

CHAPTER III

RESEARCH METHOD

The objective of this research is to examine the effect of using WhatsApp group chat in learning vocabulary. This chapter presents six topics dealing with the research method. Those are research design, population, sample and sampling, data collection method and research instrument, validity and reliability testing, and data analysis.

A. Research design

This research is aimed to examine the effectiveness of using WhatsApp group chat in learning vocabulary. Referring to the aim of the research, experimental research is conducted. Gay (1992) stated that experimental research is only of the research that can truly test hypotheses concerning cause and effect relationships. In experimental research, the researcher can manipulate at least one independent variable, control over relevant variables, and observes the effect on one or more dependent variables. Besides, Latief (2012) also said that experimental research is a powerful research method to establish cause and effect relationships involving two or more variables, the variable that becomes the cause (independent) and the variable that becomes the effect (dependent).

The quasi-experimental research design was chosen because only assign randomly different treatments in two different classes, and cannot artificially create groups for the experiment. Creswell (2012) stated that quasi-experiments include the assignment of participants to groups. The

design was chosen for two reasons: First, this study conducted in an 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 a pre-test and post-test for the experimental and control group. Creswell (2012) stated 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 before 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.

In this research, there were two variables measured. The independent variable is the use of WhatsApp Group Chat and conventional strategy, which is symbolized by X. The dependent variables are the students' vocabulary mastery, which is symbolized by Y. The research variables are shown in table 3.1 below:

Table 3.1
The Research Variable

Group	Pretest	Independent variable	Posttest
E	Y1	X	Y2
C	Y1	-	Y2

C : Control group

E : Experimental group

Y1 : Pre-test

Y2 : Post-test

X : Treatment

B. Population sample and sampling

1. Population

Gay (1992), stated that a population is a large number of groups to which is given treatment by a researcher whose results would be generalized. In this research, the population is all the students at the eleventh grade of MAN 1 Trenggalek in the academic year 2018/2019.

2. Sample

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

In this research, the researcher takes two classes to become a sample. The classes are XI IPS2 (C) and XI IPS3 (E) class, then the total

number of both two groups are 64 students. After the researcher knows about the real condition of the students like average scores presented from students' daily examination scores. The researcher divides both two groups, two classes were assigned as the experimental group and control group. The E class becomes an experimental group while C class becomes a control group.

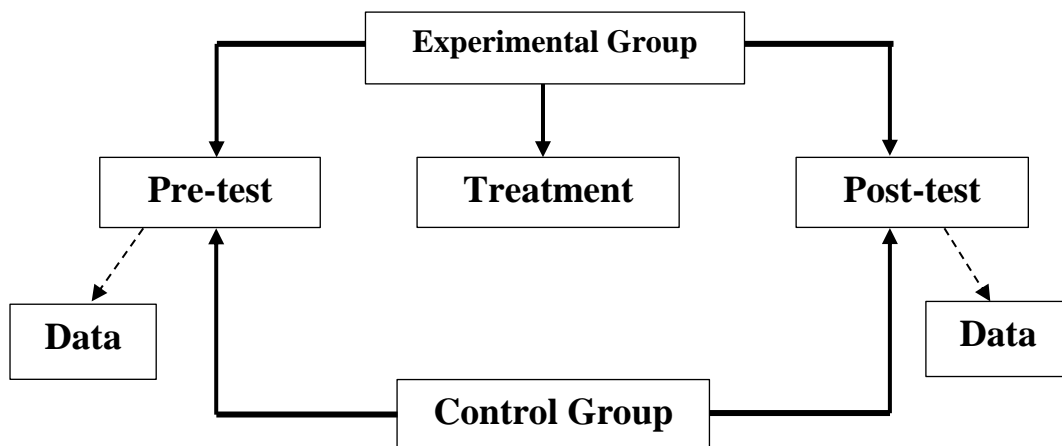
3. Sampling

Gay (1992) stated that sampling is the process of selecting some individuals for research in such a way that the individuals represent the larger group from which they were selected. The researcher selects the sample by using non-probability sampling with a purposive sampling form because the population has no chance of selection or the sampling doesn't have any chance to be selected, so the researcher takes all the sample both two groups.

C. Data collection method and research instrument

Tanzeh (2011) stated that the data collection method is a systematical and standard procedure used to collect data that is needed. In this research, the researcher collecting the data through administering the test. It means that the researcher administers the test in the form of a written test in the form of multiple-choice and matching words. The test gives twice, which is pre-test and post-test both to groups.

The diagram below represents the process of data collection. Both C and E group were given pretest and posttest, but only E group which given a treatment.



The technique of collecting data is clarified as follows:

a) Pre-test

The researcher conducts twice to test both two groups. On the first test calls pre-test. It conducted after the instrument has been validated. This test was followed by all members of class C and E that consists of 32 students in each class, so 64 students must complete the test. The researcher allocated 50 minutes for conducting the pre-test. The test is in the form of multiple-choice and matching words because it is suitable for testing vocabulary. The test contains 50 questions. It was done before the treatment process by using WhatsApp. The test was administered to know the students' ability on vocabulary.

b) Post-test

This test conducted after 2 months of treatment has done or it is held in the early of May before the students got a long holiday. The researcher conducted a post-test for both of the two groups, to know whether applying WhatsApp and conventional strategy has any effect to enhance students' vocabulary mastery or not.

All students answered the test then the researcher calculates the score by using SPSS. The test consist of 50 questions consists of multiple choice and matching word like what attached on the appendix. Every correct score will get 2 points, so there is 100 for those who answered perfectly.

c) Motivation Test

This questionnaire deals with motivation in learning English which includes intrinsic and extrinsic motivation proposed by Deci & Ryan (1985). The questionnaire is arranged based on the theory of motivation proposed by Keller (2009) which is known ARCS. It includes attention, relevance, confidence, and satisfaction.

The questionnaire uses 5 points Likert Scale to measure students' motivation which ranges from strongly disagree = 1, disagree = 2, rarely agree=3, agree=4, strongly agree=5.

d) Instrument

While the instrument, according to Gay (1992) is a tool to measure the knowledge skill, feeling, intelligence of an individual or group. Here,

the test is used to measure the students' vocabulary mastery. The researcher uses a type of achievement test, which means that the test should be representative of structure and skill that will be tested when the test must be appropriate with the grade. Here, the test used to measure the students' reading comprehension and vocabulary mastery, and the form of test is multiple choice and matching word both two forms can measure skill and component that will be tested.

A vocabulary test was developed to measure students' knowledge. The test consisted of 50 items. A total of fifty items were taken from the vocabulary lists that students had to learn for the semester. The test content and face validity of the questions as well as the difficulty level was checked by an experienced English teacher who suggested some changes to the original test, and it also validated by one of the senior English lectures in IAIN Tulungagung. (see Appendix 1)

The test was piloted with a group of 32 students who took the same English course in a different class. Necessary adjustments were made to the test. The test reliability was calculated using Cronbach Alpha. The vocabulary test was administrated as a pretest before the treatment to measure participants' vocabulary knowledge. The test is constructed by the researcher himself. In other words, the test is called the Researcher-made test, it means that the test was arranged by the researcher.

Table 3.2
Instruments of test

No	Instrument	Variable to measure	Function
1	Vocabulary mastery (pre-test)	Students vocabulary mastery before treatment	To see the homogeneity
2	Vocabulary mastery (post-test)	Students vocabulary mastery after treatment	To test the hypothesis

Here, the test used to measure the students' vocabulary mastery, the form of test is multiple choice and matching word both two forms can measure skill and component that will be tested. Besides, the test is suitable for eleventh grades because the content of the text refers to the syllabus for the eleventh grades of Senior High School.

The vocabulary mastery test was administered in experimental and control groups. This procedure covers several steps, they are the purpose recognition, establishment of the test blueprint, devising the test items, expert review, and revision, try out the test, and analysis and revision. The specification on the test included the objective, general instruction, test approach, kind of test, test type, the number of the text source, number of items, time allocation, equipment, and scoring.

The components of blueprint that has been validated by a senior lecturer include the subject matter, the grade, the specification of language areas to test, the scope of the area, the number of test items, and the proportion of the item. In the devising, the items are multiple-choice forms that were constructed to the vocabulary test in this study. There were 30 questions with multiple choice and 20 questions with matching

words. The students were asked to cross the right answer and match the word. The next step WhatsApps expert validation. The expert then finally checked whether the vocabulary test possesses evidence that meets the criteria of a good test.

Before the real test was given, the try-out of the test was done. The pilot testing was conducted on a particular subject since it has many characteristics in common with the main subject of the research. Further, the score was dissected to know the item facility/difficulty, items discrimination, reliability, and the efficiency of the distractor. After trying out the test, the test items were analyzed based on the students' scores. The test items analysis covers the analysis of item reliability, analysis of item difficulty, analysis of item discrimination, analysis of item validity.

e) Treatment

The experimental group (EG) installed WhatsApp on their cell phones, tablets, or laptops and they electronically received vocabulary instructions two sessions a week for two months (March-May). Treatment is another word for the experimental manipulation of the independent variable. The different treatment conditions administered to the subjects in the experiment are the *levels* of the independent variable.

In each session, ten vocabulary items were taught to students. All students in the EG were the member of the same WhatsApp chat group. The instructional materials initially available in the print form were

developed and converted into a mobile-based format. Different related pictures and videos were also sent to the students which conveyed the meaning of the new words and help the students with the pronunciations. Definitions, synonyms, and antonyms of the new words were also practiced in different sentences by using WhatsApp. They also received different drills to practice the new vocabulary items. The teacher asked the students to answer the questions individually and send their answers to the teacher.

This procedure followed the steps launched by Bensalem, 2019. Before the beginning of the experiment, participants were briefed about the purpose of the study. All participants took a vocabulary pretest before receiving their first list of vocabulary test items. They were assured that their test scores would not count toward their final grade. They were informed that the purpose of the test was to check their knowledge about some vocabulary. Both the experimental and control groups were given the same list of 120 words over 8 weeks with an average of 15 words per week.

The assignment consisted of looking up the meaning of new words in a dictionary and building a sentence using each word. The experimental group had the choice between using a monolingual dictionary app such as an online-dictionary or visit an online dictionary like google-translate to learn the meaning of target words. They had to build sentences using the target words and send them via WhatsApp for

correction. Participants of the control group were permitted to use any monolingual English dictionary.

After the completion of all homework assignments, participants took an unannounced vocabulary posttest during the day on which they had submitted their last homework assignment. The purpose of the posttest was to measure learners' vocabulary improvement. This helped determine the efficiency of each method of instruction in enhancing students' ability to learn new vocabulary. The pretest and the posttest were identical. The researcher had to reshuffle the items to avoid students memorizing the correct responses rather than answering the questions out of knowledge. At the end of the experiment, participants of the experimental group completed a questionnaire about their perception of learning vocabulary using WhatsApp.

D. Validity and Reliability testing

1. Validity

Gay (1992) stated that validity is the degree to which a test measures what it is supposed to measure. The researcher uses multiple choices and matching test, it means that form both of two groups can measure the skill and component of students. To measure the test has good validity, the researcher analyzed the test from *content*, and *construct validity*.

a) Face Validity

Face validity is if it looks as if it measures what it is supposed to be measured. For example, a test presented to measure pronunciation ability but, which did not require the test-takers to speak might be thought to lack face validity. This is true even if the test is construct and criterion-related validity can be demonstrated. Face validity is hardly a scientific concept, yet it is very important. A test that does not have face validity may not be acceptable by test-takers, teachers, education authorities, and employers. The researcher used face validity by consulting with the advisor and teacher.

In order to get face validity, a prototype of a vocabulary mastery test, test blueprint, and expert validation form are given to the expert to get a judgment on whether the test looks right to measure student's vocabulary mastery.

b) Content validity

On content validity, the most representative structure and skill that will be tested then the test must be appropriate with the grade. In this research, the content of questions on the test is using recount text and vocabularies related to the material (past tense/recount text).

The description of the test items used vocabulary recognition test can be seen in table 3.3 below

Table 3.3
Test Description

Objective	Types	Specific objectives	Items
To evaluate the students' vocabulary mastery of the target words	Less frequent words	Finding definitions of the target words, completing sentences, and recognizing the synonym/antonym	50

c) Construct validity

Isnawati (2011) stated that it has construct validity if it can be demonstrated that it measures just the ability which is supposed to measure. The word construct refers to any underlying ability which is hypothesized in a theory of language ability. In this research, the researcher tests the students' vocabulary by vocabulary test that have been arranged before on blueprint.

Table 3.4
Test construction

Objective	Type of test	Task
Measuring the vocabulary mastery	Vocabulary recognition test.	Students ask to answer match the right definition of the target words, fill the blank, and find the synonym/antonym

d) Criterion-related validity

Criterion-related validity applied to know how far results on the test agrees with those provided by some independent and highly dependable assessment of the candidate's ability. This independent assessment is thus the criterion measure against which the test is

validated. There are essentially two kinds of criterion-related validity: concurrent validity and predictive validity. (Hughes, 2002)

The researchers use *predictive validity*. The predictive validity of this concern the degree to which a test can predict candidates' future performance. An example would be how well a proficiency test could predict a student's ability to cope with graduation at MAN 1 Trenggalek.

The criterion measure here might be an assessment of the student's English as perceived by his or her teacher or researchers at MAN 1 Trenggalek to apply this validity, the developer or the researcher might administer a certain test before the students have begun the material about recount text explained by the researcher. 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 sets of scores and the stronger the test to predict the students' future. In this research, the researcher used the *Pearson Product Correlation Coefficient (PPMC)* though SPSS to find the correlation coefficient between the first and second scores. The correlation both two scores of vocabulary mastery evidence can be seen from table 3.5 below:

Table: 3.5
Result of Criterion-related validity (Predictive Validity)

Correlations

		Vocabulary Mastery 1	Vocabulary Mastery 2
Vocabulary Mastery 1	Pearson Correlation	1	,937**
	Sig. (2-tailed)		,000
	N	32	32
Vocabulary Mastery 2	Pearson Correlation	,937**	1
	Sig. (2-tailed)	,000	
	N	32	32

** . Correlation is significant at the 0.01 level (2-tailed).

The SPSS output suggests that the correlation coefficient is 0,937. It means that there was 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

Gay (1992), stated that reliability is the degree to which a test consistently measures whatever it measures. Or reliability is expressed numerically, usually as a coefficient, a high coefficient indicates high reliability. Then is a significant difference between the score of pre-test and the score of post-test both groups, so the result of the research is reliable.

To measure the reliability of the test, the researcher uses *Cronbach's Alpha*, if the result of *Cronbach's alpha* was higher than 0.05 (reliable index < 0.05), it means that the test was reliable.

The ideal test should be both reliable and valid. In this research, the researcher also used SPSS for Windows to know the reliability of test instruments. The criteria of reliability instrument can be divided into 2 classes as follows (RidWhatsApp, 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

The result of the reliability testing of vocabulary mastery by using SPSS can be seen from Table 3.6 below:

Table: 3.6
Result of Reliability
Reliability Statistics

Cronbach's Alpha	N of Items
,808	50

To know the items are reliable or not it can be seen from Cronbach's Alpha column. The Cronbach's Alpha score = 0, 808 means that it is very reliable.

E. Data Analysis

In quantitative research, to analyze the data by using statistical data analysis. The data gathered from the field is numerical and can be formulated using the statistical method. For practically, the SPSS program was used.