CHAPTER III

RESEARCH METHOD

This chapter presents the research method. It focus on the method used in conducting this study. The decision includes Research Design,Population and Sample, Research Instrument, Validity and Reliability Testing, Data Collecting Method, Data Analysis

A. Research Design

Research design is a strategy or way to arrange the setting of the research in order to get the valid data. In this research, the researcher used experimental research. The researcher manipulated the independent variables by setting up a treatment to be applied into theexperimental group. The independent variable was using think pair share technique, while the dependent variablewas the students' grammar achievement in past continuos tense. Therefore, the treatment applied in this research was using think pair share technique in teaching grammar focused on simple present tense.

According to Gay *et al*(2011:318), there are two major classes of experimental designs:single-variable designs, which involve one independent variable, and factorial designs, which involve two or more independent variables. This research involved only one indepedent variable, therefore it was included into single variable design. Single-variable designs it self are classified into pre experimental, true-experimentl, and quasi-experimental designs.

The type of experimental design used in this research was pre-experimental design classified into one-goup pretest-posttest design. Therefore, in the one-group pretest-posttest design, a single group was observed not only after being given by treatment, but also before being given treatment. The experimental group would be conducted by using pretest before treatment and posttest for the result of treatment as instrument to collecting data.

Table the design of One-Group-Pretest-Posttest Design

Pre-test	Independent variable	Posttest
Y1	X	Y2

X : Think Pair Share Technique

Y_1 : students' grammar achievement before being taught by using think pair share technique

Y_2 : students' grammar achievement after being taught by using think pair share technique

The procedures of the pre-experimental research with one-group pretest-posttest design in this reearch were described as follows:

1. Administering a pretest (Y_1) which purposed to measure students' grammar achievement before given a treatment.

2. Applying an experimental treatment that was using think pair share technique (X) to teach grammar focused on past continuous tense.
3. Administering a pretest (Y_2) which purposed to measure students' grammar achievement after given a treatment.

There were two kind of variables in this research, independent variable and dependent variable.

1. Independent variable is a variable which infuences another variable. In this research, the independent variable was think pair share technique.
2. Dependent variable is a variable which is influenced by another variable. In this research, the dependent variable was the students' grammar achievement in past continuous tense.

In this research, the researcher used experimental research with quantitative approach. The researcher wanted to know the effectivness of using thinkpair share technique toward students' grammar achievement in past continuous tense at the second grade of MTs Wahid Hasyim Setinggil Gandekan Blitar. The effectiveness was known after finding out the significant difference between the students grammar score before and after being taught by using think pair share technique. The significant difference was found out by comparing the pretest and posttest scores.

This research is intended to investigate the effectiveness of using think-pair-share strategy to improve students grammar at MTs Wahid Hasyim

Setinggil Blitar in academic year of 2014/2015. By applying the treatment, it is expected to know whether the scores are increasing or not. So, the effectiveness of think-pair-strategy (TPS) strategy knows by the writer after getting the score both pretest and posttest.

B. Population and sample

1. Population

According to Gay (1992: 124) population is the group of interest to the researcher, the group which she or he would like the result of the study to generalizable. In encyclopedia of educational evaluation, written "a population is a set (or collection) of all elements processing one or more attributes of interest". Based on explanation above a population are the entire subjects who have some interest for researcher.

The population of this research are all students of the second grade of MTs Wahid Hasyim setinggil Blitar in academic year 2014/2015. In which the total are 57 students and divided into four classes. Every class of second grade consisted about 28.

2. Sample

Sample is representative population for research. According to Burhan Bungi (2008:101) taking sample it means to represent the whole population. Selection of sample is very important steps

in conducting a research study. The goodness of the sample determines the generalizability of the result. According to Gay (1992: 126) a good sample is one that is representative of the population from which it was selected. From those explanations, the sample selected was second grade of class B that consists of 28 students at MTs Wahid Hasyim Setinggil Blitar in academic year 2014/2015 which consist of 14 male and 14 famale. Technique to take sample is called sampling, in this study the writer used purposive sampling. This school was chosen purposively because to apply the experimental research, the samples must not be too "good" and too "bad" in their English achievement, especially in grammar . It is intended to reduce the extraneous variable may appears since the design is pre-experimental research without control group, while in MTs Wahid Hasyim Setinggil Blitar the classes divided into two class that is class A and class B and the ability of the students in two class is same but the teacher chose the class VIIIB. So, the writer decided VIIIB as recommendation by the teacher who handle English lesson in MTs Wahid Hasyim Setinggil Blitar and hopes VIIIB is the most representative ones. Based on Ashley (2014) states a purposive sample, also commonly called judgmental sample, is one that selected based on the knowledge of a population and the purpose of the study.

C. Data Collecting method

Method is the technique or the way that used by the researcher to collect the data (Riduan, 2006: 97). Moreover, Tanzeh (2011:57) states that data collection method is a systematical and standard procedure used to collect data by doing some activities as interview, observation, documentation, questioner and test. In this research, the data collecting method is administering test that consists of pre-test and post-test. The procedure of administering test was clarified as follows:

1. Pre-test, at the first meeting, the writer gave a pre-test to the students. There 10 question; 10 questions are in the form of short answer. It was conducted to know the scores of the students grammar before being taught the treatment. The pre-test was administered on 7 Mei 2015.
2. Post-test, the post-test is given to the students after conducting the treatment of using think-pair-strategy to improve the students' grammar. Similar to pre-test, the writer ask students to answer ten question in the form of short answer. The post-test was administered on 28 Mei 2015.

D. Research Instrument

Instrument is tool of collecting data. According to Arikunto (2006: 126) instrument is the device of the researcher to use in collecting data. The instrument in this research is a test. Test is a means of measuring knowledge, skill, feeling, intelligences, or aptitude of an individual group (Gay, 1992: 154). On the other case Gay L.R explain that test is a means

of measuring the knowledge, skill feeling, intelligence, or aptitude of an individual or group. Test procedure numerical score that can be used to identify, classify, or evaluate test takers.

To obtain the required score of the studens' grammar achievement the researcher used two kind of test. There are pretest and posttest, the pretest was done bofore the treatment process by using think-pair-share technique. This is subjective test given to know the basic competence for all students before they get treatment. The second was posttest. The posttest was done after treatment process. It is done to know the final score and to know the students' difference competence before and after get treatment.The test of this study is the writer made by adapted the material module book at MTs Wahid Hasyim Setinggil Blitar.

The tests were in the form of objective test that are short answer. The writer requires 10 questions of short answer. To score the objective tests the writer treats them without any difference. Means, there was only one correct answer for each items. The scoring guide is as the formula follow:

Score = number of correct items x 100

10

The test form of short answer was not giving choice so the learner can not only guess the answer, besides they did not confuse to answer like learners do in form of multiple choices test, the learners just give the fix answer, based on their the text.

According to Alderson (2000: 227) short answer is a test that the test takers are simply asked question which requires a brief response, in a few words, as in the example below (not just yes/no or true /false). The justification for this technique is that it is possible to interpret students' response to see if they have really understood, whereas on multiple choice items students give no justification for the answer they have selected and may they have chosen by eliminating others.

The writer used achievement test, Isnawati (2012: 14) states in contrast to proficiency test, achievement test are directly related to language course. The purpose of this kind of test is to establish how successful individual students, group of students, or the courses themselves have been in achieving objectives. Its means that, achievement test that is used to measure the process that students making after learn something in achieving objectives. This test used to measure the students achievement in grammar before and after they taught by using Think-Pair-Share technique in MTs Wahid Hasyim Setinggil Blitar.

E. Try Out

The researcher implemented a try out test in MTs Wahid Hasyim Setinggil Gandekan Blitar in class A. In this try out, the form of test was past continuous tense. The questions of the try out test are about 20. The form of try out test was made in short answer consist of 20 items. The total score 100, so the score for each number is 5.

There are two points of item analysis to try out test: item facility and discrimination power.

1. Measuring the Item Facility

To measure the item facility of level of difficulty of the test items, the researcher used the following formulas:

$$P \frac{B}{JS} (Arikunto, 2012: 223)$$

Where:

P = Item Facility (Level of difficulty)

B = Number of test-takers answering the item correctly

JS = number of test-takers responding to that item

To know the classification of the difficulty level, the researcher used the classification referred by Arikunto (2012:225). Here is the following classification and interpretation of difficulty level:

Table 3.1 Classification of Difficulty Indices

Difficulty Level	**Classification**
0.00-0.30	Difficult
0.31-0.70	Fair
0.71-1.00	Easy

Based on the classification and interpretation of difficulty level proposed by Arikunto, here is the result of difficulty level analysis of the test items:

Table 3.2 The Presentation of Level of Difficulty of Try Out Test

Test Item	**B**	**JS**	**IF = B/JS**	**Classification**
1	19	25	0.76	Easy
2	18	25	0.72	Easy
3	17	25	0.68	Fair
4	16	25	0.64	Fair
5	17	25	0.68	Fair
6	16	25	0.64	Fair
7	18	25	0.72	Easy
8	16	25	0.64	Fair
9	15	25	0.6	Fair
10	3	25	0.12	Difficult
11	14	25	0.56	Fair
12	14	25	0.56	Fair
13	12	25	0.48	Fair
14	9	25	0.36	Fair
15	13	25	0.52	Fair
16	12	25	0.48	Fair
17	21	25	0.84	Easy
18	24	25	0.96	Easy
19	17	25	0.68	Fair
20	17	25	0.68	Fair

3. Measuring Discrimination Power

In order to measure the discrimination power of each item, the researcher needed to separate the students into upper and lower group in order to be applied in the following formula:

$$DP, \frac{B_A}{J_A} = \frac{B_B}{J_B}, P_{A=}P_B$$

(Arikunto, 2012:228)

Where:

DP = Discrimination Power

J =Number of Test-takers

J_A =Total participant of top test-takers

J_B =Total participant of bottom test-takers

B_A = Number of top test takers that have correct answer

B_B =Number of bottom test takers that have correct answer

P_A , $\frac{B_A}{J_A}$ = Proportion of the number of top class answering correctly

$P_{B,}$ $\frac{B_B}{J_B}$ = Proportion of bottom class answering correctly

According to Arikunto (2012:232), here is the classification and interpretation of discrimination index:

Table 3.3 Classification and Interpretation of Discrimination Indices

Discrimination Index	**Classification**
0.71-1.00	Excellent
0.41-0.70	Good
0.21-0.40	Satisfactory
< 0.20	Poor
Negative value on D	Very Poor

Table 3.4 the Data Presentation of Discrimination Power of Try Out Test

Item	BA	BB	JA	JB	PA	PB	D=PA-PB	Classification
1	7	5	8	8	0.87	0.62	0.25	Satisfactory
2	8	4	8	8	1	0.5	0.5	Good
3	6	4	8	8	0.75	0.5	0.25	Satisfactory
4	4	5	8	8	0.5	0.62	-0.12	Very poor
5	7	3	8	8	0.87	0.37	0.5	Good
6	5	5	8	8	0.62	0.62	0	Poor
7	8	3	8	8	1	0.37	0.63	Good
8	6	4	8	8	0.75	0.5	0.25	Satisfactory
9	6	4	8	8	0.75	0.5	0.25	Satisfactory
10	2	0	8	8	0.25	0	0.25	Satisfactory
11	5	3	8	8	0.62	0.37	0.25	Satisfactory
12	6	2	8	8	0.75	0.25	0.5	Good
13	6	2	8	8	0.75	0.25	0.5	Good
14	4	2	8	8	0.5	0.25	0.25	Good
15	5	1	8	8	0.62	0.12	0.5	Good
16	4	3	8	8	0.5	0.37	0.13	Poor
17	8	4	8	8	1	0.5	0.5	Good
18	8	7	8	8	1	0.87	0.13	Poor
19	6	3	8	8	0.75	0.37	0.38	Satisfactory
20	5	6	8	8	0.62	0.75	-0.13	Very poor

The result of discrimination power analysis was shown in table 3.2 showing that test items are still acceptable the criteria of good or satisfactory.

Sudjiono (1996: 408) states that following up after analyzing the discrimination power of a certain test must be done by the teacher or test-maker in order to revise the test items. The follow up proposed by Sudjiono are as follows:

a. The items which have good discrimination power; satisfactory and excellent classification; should be kept in item test bank, so that it can be used later.
b. The items which are categorized into the poor distractor should be revised and then used later.

c. The very poor discriminator of the test items then must be dropped or removed because it cannot be used later.

F. Validity and Reliability Testing

As mentioned above, the writer's instrument is test. The good instrument should be valid and reliable, the more explanation about it shows as follows:

1) Validity

The most simplistic definition of validity is that is the degree to which a test measures what is supposed to measure (Gay, 1992: 155). To measure whether the test has good validity, the researcher analyzed the test from content validity, face validity and construct validity.

a. Content Validity

According to Brown (2004) in Isnawati (2012:27), a test is said to have content validity if it contents constitutes a representative sample of the language skills, structures, etc.being tested. The test will also have content validity if it includes a proper sample of the structure or content which is relevant with the purpose of the test. Content validity is sometimes called curriculum validity. The test will be valid if the objectives of the test do not outside from the curriculum objectives that have been set by educational policy.

In this case, the researcher learn the curriculum setting to know what students had to be able to master at the second semester of the second grade of islamic junior high school. The researcher found that the students should be able to master one genre that is recount text. That text needed past continuous tense to be mastered. It was used by the researcher to select grammar in the form of past continuous tense. Next the form of the test were simple completion with inflection form based past continuous tense. It fulfills the content of grammar test, therefore the test are valid of content validity.

b. Face Validity

A test is said to have validity if it looks as if it measure what is supposed to measure. Face validity is hardly a scientific concept, yet it is very important (Isnawati, 2012: 29). In this study, the item of the tests was in the form of objective tests consists of short answer test. The writer ensured face validity by consulting to English teacher of MTs Wahid Hasyim Setinggil Blitar.

c. Construct Validity

A test is said to have construct validity if it can be demonstrated that it measures just the ability which is supposed to measure (Isnawati, 2012:29). Construct validity is one kind of validity that is measures the ability which is supposed to

measure. The word “construct“ refers to any underlying ability which is hypothesized in a teory of language ability. As mention by Brown in Isnawati (2012:30) that construct is any theory, hypothesis or model that attempts to explain observed phenomena in our universe of perception.

By basing on the theories above, in this research, the researcher created the tests based on the material which was suitable for the students at the second grade of islamic junior high school. Next, the researcher asked the students to answer questions in the form of past continuous tense to measure the students’ grammar achievement. It fulfills the construct of grammar test, therefore the tests are valid in term of construct validity.

2) Reliability

The next way to know a good test is by reliability. A Reliable test is consistent and dependable. If the students are given the same test on two different occasions, the test should produce similar results. The word “similar” is used here because it is almost impossible for the test-takers to get exactly the same scores when the test is repeated the following day. This is because of the fact that human beings do not simply behave in exactly the same way on every occasion, even when the circumstances seem identical (Isnawati,

2012:18). In this case, for reliability the writer took test from teachers test so that is assumed that the test is reliable.

Table 3.4 The preparatory to compute the standard deviation

No.	**Subject**	$\mathbf{X_t}$	$\mathbf{X_t^2}$
1	AA	17	289
2	AS	8	64
3	ADO	12	144
4	BAZ	15	225
5	BS	16	256
6	BNS	18	324
7	BS	7	49
8	DS	17	289
9	EW	17	289
10	FS	16	256
11	FN	17	289
12	IL	18	324
13	IM	14	196
14	LPK	16	256
15	MA	16	256
16	MAA	16	256
17	MFZ	13	169
18	MIR	13	169
19	MIA	14	196
20	MRA	16	256
21	MWA	14	196
22	MYE	5	25
23	NSR	14	196
24	ND	12	144
25	NA	11	121
26	UN	14	196
27	UNN	12	144
28	ZN	11	121
		389	5695

$S_t^2 = \frac{\sum xt2}{N}$

To know $\sum x_t^2$ the formula below was used:

$$\sum x_t^2 = \sum x_t^2 - \left(\frac{\sum xt}{N}\right)^2$$

$$= 5695 - \left(\frac{389}{28}\right)^2$$

$$= 5695 - 192.9$$

$$= 5502,1$$

Therefore, the standart deviation is

$$\sqrt{S_t}^2 = \frac{\sqrt{5502.1}}{28} = 14$$

After finding the result of standart deviation, the reliability can computed by using Kuder Richardson formula (KR.20)

Table 3.5 to compute the reliability by using kuder richardson formula(KR 20)

Item	**NP**	**P_1** **(NP/∑soal)**	**Nq**	**Q_1** **(NP/∑soal)**	**P_1Q_1**
1	17	0.85	3	0.15	0.1275
2	8	0.4	12	0.6	0.24
3	12	0.6	8	0.4	0.24
4	15	0.75	5	0.25	0.1875
5	16	0.8	4	0.2	0.16
6	18	0.9	2	0.1	0.09
7	7	0.35	13	0.65	0.2275
8	17	0.85	3	0.15	0.1275
9	17	0.85	3	0.15	0.1275
10	16	0.8	4	0.2	0.16
11	17	0.85	3	0.15	0.1275
12	18	0.9	2	0.1	0.09
13	14	0.7	6	0.3	0.21
14	16	0.8	4	0.2	0.16
15	16	0.8	4	0.2	0.16
16	16	0.8	4	0.2	0.16
17	13	0.65	7	0.35	0.2275
18	13	0.65	7	0.35	0.2275

19	14	0.7	6	0.3	0.21
20	16	0.8	4	0.2	0.16
21	14	0.7	6	0.3	0.21
22	5	0.25	15	0.75	0.1875
23	14	0.7	6	0.3	0.21
24	12	0.6	8	0.4	0.24
25	11	0.55	9	0.45	0.2475
26	14	0.7	6	0.3	0.21
27	12	0.6	8	0.4	0.24
28	11	0.55	9	0.45	0.2475
					$\sum P_1Q_1 =$ 5.2125

Therefore, the reliability is :

$$r_{11} = \left[\frac{n}{n-1}\right]\left[\frac{st2 - \sum p1q1}{st2}\right]$$

$$r_{11} = \left[\frac{28}{28-1}\right]\left[\frac{14-5.2125}{14}\right]$$

$$r_{11} = \left[\frac{28}{27}\right]\left[\frac{14-5.2125}{14}\right]$$

r_{11} = (1.03) (0.63)

r_{11} = 0.65

The result shows that the test was reliable with the reliability coefficient of 0.65 or 65%, it means that the reliability of test is high.

G. Data Analysis

In quantitative research to analyze the data is by using statistical data analysis. The data should in the form of number.

Data analysis is a technique to analyze data in order to know the result of a research. In this study, the writer provides some formula based on Arikunto (2010:349) as follows:

$$t = \frac{md}{\sqrt{\frac{\sum x^2}{N(N-1)}}}$$

Where:

Md : means of different pre-test and post-test

Xd : deviation in every subject (d – Md)

Σx2d : total of quadrate deviation

N : subject of sample