CHAPTER III

RESEARCH METHOD

This chapter provides several points to be explained. The first one is started from the research design, the population and sample of the research, research instrument, validity, and reliability testing, normality and homogeneity testing, data collecting method, treatment, and data analysis.

A. Research Design

The research design used in this research is based on the form of the research question. According to Cresswell (2009) research design is a research strategy and technique that includes decisions ranging from broad assumptions to specific data collecting and analysis procedures. To carry this research, the researcher employed quantitative research methods. Quantitative research is a method of studying phenomena by collecting numerical data on the spot and then using statistical procedures to analyze them. In addition, quantitative sources primarily from the psychology domain, with a focus on statistics to allow generalization from population samples (Perry, 2005: 75). Dealing with the above statements, this research adopted quantitative research methods, and the design adopted was an experimental study, particularly a quasi-experimental study.

Since this approach does not require random sampling, the quasi-experimental design was used (Jackson, 2008:318). This research method administers some tests and treatment to assess the effect of using Hangman game toward vocabulary mastery of students in ATI English course Tulungagung. As

there is no random selection, the sample in this study constituted declared nonequivalent, consisting of an experimental and a control group (Jackson, 2008: 223).

In short, the researcher took two classes in the selected Course and employing pretest and posttest to examine the outcome of the treatment (Hangman game). The outcome was going to be presented dealing with the appearance of students' vocabulary achievement. Moreover, it was involving three aspects such as *knowing, understanding,* and *application* toward the vocabulary. The design of this research can be clearly understood in Table 3.1 below:

Table 3.1 Two Groups Pretest-Posttest Design

Group	Pretest	Independent variable	Post-test
(I-1) E	Y1	X	Y2
(I-2) C	Y1	-	Y2

Y1 has defined as a pretest administered before treatment to both the Experimental (E) and Control (C) classes. The purpose of the pretest was to measure the students' vocabulary mastery before the students obtained the treatment. Then, the treatment was considered as (X) and it was called teaching vocabulary by using Hangman Game. Treatment was only given to the Experimental group (E) while the Control group (C) was not. Meanwhile, Y2 was identified as Posttest. On this occasion, the researcher followed up on the treatment by attempting to learn about the students' performance in the experimental and control groups. Using this kind of study, the impact of experimental treatment may be evaluated by evaluating all tests above.

Besides that, two groups were selected as the sample of this research. They were named the experimental group and control group. The first group (I-1) had considered as the experimental group. I-1 was provided pretest (X1) then treated by using Hangman Game (T) and closed with giving them posttest (X2). Meanwhile, the second group (I-2) was named as the control group was only given Pretest (X1) and Posttest (X2).

In this study, experimental research was employed to assess the efficiency of employing the Hangman Game in students' vocabulary mastery. The outcome was assessed by presenting a specific therapy. The efficacy would be determined by determining the significant difference between students who were trained before and after utilizing the Hangman Game.

B. Population and Sample

1. Population

The term "population" refers to the total group of people from whom data is collected. The population can be said as the whole collection of real or threatened observable elements (Dewi, 2017: 39). In other words, the population includes all people whose actions should be taken. Dealing with the above term, the researcher determined the data population was the students of I class in ATI English course Tulungagung which consisted of 70 students. Those were divided into 3 rooms. They were i-1, i-2, and i-3. It can be seen as here:

Table 3.2 Population of Research

NO	Class	Gender	
		Male	Female
1	i-1	12 students	13 students
2	i-2	11 students	14 students
3	i-3	9 students	11 students
Total students		70 students	

2. Sample

Sampling was defined as a method by which a researcher determines a set of people as a sample that describes the number of people. A part of the representative of the entire population that is observed is called a sample (Fifah: 2016). Concerning the assertions above, a sample may be defined as a collection of units chosen from a larger group, often known as a population. The researcher chose 50 participants for such a study. They were divided into two classes, first was i-1 which consisted of 25 students for the experimental group, and i-2 which consisted of 25 students as a control group. The researcher took both classes because of the recommendation from certain and trustworthy parties in that course.

The technique of selecting a sample was done by using the purposive sampling technique. Purposive sampling had been a study sample that had been chosen by choosing certain individuals for a specific reason while taking into account short period, effort, and expense so that investigators did not have to obtain a large sample that was outside of the area (Sugiyono: 2009). Understanding that utilizing purposive sampling, the researcher considered certain recommendations from trustworthy people who were well-versed in which sample should be selected by including qualification. First,

the recommendation was from the head of ATI English course Tulungagung. He recommended that "i (1, 2, and 3)" was the most suitable class to be investigated because the students were more active. Second, the recommendation was from an English tutor who taught those classes. She gave some reasons why should take "i (1, 2, and 3) classes" as a subject, the reasons were:

1. That classes were taught by Vocabulary
2. That classes were cooperative enough and more active.
3. The students' characteristics were homogeneous in mastery Vocabulary, which meant not too good and not too bad.

C. Research Instrument

This instrument was used as a means to obtain the data related to the research question. Therefore, the researcher employed the instrument of research was in the form of vocabulary tests. In one test, it was categorized by different forms such as Section 1; translating the meaning (*knowing*), Section 2; multiple choices (*understanding*), and Section 3; making sentences (*application*).

The amount of translating vocabulary was 10; multiple choices were 10, and 5 for making sentences. Those tests were given to students at the beginning (pretest) and after employing the treatment (posttest). The purpose was to measure the students' achievement in vocabulary mastery before and after being taught by employing Hangman game as the treatment. The first test

was a pre-test which was distributed on April 6th, 2021, and the second test was named as post-test which had given on April 11th, 2021. Both of the tests were about the understanding of vocabulary (Adjective, Verb, and Noun) that had been considered based on their level and the existence of standard competence and basic competence as provided in the lesson plan of ATI English course Tulungagung. The content or vocabulary (Adjective, Verb, and Noun) that must be answered between pretest and posttest were different. It aimed to avoid bias as the students can memories the previous test is given.

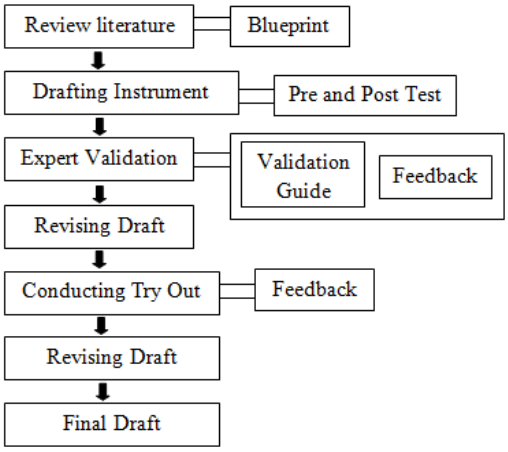
Next, as the attribute of distributing the test, the researcher needs a scoring rubric to examine and know the result in each test. The aspects of the scoring rubric were divided into three such as *knowing*, *understanding*, and *application*. First, the score of Section 1 *(knowing)* was 10 questions x 3 = 30; second was Section 2 *(understanding)* with 10 questions x 4 = 40; and the third was Section 3 *(application)* that presented 5 questions x 6 = 30. The maximum total score would be Section 1 (10x3) + Section 2 (10x4) + Section 3 (5x6) = 100 points and the minimum score was 0. Additionally, this case was no decimal result, thus any numeric score that exists should be sphered into a numerical number. The detail can be looked at in Appendix 6.

D. Validity and Reliability Testing

The validity and reliability of the tool are critical components of research because the instrument to be used must be capable and reliable before it was utilized to gather data. Then, ensuring that the instrument (test) was valid and

reliable was done by examining the tests' validity and reliability. The method of making an effective and reliable instrument is presented in figure 3.1 below:

Figure 3.1 Process in processing valid and reliable instrument



According to figure 3.1, the researcher studied the book and syllabus to create the test, which was the first step in determining the instrument's validity and reliability. Following the creation of the test, which included a pretest and a posttest, the researcher disclosed both tests to an expert checker for feedback using the validation design. The researcher then modified the test drafts in response to the input. Next, the Try out was given to the pupils in a separate class as a sample to gather feedback. Moreover, I-3 class was used to conduct the class. Eventually, after receiving input from the Tryout and adjusting with that concept, the next step was to obtain the final draft to test to i-1 and i-2 as a sample of the population of this investigation.

1. Validity

Validity was essential done by the researcher to ensure that the research instrument used was valid and applicable. Validity is defined as a

tool which is applied can be employed to measure what must be estimated (Sugiyono, 2015: 121). Meanwhile, the types of validity were various. Isnawati (2014) adds that the types of validity can be categorized as content validity, criterion validity, construct validity, and face validity. Understanding those, in this research, the researcher employed three types of them as follows.

a. Content Validity

Content validity was the first validation of instrument that focusing to the competence of content study. The extent to which a measuring instrument adequately satisfies the issue under investigation is known as content validity (Kothari, 2004: 74). The content validity of a test refers to the contents of the exam, which should reflect a representative sample of language abilities, patterns, and other factors (Idawati: 2014). As a result, the content of instrument in this research referred to the some competences such as *standar kompetensi* and *kompetensi dasar*. To ensure that, researcher designed content validity right below in Table 3.3.

Table 3.3 Content Validity

Standar Kompetensi	Mengekspresikan makna dalam percakapan transaksional dan interpersonal sangat sederhana untuk berinteraksi dengan lingkungan terdekat.
Kompetensi Dasar	Interaksi dengan lingkungan terdekat yang melibatkan tindak tutur: orang menyapa yang tidak diketahui, memperkenalkan diri / orang lain, dan memerintah atau melarang, mengetahui kata kerja, kata benda dan kata sifat
Tujuan Pembelajaran	Pada kegaiatan ini siswa dapat mengetahui vocabulary sekaligus bisa menterjemahkan kedalam bahasa inggris, dapat memahami arti vocabulary dan dapat mengaplikasikan vocabulary kedalam kalimat

Testing Objective	Untuk mengukur prestasi dalam pemahaman siswa mengenai kegunaan dari kosakata yang mereka dapatkan
Test Section	Translating Words, Multiple choice, Make Sentence
Material	Adjective, Noun, Verb
Test Score	100

b. Construct Validity

This type is the degree to that the experiments are consistent with the hypothesis that acts as the basis for developing such tests. Construct validity refers to the composing of instrument. The constructing of an instrument is referred to as construct validity. Construct validity is defined as the accuracy and significance of the inferences we make based on test scores (Bachman and Palmer, 1996: 21). Furthermore, Heaton (1975) adds that this test was used to determine the ability to measure certain characteristics in line with a linguistic approach to performance and learning.

Concerning this research, the instrument was constructed by concerning the aspects of vocabulary based on reviewing related theories. The aspects that must be concerned can be categorized as knowing, understanding, and application. Moreover, the components of vocabulary used in this research had determined based on student's level, they were Adjective, Noun, and Verb.

The aspects above had been the basis to make appropriate instruments for doing investigation about Hangman Game. As a result, before testing the instrument, the researcher spoke with qualified specialists to ensure that it was legitimate. The experts were English

tutor who handles i class in ATI English course Tulungagung and English lecturer of IAIN Tulungagung.

c. Face Validity

One of the validity kinds that may be determined is face validity. Face validity was a metric that revealed the instrument's consistency (Arikunto, 1998). Ary et al., (2010) Examiners consider the equipment as assessing what it is meant to measure, which is known as face validity.

Knowing above, in this instrument (pretest and posttest), there were some aspects to be considered for establishing a proper test such as:

1) In case of time allocation must be adequate. In this case, 45 minutes had been enough for accomplishing each test (section 1, section 2, and section 3)
2) The test content must be suitable with vocabulary that is still teaching to the students such as Adjective, Noun, and Verb.
3) The number of each section must be added 50% which means 10 Section 1, 10 section 2, and 5 section 3.

2. Reliability

After distributing the test, the reliability of the instrument must be processed. Ary et al., (2010) the level of coordination with which a measuring device holds whatever it is measuring defines its dependability. As result, it

indicates that if the students take the same exam on two separate days, the results should be equivalent. If a measurement is reliable and trustworthy, it is said to be reliable. It indicates that no matter how the exam is conducted, the results will be the same.

The researcher then utilized SPSS 16.0 for Windows to determine the test instrument's reliability. Hence, the students in i-3 who are in the same grade as the sample research subject (i-1 and i-2) were given try out. The try-out was used to see the reliability of the pretest and posttest. The result can be seen as in the table 3.4 below:

Table 3.4: The Score of Try Out Pretest and Posttest

NO	STUDENTS' NAME	PRE-TEST TRYOUT		POST-TEST TRYOUT	
		TEST	RETEST	TEST	RETEST
1	MMH	76	78	84	84
2	PND	73	78	79	82
3	NKDM	63	68	71	77
4	MHDT	71	73	75	77
5	KNA	71	73	84	84
6	MAZF	66	70	71	75
7	DPNH	74	76	79	79
8	CR	71	73	82	82
9	TRW	68	70	81	81
10	BS	73	75	80	80

Based on table 3.4 above, the researcher came to i-3 class twice on different days to obtain the data in the form of a score to reach the reliability of the instrument that would be used as pretest and posttest to experimental class (i-1) and control class (i-2). Then, the researcher sums that the comparison of scores both first test and retest was considered as same in case of difficulties level of the test. According to Ridwan (2004), the reliability instrument's criteria may be divided into five categories:

1. If the alpha Cronbach's score 0.00 – 0.20 means less reliable
2. If the alpha Cronbach's score 0.21 – 0.40 means rather reliable
3. If the alpha Cronbach's score 0.41 – 0.60 means enough reliable
4. If the alpha Cronbach's score 0.61 – 0.80 means reliable
5. If the alpha Cronbach's score 0.81 – 1.00 means very reliable

The result of the testing can be checked in the table 3.5 and 3.6 below:

Table 3.5: Reliability of Pre-test

Case Processing Summary

		N	%
Cases	Valid	10	100.0
	Excluded[a]	0	.0
	Total	10	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of items
.967	2

Table 3.6: Reliability of Posttest

Case Processing Summary

		N	%
Cases	Valid	10	100.0
	Excluded[a]	0	.0
	Total	10	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of items
.921	2

Based on the result of pretest showed that the Alpha Cronbach's score was 0.967 and posttest was 0.921 which meant the instrument used in this research both pretest and posttest were reliable to be implemented to experimental and control classes.

E. Normality and Homogeneity Testing

The data should be normal distribution and homogeneous before evaluating the significant difference score under the vocabulary of students taught using the Hangman Game against those taught using the traditional approach. One of the pre-requisite tests used to analyze the data at an independent sample of T-test, later on, is normality and homogeneity. The purpose was to see if the data had been assigned correctly and if there was any variation in variance between the two groups. In addition, the researcher attended both tests to evaluate the data computation continued normal distribution and homogeneity. Chapter 4 provides the results of both exams in chronological sequence. To ensure both terms, it can be defined deeply as below:

1. Normality Testing

This test was performed to see if a data collection can be properly represented by a normal distribution. The purpose of a normality test was to give sample data from a normally distributed population to determine its normalcy. As a result, the researcher utilized the Kolmogorov-Smirnov test

by significant level in IBM SPSS Statistic 16.0 in this case (0.050). Testing of the normality is provided by the points below:

1. When the significant value is more than 0.050, the data distribution is considered normal.
2. When the significant value is less than 0.050, the data distribution is not normally distributed.
3. Whereas if the data distribution is normal, the researcher moves on to the homogeneity test.

2. Homogeneity Testing

This test was proposed to ensure the data variation. Arikunto (2010: 98) supports that Homogeneity is a measurement that can be used to determine the data variation. There were so many ways that can be used to calculate the homogeneity such as analysis test or one-way ANOVA. In this research, the researcher employed one-way ANOVA particularly *Levene statistic* to check out the variances of data.

Homogeneity is meant as a parameter that may be used to identify data variation (Arikunto, 2010: 98). Several methods for calculating homogeneity can be included in analytical tests and one-way ANOVA. To examine the variants of data in this study, the researcher used a one-way ANOVA, which is a very *Levene* statistic.

The reason was that this research only used two variables. The researcher used SPSS 16.0 to measure it. The fundamental decision in presenting homogeneity testing can be categorized as below:

1. The data distribution is homogenous if the significance value is greater than 0.050.
2. When the significant value is less than 0.050, the data distribution is not homogeneous.

F. Data Collecting Method

The data collection was utilized to confirm that the research procedure was followed. Because the researcher cannot execute more than that time allotment, and the class does not belong to the researcher, the researcher conducted three sessions and two tests throughout this study. So, shortly the first stage of collecting data was done with a pretest then giving treatment three times, and finished with a posttest. To know the clear schedule of collecting data in this study as follows:

Table 3.7: The Schedule of the Research

No.	Class	Meeting	Date	Activity
1	Control	I	Tuesday, April 6th 2021	Pretest
2	Experimental			
3	Control	II	Wednesday, 7th April 2021	Conventional Teaching
4	Experimental			Hangman Game (Verb)
5	Control	III	Thursday, 8th April 2021	Conventional Teaching
6	Experimental			Hangman Game (Noun)
7	Control	IV	Friday, 9th April 2021	Conventional Teaching
8	Experimental			Hangman Game (Adjective)
9	Control	V	Sunday, 11th April 2021	Posttest
10	Experimental			

Understanding the above table, the tests were administered twice namely pretest before treatment and posttest after it. While the treatment or as named as Hangman Game was conducted in the second, third, and fourth meeting. The class

that was taught by Hangman Game was only an experimental class or i-1. On the contrary, i-2 class was taught by conventional teaching without any special treatment. The procedures were as follow:

1. Pretest on Tuesday, April 6th, 2021

Pre-test was administered in this research to know the students' capability on mastery of English vocabulary in the earlier knowledge before they obtained the treatment of Hangman Game. This test was conducted on Tuesday, April 6th, 2021 toward both experimental and control classes. The test consisted of three aspects of vocabulary mastery such as knowing, understanding, and application. Each aspect contained three parts of speech chosen before such as Adjective, Verb, and Noun. Moreover, the questions sections were divided into three kinds such as 10 numbers for *translating the meaning of the word* (Section 1), 10 numbers for *multiple choices* (Section 2), and 5 questions for *making sentences* (Section 3). All of the sections were in Simple present tense as same as students' level. By the number of tasks that must be answered by students, they must finish them in 45 minutes. The total number of pupils who received a pre-test was 25. The researcher utilized a scoring rubric to compute the pre-test score after completing the exam. The goal of calculating was to find out the results of the pre-test before teaching the Hangman Game as a treatment.

2. Post-Test on Sunday, 11th April 2021

After the series of treatments had been well done, the last section was distributing the posttest to the students. The date was on Sunday, 11th April 2021. The content of the test was different from than pretest. However, the sections were still the same such as 10 numbers for *translating the meaning of a word* (Section 1), 10 numbers for *multiple choices* (Section 2), and 5 questions for *making sentences* (Section 3). Each aspect contained three parts of speech chosen before such as Adjective, Verb, and Noun. All of the sections were in Simple present tense as same as students' level. By the number of tasks that must be answered by students, they must finish them in 45 minutes. The total of pupils who obtained the pre-test was 25 students.

G. Treatment

After presenting the pre-test, the researcher conducted the treatment to the students (i-1 class as an experimental class). The treatment was done three times. The first treatment had done on Wednesday, 7th April 2021. On that date, the researcher explains the way to play the Hangman game and practice it. The vocabulary component that was concerned was the use of *Verb*. Hence, the researcher did not only describe the game but also the knowledge of Verb to the students. During this phase, the students were classified into five groups, each with five members. In the last section, the students were asked about their

difficulties dealing with the material that had already been given by the researcher.

The second treatment had done on Thursday, 8th April 2021. This second treatment time has a similar way to the previous one dealing with the way of implementing Hangman Game. However, the researcher was concerned about the use of *Noun*. Moreover, the chosen group was asked to make a series of dotted lines according to the number of words to be guessed. Meanwhile, other groups guessed letters to other letters. At this moment, students did reading and translating the words.

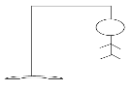
The third treatment was on Friday, 9th April 2021. In this time, the researcher taught vocabulary concerning the use of *Adjective* by using Hangman Game. Then, as in the previous meeting, they must make a series of dotted lines according to the number of words to be guessed. After that, the students did reading and translating the words. Then, the researcher asked about any difficulties that students had experienced dealing with the material that day. To close the meeting, the researcher concludes learning during that day.

Furthermore, the procedure of Hangman Game in teaching vocabulary can be known as follow:

1. Determine the vocabulary that will be given to students. between verb, noun, and adjective
2. Draw some rows representing the number of the letters in vocabulary words.
3. Have the students guessing the letter.

4. If the letter is in the word, write the letter in the appropriate now.
5. If the letter is not in the word, draw the hanged man (1-head, 2-torso, 3-arm, 4-arm, 5-leg, and 6-leg).
6. If they can guess the word before the drawing of a hanged man is done, then they win.
7. The complete drawing looks something like this:

Furthermore, the complete stages of the Hangman Game of this research can be seen in Appendix 9.



H. Data Analysis

After the whole process above, the data that had been collected must be analyzed to know the effectiveness of using Hangman Game toward students' vocabulary achievement. The researcher in this circumstance split the test results into two sections. They were in both the experimental and control classes. The Independent-Samples T-Test in IBM SPSS Statistic 16.0 was used to statistically evaluate both sets of data. The T-test was used to determine whether the significant value was more than or less than 0.050. If the result of significant value was higher than 0.050, it meant that using Hangman Game to teach students vocabulary mastery in English courses was considered effective.