

CHAPTER III

RESEARCH METHOD

In this chapter, the researcher presented research method. It focused on the method used in conducting this study. This chapter presented some topics dealing with the research method. Those are: research design, data and data source, population, sample and sampling, data collecting method, validity and reliability testing, method of data analysis, and hypothesis testing.

A. Research Design

This research study used quantitative approach and pre-experimental design. The following definition, taken from Aliaga and Gunderson (2002), Muijs (2004: 1) describes that the meaning of quantitative research methods as follows:

Quantitative research is ‘Explaining phenomena by collecting numerical data that are analyzed using mathematically based methods (in particular statistics).’

This research was using quantitative approach because the purpose of this research was to test a theory, to test whether Blindfold Game was effective used or not toward students’ vocabulary mastery. This research was conducted without a control group.

Variable is defined as “*characteristics that tend to differ from individual to individual, though any two or more individuals may have the same variable trait or measure*”. (Charles, C.M. 1995:29, as cited by Latief, 2017: 11). This research had two variables, they were:

1. Independent variable in this research was the researcher uses Blindfold Game.
2. Dependent variable in this research was student's vocabulary mastery at the seventh grade of MTs Al Huda Bandung.

The researcher used the one-group pretest-posttest design. The pre-test was given before treatment. The pre-test contained multiple choice questions and matching words. After the pre-test, the researcher provided treatment. This treatment done by teaching lessons and Blindfold Games. After treatment, the researcher provided a post-test. The questions in the post-test also contained multiple choice questions and matching words.

Data is all information in the form of numbers and writings that will be collected to be processed and have function to answer research questions then draw final conclusions of the research. In this research, the quantitative data was in the form of numbers. So, the data of this study were students' vocabulary mastery scores before and after the treatment given. The data were collected in three sources as follows:

- Person: 7A students of MTs Al Huda Bandung in which the treatment will be given.
- Place: 7A classroom where the tests will be administered
- Paper: Students' vocabulary mastery tests

Latief (2017: 97) stated the effectiveness of the instructional treatment is measured by comparing the average score of the pre-test and the post test. When it turns out that the post test average score is significantly higher than the average score of the pre-test, then it is concluded that the instructional treatment is effective.

So, at the end of the research, the researcher had to conclude the effectiveness of Blindfold Game in teaching vocabulary in pre-experimental design by comparing pre-test and post test results. The effectiveness was known through differentiate scores between pre-test and post-test. If the post test scores were higher than pre-test scores, the technique would be effective and vice versa.

B. Population, Sample, and Sampling

The population is the group of people the researchers to *generalize* to (Muijs, 2004: 15). Target population in educational research is usually defined as *all the members of a real or hypothetical set of people, events, or objects to which educational researchers wish to generalize the results of the research* (Borg, W.R., Gall, M.D., 1989: 216, as cited by Latief, 2017: 195). The population of this research were all of the seventh grade students at MTs Al Huda Bandung, there were 4 classes (A-D).

Latief (2017: 195), as cited by Charles, C.M. (1995: 87), defines a sample as *a small group of people selected to represent the much larger entire population from which it drawn*. While the process of selecting just a small group of cases from out of large group is called sampling (Muijs, 2004: 93). This research took one class of four classes which used purposive sampling technique. The researcher interviewed with the teacher to ask her suggestions toward this research.

C. Validity and Reliability Testing

a. Validity

Latief (2017: 236) states valid means correct. For example, when the result of students' writing assessment is valid, then the students' writing assessment reflects the students' writing skills as distinct from their reading skills, or from their knowledge of vocabulary. Charles, C.M. (1995:101), as cited by Latief (2017: 236), states that *"Scores obtained from a test of critical thinking are valid if they represent ability to think critically, as distinct from knowledge of vocabulary or reading skill"*. The correctness of the assessment is called validity and the evidence to support the correctness of the assessment is called validity evidence.

According to Ary et al (2010: 225), validity was defined as the extent to which an instrument measured what it claimed to measure. Heaton (1988: 159) states that the validity of test is the extent to which it measures and nothing else. There are four types of validity, they are content validity, criterion-related validity, construct validity, and face validity. In this research, the researcher used content validity and construct validity to measure whether the test had a good validity or not. The description of both types of validity are described below.

- Content Validity

According to Latief (2017: 239), in the construct validity, the task of the students are required to perform becomes the evidence. If the task of the students are required to perform matches the purpose of the assessment, then the match becomes the construct validity evidence.

The absence of the match between the task and the purpose of the assessment leads to the construct validity problem to the result of assessment. In the content validity, the coverage of the tasks becomes the evidence. If the task of the students are required to perform covers all samples of the domain in the skill or in the knowledge to be assessed, the representativeness of the samples becomes the supporting content validity evidence. Gronlund (1985:59), as cited by Latief (2017: 239), states that *“The essence of content validation is determining the adequacy of sampling. More formally, content validation is the process of determining the extent to which a set of test tasks provides a relevant and representative sample of the domain of tasks under consideration”*. In developing assessment instruments, the test developer includes all samples of the content domain of the skill or knowledge to be assessed.

This research used test as the instrument. The content validity of the test was done by comparing the content of the test and the material that would be learned. The researcher made the pre-test and post test based on the core competence and basic competence. After that, the researcher made blueprint that contained basic competence, dimension, variable, and indicators. The core competence and basic competence could be seen as follows:

Table 3.1 The Core and Basic Competence

Core Competence	Basic Competence
3. Understanding knowledge (factual, conceptual, and	3.7 Understanding the purpose, structure, and linguistic elements

procedural) based on the curiosity about science, technology, art, culture related to phenomena and events that real.	of oral and written texts for the mention of the nature of people, animals and things.
---	--

The content validity of this research could be seen as follows:

Table 3.2 Test Specification on Pre-test

Indicator	Description of question	Kind of test	Number of test	Total of test
1. Identifying information on people and animals contained in the sentence	a. Determining the characteristics of people based on picture	Multiple choice	11-15	5
	b. Determining the name of characterized part of animal description based on picture	Multiple choice	16-20	5
2. Identifying information on things contained in the sentence	Determining the characteristics of things based on picture	Multiple choice	1-10	10

3. Find the meaning of information on people, animals and things contained in the table	Determining the meaning of words that related to information of people, animals, and things in the table	Matching	1-5	5
	Total			25

Table 3.3 Test Specification on Post test

Indicator	Description of question	Kind of test	Number of test	Total of test
4. Identifying information on people and animals contained in the sentence	c. Determining the characteristics of people based on picture	Multiple choice	11-15	5
	d. Determining the name of characterized part of animal description based on picture	Multiple choice	16-20	5
5. Identifying information on	Determining the characteristics of	Multiple choice	1-10	10

things contained in the sentence	things based on picture			
6. Find the meaning of information on people, animals and things contained in the table	Determining the meaning of words that related to information of people, animals, and things in the table	Matching	1-5	5
	Total			25

- Construct Validity

Latief (2017: 238) states construct validity evidence comes from the assessment instrument used. An assessment instrument is always designed to measure specific knowledge or skills of a group of people. So, the development of an assessment instrument must always be started from the defined construct of the knowledge or skills to be assessed. The correct definition of the construct will lead to the correct selection of the task, which will result in correct data, which has strong validity.

Therefore, the researcher created the test based on the material which was suitable to the students at the seventh grade of MTs Al Huda Bandung Tulungagung. Then, the researcher constructed questions of the test from the simple one to the complex one in order to measure their

vocabulary mastery. In the multiple choice, the students answered the questions from the selection answer. In the matching item, the students answered the questions by matching the words on the left side of the table and their right meanings on the right side of the table.

b. Reliability

Latief (2017: 226) stated while validity refers to the degree of correctness of the assessment result in representing skill being assessed, reliability of the result of language skill assessment refers to the preciseness of the language skill assessment result in representing the actual level of the skill of the examinees. The result of a language skill assessment has high reliability if the result precisely represents (is very closed to, or is not too far away from, or gives good estimate of, or does not overestimate or underestimate) the true level of the skill being assessed. In other words, if the language skill assessment result is too far away different from the true level of the skill being assessed, then the assessment result has low reliability. The distance between the true level of the skill and the assessment result, then, determines the degree of reliability; the bigger the distance is between the language skill assessment result and the actual level of the skill being assessed, the lower the reliability of that assessment result is.

According to Raharjo (2019) on his blog SPSSIndonesia.com, as cited by Widiyanto (2010: 43), stated that the basis for decision making in reliability testing is as follows:

1. If the value of Cronbach's Alpha $> r_{table}$, then the questionnaire is declared reliable.

2. If the value of Cronbach's Alpha is $< r_{table}$, then the questionnaire is declared unreliable.

In this research, the researcher used test as research instrument. Therefore, the statements could be seen as follows:

1. If the value of Cronbach's Alpha $> r_{table}$, then the test is declared reliable.
2. If the value of Cronbach's Alpha is $< r_{table}$, then the test is declared unreliable.

Table 3.4 Pre-test Reliability Statistics

Cronbach's Alpha	N of Items
.733	25

Based on the table 3.4, the SPSS 16.00 calculation showed that the number of questions in pre-test was 25 and the Cronbach's Alpha was 0.733. The value of df can be calculated with $N-2$ (the total students minus 2), so the df was nine ($11-2$). The value of r_{table} for one tailed with Sig. 0.05 was 0.5214.

Table 3.5 Item-Total Statistics (Pre-test)

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	r_{table}
<i>Soal1</i>	20.00	7.800	.426	.710	0.5214

<i>Soal2</i>	20.09	8.691	.125	.740	0.5214
<i>Soal3</i>	19.45	8.873	.162	.731	0.5214
<i>Soal4</i>	19.55	8.073	.443	.710	0.5214
<i>Soal5</i>	19.55	8.873	.091	.740	0.5214
<i>Soal6</i>	19.36	9.255	.000	.734	0.5214
<i>Soal7</i>	19.73	8.218	.270	.727	0.5214
<i>Soal8</i>	19.55	9.673	-.231	.765	0.5214
<i>Soal9</i>	19.36	9.255	.000	.734	0.5214
<i>Soal10</i>	19.36	9.255	.000	.734	0.5214
<i>Soal11</i>	19.36	9.255	.000	.734	0.5214
<i>Soal12</i>	19.36	9.255	.000	.734	0.5214
<i>Soal13</i>	19.36	9.255	.000	.734	0.5214
<i>Soal14</i>	19.36	9.255	.000	.734	0.5214
<i>Soal15</i>	19.36	9.255	.000	.734	0.5214
<i>Soal16</i>	19.36	9.255	.000	.734	0.5214
<i>Soal17</i>	19.36	9.255	.000	.734	0.5214
<i>Soal18</i>	19.45	8.273	.514	.709	0.5214
<i>Soal19</i>	19.55	8.473	.263	.726	0.5214
<i>Soal20</i>	19.45	8.273	.514	.709	0.5214
<i>Soal21</i>	19.64	7.455	.620	.690	0.5214
<i>Soal22</i>	19.82	6.964	.732	.673	0.5214
<i>Soal23</i>	20.09	7.291	.692	.682	0.5214
<i>Soal24</i>	19.55	8.473	.263	.726	0.5214

<i>Soal25</i>	19.64	7.655	.535	.699	0.5214
---------------	-------	-------	------	------	--------

Based on the table 3.5, it showed that all questions of pre-test were reliable because the value of Cronbach's Alpha was bigger than r_{table} ($0.733 > 0.5214$).

Table 3.6 Posttest Reliability Statistics

Cronbach's Alpha	N of Items
.864	25

Based on the table 3.6, the SPSS 16.00 calculation showed that the number of questions in posttest was 25 and the Cronbach's Alpha was 0.864. The value of df could be calculated with $N-2$ (the total students minus 2), so the df was nine ($11-2$). The value of r_{table} for one tailed with Sig. 0.05 was 0.5214.

Table 3.7 Item-Total Statistics (Posttest)

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	r_{table}
<i>Soal1</i>	21.82	9.764	.921	.838	0.5214
<i>Soal2</i>	21.82	9.764	.921	.838	0.5214
<i>Soal3</i>	21.64	12.255	.000	.865	0.5214
<i>Soal4</i>	21.91	9.491	.884	.838	0.5214
<i>Soal5</i>	21.64	12.255	.000	.865	0.5214

<i>Soal6</i>	21.64	12.255	.000	.865	0.5214
<i>Soal7</i>	21.73	11.018	.572	.854	0.5214
<i>Soal8</i>	21.82	12.364	-.096	.881	0.5214
<i>Soal9</i>	21.73	11.018	.572	.854	0.5214
<i>Soal10</i>	21.64	12.255	.000	.865	0.5214
<i>Soal11</i>	21.64	12.255	.000	.865	0.5214
<i>Soal12</i>	21.64	12.255	.000	.865	0.5214
<i>Soal13</i>	21.64	12.255	.000	.865	0.5214
<i>Soal14</i>	21.64	12.255	.000	.865	0.5214
<i>Soal15</i>	21.64	12.255	.000	.865	0.5214
<i>Soal16</i>	21.64	12.255	.000	.865	0.5214
<i>Soal17</i>	21.64	12.255	.000	.865	0.5214
<i>Soal18</i>	21.73	11.018	.572	.854	0.5214
<i>Soal19</i>	21.91	10.091	.656	.850	0.5214
<i>Soal20</i>	21.82	9.764	.921	.838	0.5214
<i>Soal21</i>	21.82	10.764	.500	.857	0.5214
<i>Soal22</i>	21.73	11.018	.572	.854	0.5214
<i>Soal23</i>	21.82	11.164	.343	.863	0.5214
<i>Soal24</i>	21.73	12.418	-.120	.875	0.5214
<i>Soal25</i>	21.91	9.491	.884	.838	0.5214

Based on the table 3.7, it showed that all questions of posttest were reliable because the value of Cronbach's Alpha was bigger than r_{table} ($0.864 > 0.5214$).

D. Data Collecting Method

Data collection method is a systematical and standard procedure used to collect the data. The researcher collects the data by administering test. Wikipedia.com (2016) defines a test or examination (informally, exam or evaluation) as an assessment intended to measure a test-taker's knowledge, skill, aptitude, physical fitness, or classification in many other topics. A test may be administered verbally, on paper, on a computer, or in a predetermined area that requires a test taker to demonstrate or perform a set of skills. Tests vary in style, rigor and requirements.

The researcher used test as the research instrument. The researcher used two kind of test, they were pre-test and post-test. The kind of test was multiple choice-test and matching words. Both pre-test and posttest consist questions that related to vocabulary. The total question was 25 questions in the form of multiple choice (20 questions) and matching item (5 questions). The topic in the pre-test and posttest were about identifying information or the characteristics of people, animals, and things around the students' daily life.

The data was collected through this research by administering pre-test. Pre-test was conducted before the students were given the treatment. The researcher needed to know their vocabulary mastery before doing Blindfold Game. Before the researcher administered the pre-test, the researcher have tried out the test on seventh grade students in Madrasah Diniyah Al-Huda Kedungwaru Tulungagung in order to know the validity and reliability of the test. The schedule of researcher as follows:

Table 3.8 The Schedule of Research

No	Date	Activity
1.	March 21 st 2019	Administering pre-test
2.	March 26 th 2019	Giving treatment
3.	March 28 th 2019	Giving treatment
4.	April 2 nd 2019	Administering posttest

The first step in collecting research data was administering pre-test. The researcher gave the pre-test on March 21st 2019. Some materials in pre-test has been taught by the teacher before the researcher gave the pre-test.

After administering the pre-test, the second step was giving treatment. The researcher gave the treatment into two meetings. The first meeting was conducted on March 26th 2019. On the first meeting, the researcher gave an English lesson contained materials about present tense and pronoun. The researcher also helped the students to know the meaning of vocabularies they didn't understand in pre-test in order to increase their vocabulary mastery. Beside that, the researcher also taught about vocabularies used in describing things, animals, and person. In describing things, the vocabularies divided into some groups, such as size, color, shape, material, number, and their function. In the end of the lesson, the researcher told the students that there would be a game to play on the next meeting. Therefore, the researcher divided the students into three small groups. The groups were group A (4 students), group B (4 students), and group C (3 students). Then, the researcher asked each group to bring one thing around their daily lives.

The second meeting of the treatment was conducted on March 28th 2019. In this meeting, the researcher asked the students to play a game. The game was Blindfold Game. The game was used for increasing students' vocabulary mastery. The objective of this game was guessing things by using blindfold. Every group have brought the thing and had to keep it from others. The game started with group A which tried to guess group C's thing using blindfold. Each member of group could guess the thing at least from its shape, material, size, weight (heavy or light), and other characteristic point from things. The next guessing was done by the other groups. Group B tried to guess group A's thing and group C tried to guess group B's thing.

The last step was administering posttest. It conducted on April 2nd 2019. The researcher gave posttest to the students after the treatments were done. The researcher needed to know students' vocabulary mastery after doing Blindfold Game. If the post test scores was higher than pre-test score, then the Blindfold Game was effective toward students' vocabulary mastery.

The researcher wanted to measure the effectiveness of Blindfold Game toward students' vocabulary mastery, and to know whether the game was effective or not. The result of the study can be decided after comparing scores of pre-test and post test. The material of the test was from English book and other resources to add more vocabularies.

E. Method of Data Analysis

Method of data analysis is the way data analyzed by the researcher. In managing and analyzing the data collected, the researcher used quantitative data analysis so the researcher analyzed the data by using statistical technique. In this study, the researcher used t-test to analyze the data in order to know whether the seventh grade students' vocabulary mastery in MTs Al Huda Bandung Tulungagung are better after being taught by using Blindfold Game. The SPSS 16.00 application was used to analyze the data.

F. Hypothesis Testing

The hypothesis of this research was as follows:

1. If t_{count} was bigger than t_{table} , the null hypothesis (H_0) was rejected. It means that the seventh grade students in MTs Al Huda Bandung Tulungagung achieved their vocabulary mastery better after being taught by using Blindfold Game.
2. If t_{count} was smaller than t_{table} , the null hypothesis (H_0) was accepted. It means that the seventh grade students in MTs Al Huda Bandung Tulungagung didn't achieve their vocabulary mastery better after being taught by using Blindfold Game.