CHAPTER III

RESEARCH METHOD

This chapter presents the description of the research method used in the study. It consists of (1) Research Design, (2) Population, Sampleand Sampling, (3) Data Collection Method, (4) Research Instrument, (5) Validity and Reliability Testing, (6) Normality and Homogeneous Testing, (7) Data Analysis.

A. Research Design

A research design encompasses the methodology and procedure employed to conduct scientific research. It is the blue print of the overall strategy that the researcher chooses to integrate the different components ofthe study in a coherent and logical way thereby, ensuring the effective addressing of the research problem.

Moreover, this research uses quantitative method. Quantitative research designs emphasise objectivity in measuring and describing phenomena (McMillan & Schumacher, 2010).According to Aliaga and Gunderson (2002), quantitative research is 'explaining phenomena by collecting numerical data that are analysed using mathematically based methods (in particular statistics). In another definition according to Muijs (2004) quantitative research is essentially about collecting numerical data to explain a particular phenomenon. Thus, the goal of quantitative research is to test the hypothesis.

Thus, the goal of this research is to examine the effectiveness of "Jeopardy" game-based learning and conventional strategy in students' vocabulary mastery and grammar achievement.Referring to the aim of the

research, an experimental research is conducted. As Robin (2009) states that the experimental research offers the best method available toresearchers to be able to investigate causality due to the high degree of control. Then, Ary (2010) stated that experimental research involves a study of the effect of the systematic manipulation of one variable(s) on another variable.

In an experimental research, the researcher can manipulate at least one independent variable, control over relevant variables and observes the effect on one or more dependent variable. In addition, Latief (2009) also said that experimental research is a powerful research method to establish cause and effect relationship with involving two or more variables, the variable that becomes the cause (independent) and the variable that becomes the effect (dependent). Hamid (2011) strengthen the idea above, that this research method is called as productive research because it intends to answer the hypothesis that related to cause and effect.

Quasi-experimental research design was chosen because only assign randomly different treatments in two different classes, and cannot artificially create groups for the experiment. Ary (2010) stated that this design is similar to randomized experimental design in that they involve manipulation of an independent variable but differ in that subjects are not to treatment groups. Creswell (2012), state that quasi-experiments include assignment of participants to groups. The design was chosen for two reasons: First, this study conducted in the organized classroom setting in which the classes were not allowed to rearrange the students or subjects for research. Second, the school schedules which have been arranged by the school cannot disrupt.

This design employed non-randomized control group pre-test and post-test. Creswell (2012) state that the researcher assigns intact groups of the experimental and control groups, administers a pre-test to both groups, conduct experimental treatment activities with only the experimental group, and then administers a post-test to assess the differences between the two groups. The pre-test in this research was used to measure the students' vocabulary mastery and grammar achievement prior to the treatment and to check whether the two groups are equal or not before the treatment, while the post-test in this research was to find out the effectiveness of the strategy employed.

Therefore, every research has variables. Variable is an object of study that became important points in research. It is divided into two, they are dependent variable and independent variable. Dependent variable is the variable that depends on other factors that are measured. These variables are expected to change as a result of an experimental manipulation of the independent variables. It can be stated as effect. In this research, vocabulary mastery and grammar achievement as dependent variables. While independent variable is the variable that is stable and unaffected by the other variables the researcher is trying to measure. It refers to the condition of an experiment that is sytematically manipulated by the investigator. It is called as cause. In this case, Jeopardy as independent variable.The research variables shown in the following table:

Table 3.1 Independent Variable

Group	Pretest	Independent variable	Posttest
E	Y1	XI	Y2

C	Y1	-	Y2

Notes:

E   : Experimental group, the group that is taught by using "Jeopardy" game-based learning.

C   : Control group, the group that is taught by using conventional strategy

Y1   : Pre-test is given to both groups to measure students' vocabulary and grammar achievement prior to the treatment.

Y2   : Post-test is given to both groups to know whether or not applying "Jeopardy" game-based learning and conventional strategy has any effect to enhance students' vocabulary mastery and grammar achievement.

X1   : Treatment, applying "Jeopardy" game-based learningto the students of experimental group

B. Population, Sample and Sampling

1. Population

It is not possible to study the entire population of a single school. A population is defined as a group of individuals, with at least one common characteristic which distinguishes that group from other individuals (Best & Kahn, 2006). According to Ary (2010) Population is a set (or collection) of all elements processing one or more attribute of interested.Mugo(2002) states that population is a group of individuals persons, objects, or items from which samples are taken for measurement. The population would firstly, be too large for a study of this limited scope and secondly, too diverse to be able to generalise the findings. It is for this reason that it is necessary to have a target population. A

target population consists of a specific group to whom findings might be generalisable. In this research, the target population is all the students at seventh grades of State Islamic Junior High School of 3 Trenggalek in the academic year 2018/2019.

2. Sample

Gay (1992) states that sample is the individual selected comprise. Selection of a sample is very important step in conducting a research. Regardless of the specific technique used, the steps in sample include identification of the population, determination of required sample size and selection sample.

Related to the sample, in this research, the researcher takes two classes become a sample. The classes are A1 and A2 class, then the total number both two groups are 58 students. After the researcher know about the real condition of the students like average scores presented from students' daily examination scores. The researcher divides into two groups. Two classes were assigned as the experimental group and control group. The A1 class becomes experimental group while A2 class becomes control group. A1 consists of 30 students, while A2 class consists of 28 students. They were chosen with the age range was 13 to 15 years old. As for gender, 40 females and 18 males in this study.

3. Sampling

Sampling is the act, process, or technique of selecting a suitable sample, or a representative part of a population for the purpose of determining parameters or characteristics of the whole population(Mugo, 2002). Gay (1992) states that sampling is the process of selecting a number of individual for a research in such a way that the individuals represent the larger group from which they selected.

In this case, the researcher selects the sample by using non-probability sampling with purposive sampling form, because the population have no chance of selection or the sampling don't have any chance to be selected, so the researcher takes all the sample both two groups.

C. Data Collection Method

Quantitative data collection methods rely on random sampling and structured data collection instruments that fit diverse experiences into predetermined response categories. They produce results that are easy to summarize, compare, and generalize. Data collection is a process of collecting information from all the relevant sources to find answers to the research problem, test the hypothesis and evaluate the outcomes.

The researcher collected the data through administering test. It means researcher administered the test in the form of vocabulaty test and grammar test. To carry out this research, fifty eight (58) students were divided into two groups with thirty and twenty eight students. One group was the experimental group and the other one was the control group. For the experimental group we had an introductory session before four-week experiment. In this session, the researcher explains what is *Jeopardy*and how it works including its rules. Then, the researcher gives list of vocabulary in term of adjective and its meaning and its synonym and material of grammar. Next, the researcher asks to the participants to read and memorize that vocabulary in a week (first week) and tells them that evaluation using *Jeopardy*will be conducted in second, third, fourth, fifth, and sixth week. No such session was held for the control group.

It is non-equivalent control group designs, that is research designs having both experimental and control groups but the participants are not randomly assigned to these groups. Moreover, observed differences between the two groups are assumed to be a result of the treatment. Both groups are pretested for the independent variable. The experimental group receives the treatment and both groups are post-tested to examine the effects of manipulating the independent variable.The first and the second score compared to know students' progress.

The data of this research is the score of the students' vocabulary and grammar score. The data is gained after administering pretest and posttest in both two groups. The resultsof the tests are calculated by giving one (1) score for correct answer and zero (0) for wrong answer.

D. Research Instrument

According to CINAHL (Cumulative Index to Nursing and Allied Health Literature(2018))research instruments are measurement tools (for example, questionnaires or scales) designed to obtain data on a topic of interest from research subjects. In this research, as the research instrument, the researcher administered the vocabulary test and grammar test.It is achievement test. The researcher wanted to measure the instrument through validitytesting, reliability testing, normality testing and homogeneity testing. The test is constructed by researcher herself. In other words, the test is called Researcher-made test, it means that the test was arranged by the researcher.

E. Validity And Reability Testing

1. Validity Testing

Based on the term validity refers to whether or not a test measures what it intends to measure. On a test with high validity the items will be closely linked to the test's intended focus (Rosynella Cardozo and Jonathan Magdalena, no year). To measure whether the test has good validity, the researcher analyzed the test from face validity, content validity, and construct validity.

a) Face validity

Hughes (1989) says that a test is said to have face validity if it looks as if it measures what it is supposed to measure. It could be said a test should be match between skill and component and test form. Face validity is the extent to which a test is subjectively viewed as covering the concept it purports to measure. It refers to the transparency or relevance of a test as it appears to test participants.

In this research, validating for face validity is conducted by knowing the form of test. The researcher uses multiple choice and describing term for vocabulary test and multiple choice and analyzing error for grammar test.

b) Content validity,

A test is said to have content validity if its content constitutes a representative sample of the language skills, structure, etc. with which it is meant to be concerned (Hughes, 1989). In other words, the test must represent language skills and component that will be testing and also it must be appropriate with the grade.

Related to this kind of validity, the content of this test is using daily adjectiveand verb vocabulary and for grammar it is about simple present tense in the form of nominal and verbal sentence that suitable with the material of seventh gradestudents. The test also refers to syllabus for seventh grade of Junior High School.

c) Construct validity

Based on Hughes (1989) a test, a part of a test, or a testing technique is said to have construct validity if it can be demonstrated that it measures just the ability which it is supposed to measure. The researcher tested the students' vocabulary mastery and grammar achievement.

Thus, the test of this research will be measure based on the hypothesis of vocabulary and grammar. On vocabulary mastery, Rosenfeld's (1994) list of strategies of successful readers includes a few vocabulary learning strategies, such as guessing a word's meaning from the context, identifying the grammatical category of a word, looking up words or recognizing cognates. While for grammar, some linguists (e.g. Hale and Keyser 1993) have argued that apparently simple words are syntactically complex and are the product of derivations involving movement and deletion.

d) Criterion-related validity

Criterion-related validity or criterion based on the validity is a validity measurement by comparing test scores with the performance. There are essentially two kinds of criterion-related validity, they are concurrent validity and predictive validity. In this research, we can

indicate that this research is predictive validity. It concerns the degree to which a test can predict candidates' future performance. In this case, we can predict students' achievement on first semester by this final test.

To apply this validity, first, the researcher explained the material about vocabulary related to adjective and verb and grammar focused on simple present tense before the test. After several times, the same group of students might take the same test and the scores, resulted from the first score and the second score are calculated for the correlation coefficient. The closer the correlation, to know the stronger the relationship between the two set of scores and the stronger the test to predict the students' future. In this research, the researcher used *Pearson Product Correlation Coefficient (PPMC)* though SPSS 25.0 version to find the correlation coefficient between the first and second scores.

According toSugiono (2013) the interpretation of correlation coefficient is divided into five classes as follow:

Coefficient interval	Meaning
0.00 – 0.199	Very Weak
0.20 – 0.399	Weak
0.40 – 0.599	Medium
0.60 – 0.799	Strong
0.80 – 1.00	Very Strong

Then, the score of r as follow:

1) If $r = 1$ or closer to 1, then the relation between X variable and Y variable is very strong and positive.

2) If r = -1 or closer to -1, then the relation between X variable and Y variable is very strong but negative

3) If r = 0, or closer to 0, then there is no relation between X variable and Y variable or the relation is very weak.

The correlation both two scores of vocabulary mastery evidence can be seen from the table:

Table: 3.2The Result of Criterion-related validity (Predictive Validity)

Correlations

		Vocabulary1	Vocabulary2
Vocabulary1	Pearson Correlation	1	.915**
	Sig. (2-tailed)		.000
	N	27	27
Vocabulary2	Pearson Correlation	.915**	1
	Sig. (2-tailed)	.000	
	N	27	27

The SPSS output suggests that the correlation coefficient is 0,915. It means that there is a positive correlation between variables. It also suggest that the ρ -value is 0.000. Considering that 0.000 is smaller than 0.05, so the null hypothesis is rejected.

While, the correlation both two scores of grammar achievement evidence can be seen from the table:

Table: 3.3Result of Criterion-related validity (Predictive Validity)

Correlations

		Grammar1	Grammar2
Grammar1	Pearson Correlation	1	.807**
	Sig. (2-tailed)		.000
	N	27	27
Grammar2	Pearson Correlation	.807**	1
	Sig. (2-tailed)	.000	
	N	27	27

The SPSS output suggests that the correlation coefficient is 0, 807. It means that there is a positive correlation between variables. It also suggests that the ρ -value is 0.000. Considering that 0.000 is smaller than 0.05, so the null hypothesis is rejected.

2. Reliability Testing

Moreover, a good test should be reliable. Rosenthal(1991) states that reliability is a major concern when a psychological test is used to measure some attribute or behaviour.Reliability refers to the consistency of scores obtained by the same individuals when re- examined with test on different occasions, or with different sets of equivalent items, or under other variable examining conditions (Anastasi, 1968). In other word, the more similar the scores would have been, the more reliable the test is said to be. Moreover, reliability is expressed numerically, usually as coefficient, a high coefficient indicates high reliability.

As Hughes (1989) stated that there are several things should be paid attention to make tests more reliable such as take enough samples of behaviour, do not allow candidates too much freedom, write unambiguous items, provide

clear and explicit instructions, ensure that tests are well laid out and perfectly legible, candidate should be familiar with the format and testing techniques, provide uniform and non-distracting conditions of administration, use items that permit scoring which is as objective as possible, make comparisons between candidates as direct as possible, provide a detailed scoring key, train scorer, agree acceptable responses and appropiate scores at outset of scoring, identify candidates by number, not name, and employ multiple, independent scoring.

Here for reliability to know the homogized participants, it employed with the try out to the students and the test score was calculated using *Cronbachs' Alpha Coefficient.* If the result of *Cronbachs' Alpha*is higher than 0.05 (reliable index < 0.05), it means that the test is reliable.

Obviously, the ideal test should be both reliable and valid. In this research, the researcher also used SPSS 25.0 for window to know the reliability of test instruments. The criteria of reliability instrument can be divided into 5 classes as follows (Ridwan, 2004), those are:

1. If the *alpha cronbach* score $0.00 - 0.20$: less reliable
2. If the *alpha cronbach* score $0.21 - 0.40$: rather reliable
3. If the *alpha cronbach* score $0.41 - 0.60$: enough reliable
4. If the *alpha cronbach* score $0.61 - 0.80$: reliable
5. If the *alpha cronbach* score $0.81 - 1.00$: very reliable

In this research, the researcher uses two types of vocabulary test—multiple choice and matching meaning test. The result of reliability testing of vocabulary mastery of multiple choice and matching meaning by usingSPSS 25.0 can be seen from the table:

Table: 3. 4Result of Reliability

Reliability Statistics

Cronbach's Alpha	N of Items
.875	50

From the table above, we can conclude that the Cronbach's Alpha score of multiple choice and matching meaning test type of vocabulary mastery is 0,875 and it is higher than 0.05 then, it means reliable.

While, grammar achievement is also divided into two types, they are multiple choice and analysis word. The result of reliability testing of multiple choices and analysis wordby using SPSS 25.0 can be seen from the table:

Table: 3. 5Result of Reliability

Reliability Statistics

Cronbach's Alpha	N of Items
.809	50

From the table above, we can conclude that the Cronbach's Alpha score of multiple choice and analysis word test type of grammar achievement is 0,809 and it is higher than 0.05 then, it means reliable.

F. Normality andHomogeneity Testing

1. Normality testing

This research works with statistic nonparametric to analyze the hypothesis. In this statistic nonparametric the data that will be analyzed should in normal distribution. One technique is that can be used to test normality by using *Shapiro Wilktechnique* because each group consist less than 50 students. The data

analyzed students' pre-test and post-test scores. If the sig/p is higher than 0.05 (sig/p > 0.05), it means that the test is normal distribution and H_1 is rejected. In other word, it can be stated the rules as follow (SinggihSantoso (2014)):

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

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

Table: 3.6The Statistical Correlations of Try OutScore of Vocabulary Mastery and Grammar Achievement with Shapiro Wilk Test

Tests of Normality

	group	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
score	V.M	.115	27	.200*	.952	27	.238
	G.A	.123	27	.200*	.982	27	.905

Based on the output table above, the result of df is 27 students, it means the data sample less than 50. Then, Shapiro wilk technique detects the normality of the research.

Therefore, the score of significant of vocabulary mastery is 0.238 and it is higher than 0.05 means the test of this group i normal distribution. While, the sig. of grammar achievement is 0.905 and it is higher than 0.05. Finally, we can conclude that the test is normal distribution.

2. Homogeneity testing

Homogeneity testing 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 *Levene's statistic*. If the result of sig higher than 0.05 (sig > 0.05), it means that the test is homogeneous.

Table 3.7 The Statistical Correlations of pre-test score (try out) of Vocabulary Mastery and Grammar Achievementwith *Levene's Statistic*

Test of Homogeneity of Variances

Score

Levene Statistic	df1	df2	Sig.
4.621	1	52	.063

Levene's Statistic

score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	294.000	1	294.000	1.797	.186
Within Groups	8509.037	52	163.635		
Total	8803.037	53			

From the table above, it can be assumed that the score of significant (Sig.) is 0.063 and it is lower than 0.05. Then, the variance of the data of vocabulary mastery and grammar achievement is homogenous.

G. Data Analysis

The researcher uses quantitative analysis by using statistical technique, because the data gathered from field is numerical data and can be formulated using statistical method. The scale measurement of the data was interval that focused on the students' ability score. Then, the raw data was fed into SPSS program 25.0 version software for further analysis. The data is analyzed by *MANOVA test*. MANOVA is used because there are one independent variable, it is Jeopardy and two dependent variables, they are vocabulary mastery and

grammar achievement. In other word, it used to test the hypothesis regarding the effect of one or more independent variables on two or more dependent variables. In this case, this test is used to test the hypothesis regarding the effectiveness of Jeopardy game base learning on students' vocabulary mastery and grammar achievement.

Homogeneity test of variance used to examine whether or not the variance between the independent variable group were equal. This test was one of the prerequisite tests of MANOVA. Levene's test of Equality of Error Variances was used based on the decision, if the significance value (Sig.) is > 0.05. It means that the variance between the independent variable groups were not equal.

Homogeneity test of variance that determined the variance between the independent variable groups, homogeneity test of covariance matrices should be conducted. The covariance matrices between the independent variable groups had to be equal too. The homogeneity test of covariance matrices could be done through Box's M test based on the decision, if the significance value was > 0.05, it means that the covariance matrices between the independent variable groups were equal. Yet, if the significance value was < 0.05, it means that the covariance matrices between the independent variable groups were not equal.

There was research hypothesis that should be tested. The hypothesis was the students' vocabulary and grammar had better improved achievement significantly by using Jeopardy game based learning than those improved with using conventional strategy. To test the hypothesis, the hypothesis was transformed into null hypothesis (symbolized $H_{0)}$. The Howas the students' students' vocabulary mastery and grammar achievement improved by using

Jeopardy game based learning has no difference from that improved by conventional strategy.

To reject the null hypothesis the researcher stated the alternative hypothesis (symbolized Ha). There was alternative hypothesis. The Ha was the students' reading ability and vocabulary improved by using task-based language teaching was better than that improved by conventional strategy.