11th March 2017. Post-test was given after the students were taught by using Think-Pair-Share strategy to know the students' speaking ability after give the treatment. In post-test, the students were given the topic, they were given 3 minutes to present their idea about the topic in front of the class. The topic of pre-test is the effect of using media social.

G. Data Analysis

Data analysis is a technique to analyze data to know the result of a research. In analyzing data, the researcher used quantitative data by using statistical program SPSS 16.0. The quantitative data analysis was used to know the significant differences on the students' speaking ability after teach by using think pair share strategy.

Data that was obtained from the pre-test and post-test would be analyzed statistically using *Paired-Sample T Test* through SPSS 16.0. The steps of analyzing data are as follow:

1. The researcher opened the program SPSS 16.0
2. The researcher computed the students' speaking score of pre-test and post-test and analyzed by click *Analyze > Compare Means > Paired Samples T-Test*

3. The researcher choose option to decide confidence interval percentage 95%
4. After that click *OK* to get the result.