CHAPTER III

RESEARCH METHOD

This chapter describes the research design, Population and Sample, Research Instrument, Validity Testing, Reliability Testing, Data Collecting Method and Data Analysis.

A. Research Design

This study used pre-experimental research design with quantitative approach because the researcher intended to investigate the effect of using word search puzzle game for improving students' vocabulary mastery of the second grade at MTsN Tulungagung.

According to Ary et al (2010: 265) experimental research is a scientific investigation in which the researcher manipulates one or more independent variables, controls any relevant variable, and observes the effect of the manipulations on the dependent variables. The present research conducted an experimental teaching using a certain game, word search puzzle game. It had two variables, the teaching game as the independent variable, meanwhile the student's vocabulary mastery as the dependent variable.

There are two designs included in pre-experimental. They are one-group pretest-posttest design and static group. The researcher used one group pretest-posttest design as experimental group. In pretest and posttest group the observation do two times, before giving treatment called pretest and after giving treatment called posttest. The one group pretest-posttest design usually involves

three steps: (1) administering a pretest measuring the dependent variable, (2) applying the experimental treatment x to the subject, and (3) administering a posttest, again measuring the dependent variable. (Ary, 2010: 327) That is based on the diagram below:

Table 3.1 Diagram of one group pretest and posttest design

Pretest	Independent variable	Posttest
Y1	X	Y2

It means that there was one class only in this research. The students had two tests: before being taught by using word search puzzle game and after being taught by using word search puzzle game. Then both of score is compared to know the significant different.

B. Population and Sample

Population is the groups want to generalize your findings. Population is very important to determine who in most cases going to sample (Muijs, 2014: 37). Based on that statement above, the population of this research was all of the second grade students of MTsN Tulungagung, which consist of 9 classes from A class until I class. That is list of the students total on the table below:

Table 3.2 Diagram of list of the students' total

Class	Total
A	36 students
B	36 students
C	35 students
D	35 students
E	39 students
F	40 students
G	40 students
H	39 students
I	39 students

Sampling whole population as a census (Muijs, 2014: 38). It is also possible to sample even a large population given enough resources. This sampling had function to get information about population. So that sampling was the technique to take a sample. In this research, random sampling was chosen as a technique of choosing sample, because the random sampling was the best single way to obtain a representative sample. One class was finally selected as the sample of the study, which is B class. Besides, the researcher also admitted that subjects were homogeny in their skills especially in their vocabulary.

C. Research Instrument

In this study, the instrument used to get and collect information will be in form of test. In this research, test was used as a main instrument. Because the writer uses pre experimental design, there will be pre test and post test here. Pre test and post test will be given to the all of second grade students before get a treatment and after get a treatment. The form of test was vocabulary test. The questions of the test are about 20. The form of test was made in the following form:

1. Fill in the blanks of 15 items. The students must choose the correct answer in the box to complete the paragraph and sentence. The total score is 60, so the score for each number is 4.

2. Arrange the words that consist of 5 items. The student must arrange the scramble word. The score for all items is 8, so the total scores 40.

D. Try Out Test

The researcher implemented a try out test on May 04, 2015 in MTs Negeri Tulungagung in class A. In this try out, the form of test was vocabulary test. The questions of the try out test are about 35. The form of try out test was made in the following form:

1. Fill in the blank of 25 items. The total score 50, so the score for each number is 2.

2. Arrange the words that consist of 10 items. The score for all items is 5, so the total scores 50.

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:

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

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 3.3 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 3.4 The Presentation of Level of Difficulty of Try Out 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
21	14	25	0.56	Fair
22	3	25	0.12	Difficult
23	16	25	0.64	Fair
24	23	25	0.92	Easy
25	21	25	0.84	Easy
26	25	25	1	Easy
27	16	25	0.64	Fair
28	2	25	0.08	Difficult
29	24	25	0.96	Easy
30	22	25	0.88	Easy
31	25	25	1	Easy
32	25	25	1	Easy
33	20	25	0.8	Easy
34	23	25	0.92	Easy
35	25	25	1	Easy

2. 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 3.5 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 3.6 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
21	6	3	8	8	0.75	0.37	0.38	Satisfactory
22	3	0	8	8	0.37	0	0.37	Satisfactory
23	5	7	8	8	0.62	0.87	-0.25	Very poor
24	8	6	8	8	1	0.75	0.25	Satisfactory
25	8	5	8	8	1	0.62	0.38	Satisfactory
26	8	8	8	8	1	1	0	Poor
27	6	4	8	8	0.75	0.5	0.25	Satisfactory
28	0	1	8	8	0	0.12	-0.12	Very poor
29	8	7	8	8	1	0.87	0.13	Poor
30	8	5	8	8	1	0.62	0.38	Satisfactory
31	8	8	8	8	1	1	0	Poor
32	8	8	8	8	1	1	0	Poor
33	7	5	8	8	0.87	0.62	0.25	Satisfactory
34	8	7	8	8	1	0.87	0.13	Poor
35	8	8	8	8	1	1	0	Poor

The result of discrimination power analysis was shown in table 3.4 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.

E. Validity Testing

The validity of an assessment measure or tool indicates the extent to which it is an adequate measure of the curriculum and objectives it represent. Validity is the most important consideration in developing and evaluating measuring instruments. The focus of recent views of validity is not on the instrument itself but on the interpretation and meaning of the scores derived from the instrument (Ary et al, 2010: 225).

There are several types of validity, but the most important type for classroom teachers to understand was content validity. Content validity is not separated from construct validity. In fact, construct validity is the prerequisite for content validity (Latief, 2012: 227). It means that content validity attained when there was a close matching between the content of the assessment and of the

curriculum and instruction. When students are assessed on the material that they had been studying in the classroom, the assessment measure can be had content validity. In this test, the researcher asked students to answer the matching test and arranging the scrabble vocabulary to measure students' mastery in learning vocabulary. Therefore, the materials used in this research were taken from English syllabus in second grade at Junior High School. This instrument of this research will have a content validity since the test is designed based on SKKD in KTSP 2006. This material just covered the "Narrative" that was described below.

Table 3.7 the Materials for Junior High School

No	Topic	Objective	Learning activity	Vocabulary
1	Narrative Text	Students are able to respond the meaning in short simple monologue accurately, fluently and acceptably to interact with the environment in the form of narrative	Introducing vocabulary related with topic (Narrative Text)	The story of Malin Kundang

The material Junior High School in the book students (LKS). Besides, the instrument used had face validity too. Face validity refers to the degree to which a test looks right, and appears to measure the knowledge or abilities it claims to measure, based on subjective judgment or the examines who take it, the administrative personnel who decide on its use, and other psychometrically unsophisticated observer (Mousavi, cited in Brown, 2004). In the test, the researcher asked the students to answer the test with topic narrative text to measure the students' mastery in vocabulary.

F. Reliability Testing

Reliability of a measuring instrument is the degree of consistency with which it measures whatever it is measuring. This quality was essential in any kind of measurement. The teacher use such measuring instruments must identify and use techniques that will help the teacher determine to what extent their measuring instruments are consistent and reliable (Ary et al 2010: 236). To make sure that the instrument (test) was reliable, the researcher set test vocabulary prompt and scoring rubric it advances so that students can use it as their borders to produce a final version of their vocabulary.

In this study, the internal consistency of reliability of estimated through Kuder-Richardson Reliability Coefficient. The result shows that the test was reliable with the reliability coefficient of 0.78 or 78%, it means that the reliability of test is high. Can be seen in the Appendix.

G. Data Collecting Method

Data collecting method is the way that is used to get the data. (Muijs, 2004: 41) state that data collecting is the next phase and another one where problems can occur in survey studies. The aim of the data collecting in conducting a scientific research was to get the materials needed. The materials must relate to each other and to solve the problem. There was method of data collection used in this research. It was administering test.

Administering test is the data collecting method of test. Before testing it, the researcher would do the pretest to know ability's student before giving treatment. After giving pretest, the researcher gave treatment to drill ability's

students about vocabulary related with the topic. After giving treatment, the researcher gave posttest. Posttest was used to know ability's students after getting treatment. Then, the result of test would be compared between pretest and posttest score weather differences or not. If there any differences score, it showed that treatment was successful and if there was no differences score, it showed that treatment was unsuccessful.

In the teaching procedure of word search puzzle game, there were steps. The following section was described each stage. Steps in word search puzzle game will be presented below.

Table 3.8 the Teaching Procedure

No	Steps	Teacher activities	Learners activities
1	Opening	Greeting	✦ Answering greeting ✦ Brain storming
2	Main teaching	Introducing new vocabulary related with topic	✦ Pay attention
		Teacher introduce the word search puzzle game and explained instruction to do it correctly	✦ Pay attention
		Teacher gave the handout copies to be tried by students to review those new vocabulary through word search puzzle game	✦ All students tried to do it carefully to find, arrange horizontally and vertically in word search puzzle game
3	Closing	Pointing and asking the students randomly to memorize those new vocabulary related with topic	✦ Answering the teacher and memorizing new vocabulary
		Giving evaluation	✦ Express their difficulty in learning using word search puzzle game

The one group pretest and posttest was exposed to the treatment in the same time with the same topics also. The schedule of the treatment was presented below.

Table 3.9 the Schedule of Test and Treatment

Treatment and test	One group pretest and posttest
Subject (N)	36 subjects
Topic : Narrative	**Pretest**
	May 9, 2015
	Treatment
	May 14, 2015 May 16, 2015 May 18, 2015
	Posttest
	May 21, 2015

H. Data Analysis

In managing and analysis quantitative data collected from the research, the researcher used quantitative data analysis by using statistical program. The quantitative data analysis was used to know the student's achievement in vocabulary after using word search puzzle game in teaching vocabulary. The researcher conducted test to the student before and after they were taught by using word search puzzle game. Here the researcher use t-test formula to analyze the data to know the students test result which are conducted before and after using word search puzzle game. The description of t-test is as follow:

$$t = \frac{M_1 - M_2}{\sqrt{\frac{N\left(\sum D^2\right) - \left(\sum D\right)^2}{N^2(N-1)}}}$$

M_1 = average score of post test

M_2 = average score of pre test

N = the number of pair

$\sum D$ = the sum of squared differences score