

## **CHAPTER III**

### **RESEARCH METHOD**

In this chapter, the writer presents discussion about research design, population and sample, research instrument, validity and reliability testing, normality and homogeneity testing, data collecting method, and data analysis

#### **A. Research design**

J.W Creswell(2009: 143) defines experimental studies as “the basic intent of an experimental design is to test the impact of a treatment (or an intervention) on an outcome controlling for all other factors that might influence that outcome.” In this study, the writer intended to find out whether List-Group-Label (LGL) which were given to the experimental group was effective and contributed to the students achievement to academic vocabulary development.

There are many designs in experimental research, such as pre-experimental, true experimental, single-subject, and quasi-experimental designs. The writer conducted this study by using quasi-experimental research design.

According to J.W Creswell (2009: 158) in Quasi-experiments, the investigator uses control and experimental groups but does not randomly assign participants to group (e.g. they may be intact groups available to the researcher). It has many examples of designs, for instance, single-group interrupted lime-series design, control-group interrupted lime-series design, and non-equivalent (pre-test and post-test) control-group design.

In this study, the writer administered pre-test and post-test with experimental and control groups design. J.W Creswell (2009: 160) states that the

experimental group A and the control group B are selected without random assignment. Both of groups took a pre-test and post-test, but only the experimental group received the treatment. Therefore, the writer administered a pre-test and a post-test for experimental and control groups, but there was only one group (the experimental group) which was taught by using List-Group-Label (LGL) strategy. The design of the experiment can be described as follows:

Group	Teaching vocabulary	Treatment	Teaching vocabulary
Experimental	Pre-test	Using LGL strategy	Post-test
Control	Pre-test	Without LGL strategy	Post-test

Based on the research design above, the writer involved two groups to conduct this study. First, a experimental group which received treatment by using List-Group-Label (LGL) strategy. Second, an group which was taught by without using List-Group-Label as a control group.

In order to measure the two group initial ability, a pre-test was carried out. The experimental group received the treatment by using List-Group-Label (LGL) strategy, but the control group was taught by using without List-Group-Label. The writer administered post-test for both groups to measure the students' achievement after they received the treatment.

## **B. Population and samples**

Sugiyono (2014: 117) also defined that population is generalization area that consist of object or subject which having certain qualities and characteristics determined to learn then concluded by writer. Meanwhile, Fraenkel *et al* (2012: 91) explained that population was the larger group to which one hopes to apply

the results. So that, population was the larger group of individual that has one more characteristics in common that are of interest to the writer.

In this research, the writer can take the population was all students coming from the seventh grade students (class 7) of MTs N 1 Blitar of academic year 2017/2018. They are all students who supposed have been actively treated with English instruction. There are 8 (eight) classes in that school, each class consists of 38 students. So the total number of the population in that 7<sup>th</sup> year level is 304 students.

In this study, the writer used purposive sampling technique to take the sample. Purposive sampling was based on the judgement of the writer as to who will provide the best information to succeed for the objectives study. According to J.W Creswell (2009:146), purposive sampling techniques that have also been referred to as nonprobability sampling techniques, involved selecting certain units or cases “based on a specific purpose rather than randomly.”

The sample was taken two classes from seventh graders of MTs N 1 Blitar. They were VII-7 class as the experimental group that taught by List-Group-Label and VII-8 class as the control group that taught without using List-Group-Label (LGL). The number of student was 30 for VII-7 Class and 30 for VII-8 Class. The writer used this purposive sampling due to suggestion from the English teacher that both classes have the same number of students and that both classes have equal of English ability and have same gender is male.

### **C. Research instrument**

In order to collect the data, the writer designed an academic vocabulary test in multiple choice test items format. This research used a test as instrument. According to Ary *et al* (2010: 201), “Test is a set of stimuli presented to an individual in order to elicit responses on the basis of which a numerical score can be assigned”. In this case, the writer gave vocabulary test to get the data or information. Vocabulary test As in Chapter 2, there were three languages testing in vocabulary such as multiple choice, completion, gap-fill. In this research, the writer used multiple choices test. The reason for using multiple-choice items was they were undoubtedly one of the most commonly used types of item in objective test. The questions consisted of 30 multiple choices item test.

The instruments of this use test multiple choices for collecting data. The vocabulary test is animal’s name and body’s name. This consists of 30 items for maximal 45 minutes. The score processes that will be given if the student answer with correct they will get 2 then the score will be  $30 \times 2 = 60$ , and if the study’s answer with incorrect they will get 0. Every right answer was be scores two values. The result of match can see the table below.

### 3.1 Table of scoring rubric

Jumlah	Nilai	Jumlah	Nilai	Jumlah	Nilai	Jumlah	Nilai
60	100	47	80	34	58	21	39
59	98	46	78	33	57	20	38
58	97	45	77	32	55	19	36
57	95	44	75	31	54	18	34
56	94	43	73	30	52	17	33
55	92	42	72	29	50	16	31
54	91	41	70	28	49	15	30
53	89	40	69	27	48	14	28
52	88	39	67	26	46	13	27
51	86	38	66	25	44	12	25
50	84	37	64	24	43	11	23
49	83	36	62	23	41	10	22
48	81	35	60	22	40	9	20

#### D. Validity and Reliability Testing

The best instrument had to fulfill two importance requirements, these were validity and reliability. Validity and reliability were used to test the legality of data. These were the explanations of validity and reliability below:

##### a. Validity

The writer used validity to know her research instrument was valid or not. According to Gay (1992:52) validity was measure what to be measured, a measurement showed levels of research instrument validity.

These are four types of validity; 1) Content validity, 2) Criterion related validity, 3). Construct validity, 4). Face validity. In this research, to measure whether the test has a good validity, the writer analyzed the test from content validity and construct validity.

a) Content validity

The test was called content validity if the content of test was suitable to what was going to be tested. According to Gay (1992:56) content validity was the degree to which a test measures an intended content area. Content validity was guaranteed if contents of test constituted representative sample of the language skill, structures, etc. being tested. It means that the test must be appropriate with the material in certain level of education. In developing items of test, writer had to know the curriculum of school which used to do research. Because the school uses school-based curriculum, the writer developed the test based on the school-based curriculum by reading the core-competence and base-competence. The writer only used the core-competence and base-competence of reading skill because this research wanted to know the students' achievement of vocabulary integrated in teaching of reading.

In this research, the test which was given twice at pre-test and post-test was in the form of multiple choices. The test was made up based on course objective the syllabus of seventh Graders of MTs N 1 Blitar. The contents validity in this research could be shown from the table:

**Table 3.2 Blue Print of Test.**

Kompetensi inti	Kompetensi dasar	Kls/smt	Materi	Indikator	Sub-indikator soal	Bentuk soal	No. soal
Mencoba, mengolah, dan menyaji dalam ranah konkret (menggunakan, mengurai, merangkai, memodifikasi, dan membuat) dan ranah abstrak (membaca) sesuai dengan yang dipelajari di sekolah dan sumber lain yang sama dalam sudut pandang/teori	Membandingkan fungsi sosial, struktur teks, dan unsur kebahasaan beberapa teks deskriptif tulis dengan memberi dan meminta informasi terkait dengan deskripsi orang, binatang, dan benda sangat pendek dan sederhana, sesuai dengan konteks penggunaannya	VII/2	Descriptive text	Parts of body	Identify the word to be one word group in group parts of body and find synonym	PG	1, 2, 3, 4, 5.
				Parts of body	Identify the word to be one word in group parts of body and translation.	PG	6, 7, 8, 9, 10, 11, 12, 13, 14, 15.
				Kind of animal	Identify the word to be one word group in group name animals and find synonym	PG	16, 17, 18, 19, 20.
				Kind of animals	Identify the word to be one word in group name animals and translation.	PG	21, 22, 23, 24, 25, 26, 27, 28, 29, 30.

## b) Construct validity

A test was said to have construct validity if it can be demonstrated that was measure just the ability which was supposed to measure. Gay (1992:157) said that construct validity was degree to which a test measured an intended hypothetical construct. Construct validity cannot be seen but the effect can be observed. It means that construct validity was used to explain students' behavior. In this research, the instrument has been constructed based on vocabulary mastery theory. After the instrument was constructed, the test was tried out and then the writer used SPSS 16.0 of Pearson Correlation to count the validity test per items.

Basic decisions making in validity testing per items are as follows:

1. If the score of  $r_{hitung} > r_{table}$  in score signification 5%, then the test items is valid.
2. If the score of  $r_{hitung} < r_{table}$  in score signification 5%, then the test items is not valid.

The process calculation of validity testing (see appendix 9) by using SPSS 16.0 version for windows found that from the 30 questions of multiple choices which had been tried out, there were only 21 questions valid. The result of validity can be seen as follows:

**Table 3.3 Result of Construct Validity**

No. Item	$r_{hitung}$	$r_{table\ 5\%}$	Kriteria
1	0,208	0,482	Invalid
2	0,637	0,482	Valid
3	0,658	0,482	Valid
4	-0,045	0,482	Invalid
5	0,511	0,482	Valid
6	0,508	0,482	Valid
7	0,530	0,482	Valid
8	0,556	0,482	Valid
9	0,299	0,482	Invalid
10	0,574	0,482	Valid
11	0,602	0,482	Valid
12	0,658	0,482	Valid
13	0,283	0,482	Invalid
14	0,697	0,482	Valid
15	0,188	0,482	Invalid
16	0,555	0,482	Valid
17	0,574	0,482	Valid
18	0,583	0,482	Valid
19	0,581	0,482	Valid
20	0,496	0,482	Valid
21	-0,151	0,482	Invalid
22	0,136	0,482	Invalid
23	0,715	0,482	Valid
24	0,600	0,482	Valid
25	0,581	0,482	Valid
26	0,080	0,482	Invalid
27	0,574	0,482	Valid
28	0,608	0,482	Valid
29	0,556	0,482	Valid
30	-0,020	0,482	Invalid

The table above presented there were 21 valid items of 30 items and 9 invalid items. The writer eliminated those 9 invalid items and selected 20 of 21 valid items which were used as the pre-test and the post-test. There were 13 valid items which had high discriminating power and it was needed 7 valid items more, so the writer chose 7 of 8 valid items which had low discriminating power. Based on its criteria of the item facility, the test items number 5 and 7 belonged to easy

items because more student have done correct answer, and the test items number 3, 6, 11, 12, 18, and 23 were appropriate items. Because there were many appropriate test items which had been already selected, the writer preferred to select the test items number 3, 5, 6, 7, 11, 12, and 18. As a result, the pre-test and the post-test consisted of 15 appropriate test items, 3 difficult test items, and 2 easy test items.

### **b. Reliability**

A test which gave the same result of measurement was reliable. According to Gay (1992:161) reliability was the degree to which a test consistently measured whatever it measure, an absolute requirement to determine one variable effect to another.

Reliability was also requirement for validating a test. It means that the test which was not reliable, it cannot be valid automatically. Reliability was expressed numerically, usually as a coefficient. A high coefficient indicated high reliability and a low coefficient indicated low reliability (Gay, 1992:162).

Reliability test instrument can be done by using Cronbach's Alpha. The instrument has a high degree of reliability if the value of Cronbach's Alpha obtained as follows.

**Table 3.4 Cronbach's Alpha Interpretation**

<b>Cronbach`s Alpha</b>	<b>Interpretation</b>
0,00 – 0,20	Less Reliable
0,21 – 0,40	Rather Reliable
0,41– 0,60	Quite Reliable
0,61 – 0,80	Reliable
0,81 – 1,00	Very Reliable

The result of reliability testing by using SPSS 16.0 can be seen from the table:

**Table 3.5 result of reliability**

<b>Reliability Statistics</b>	
Cronbach's Alpha	N of Items
.906	20

From the table above, the value of Cronbach alpha is 0,906. It means that the test is very reliable.

### **E. Normality and Homogeneity**

In this part the writer discussed about the result of normality and homogeneity testing.

#### **1. The Normality Testing**

Normality testing is conducted to determine whether the data are normal distribution or not. The writer used SPSS.16 one sample kolmogrov-smirnov Test by the value of significance ( $\alpha$ ) = 0.050.

Basic decisions making in normality testing are as follows:

1. If the significance value  $> 0.050$ , then the data has normal distribution
2. If the significance value  $< 0.050$ , then the data does not have normal distribution

The writer conducted normality testing from gain experimental class and control class. The result can be seen below:

**Table 3.6 Result of normality test gain score**

<b>One-Sample Kolmogorov-Smirnov Test</b>		Gain
N		60
Normal Parameters <sup>a</sup>	Mean	3.4500
	Std. Deviation	3.07767
Most Extreme Differences	Absolute	.141
	Positive	.141
	Negative	-.130
Kolmogorov-Smirnov Z		1.089
Asymp. Sig. (2-tailed)		.186
a. Test distribution is Normal.		

Based on the table above, the significant value in gain score from experimental group and control group was 0.186 which was higher than the significant 0,050. It means that the pre-test and post-test data experimental groups and control group were normally distribution.

## 2. Homogeneity

Homogeneity testing is conducted to know whether the data has a homogeneous variance or not. To know the homogeneity, the writer used Homogeneity of Variances Test by using SPSS.16. The value of significance ( $\alpha$ ) = 0.050.

Basic decisions making in homogeneity testing are as follows:

1. If the significance value  $> 0.050$ , then the data distribution is homogeneous
2. If the significance value  $< 0.050$ , then the data distribution is not homogeneous

The writer conducted Homogeneity testing from gain experimental class and control class. The calculation result of homogeneity testing from experimental and control class is as follows:

**Table 3.7 result of homogeneity**

**Test of Homogeneity of Variances**

Gain

Levene Statistic	df1	df2	Sig.
.480	1	58	.491

The data was homogeneous, if the significant value was higher than the significant 0,050. From the table above, the significant value was 0.491. Because the result was higher than significant 0,050, the gain between the experimental and the control groups was homogeneous.

## **F. Data Collecting Method**

One of the most important steps in doing research is collecting data to find out the result of the research. The procedures of data collection followed some steps.

### **1) Try-out test**

The writer conducted the try-out test to design a good test and it was administered to the students who belong to neither the experimental nor the control groups. By using the students' scores of the try-out test, the writer did item analysis and an analysis to find out the validity and reliability of the test. There were 30 items in the form of multiple choice items (see appendix 4) and the students were given 60 minutes to do the try out.

The try-out test was conducted on Saturday, February 24<sup>th</sup> 2018 and involved 17 students of VII-G they belong to neither the experimental nor the control group. They had to complete a test which consisted of 30 items in 60 minutes (see appendix 4). By using the students' scores of the try-out test (see appendix 1), the item analysis and instrument analysis (to find out the validity and reliability of the test) were performed to construct a good test.

### **2) Pre-test**

Pre-test test was carried out to find out the initial students' knowledge of academic vocabulary and it was given once before both the experimental and control groups got a treatment. There were 20 items in the form of multiple choice items (see Appendix 6) and the students were given 45 minutes to do the pre-test.

The writer conducted the pre-test to measure the students' initial academic vocabulary mastery before getting a treatment on Monday, March 05<sup>th</sup> 2018 for the control group and experimental group have same schedule. There were 30 students of each group who participated in the pre-test and they were given 45 minutes to answer 20 multiple choice items question about vocabulary. The students' pre-test scores of the experimental and the control groups (see Appendix 2-3).

### 3) **Post-test**

The purpose of administering post-test was to measure the students' ability in developing their academic vocabulary after they received the treatment. By analyzing the students' post-test scores, the writer could measure the significant difference in students' achievement between the experimental and control groups. In addition, the test items in the post test and the pre-test were the same. There were 20 items with multiple-choice test items format in this test.

The writer administered the post-test to measure the students' academic vocabulary achievement after the treatment was given. It was conducted on Tuesday, 13<sup>th</sup> 2018 for the control group and the experimental group have same schedule. The students' post-test scores of the experimental and the control groups (see Appendix 2-3).

### **G. Procedure of Treatment**

In this study, the different treatment was given to two groups. The experimental group was given a treatment by using List-Group-Label (LGL) strategy in learning academic vocabulary. Meanwhile, the control group was

taught by using without List-Group-Label strategy. The treatment was given to the students before the post-test was administered. The lesson plans of each group which explained the detail implementation of the treatment are presented in Appendix 17-18.

The first treatment was conducted on Tuesday, March 06<sup>th</sup> 2018 and involved 30 students of VII-7 who enrolled English subject in the cross interest program. Because it was the first time for the students applied List-Group-Label (LGL) strategy, the writer gave a brief explanation to the students how the learning activities were carried out. The topic which was given in the first treatment was “*person*”.

First activity is listing. Each student brainstorms words by reading a descriptive text about given by teacher. Ask students to list as many words as they found that they do not know the meaning.

The second is Group. When each student completes the brainstorming part of this activity (List), divide the class into small group. Students then work in small group to share and combine their words into logical categories. Each group will work to cluster the class list of words into subcategories. As students attempt to combine three to five individual words lists, they begin to discover patterns of words. In this process, they are refining their knowledge of the concept. Students have to work together to combine their individual list into a common group list that cover all of their words. Once students create their categories, they give label each of them by giving part of speech. After that ask students to look for the meaning of each words

After students know the meaning, they should find the synonym each of them. Synonyms are group of words that have the same or almost the same meaning (Smith:2013:1)

The second treatment was processed on Monday, 12<sup>nd</sup> March 2018. The writer gave the different topics in each treatment and the topic is animal given in the second treatment. The learning activities were carried out by following the steps of List-Group-Label (LGL) strategy. In the second treatment, the writer found that the students could accomplish the activities better than in the first treatment. It showed that they wrote more list of words than what they listed in the first treatment. The writer implied that it happened because of some reasons. One of the reasons was the students started to get used to apply List-Group-Label (LGL) strategy in their learning activities. Besides, it appeared that the students learned the learning materials related to the topic animal.

#### **H. Data Analysis**

Processing the data leads the writer to prove whether the hypothesis stated can be confirmed or rejected completely. The data obtained in this investigation is analyzed statistically. The statistical analysis used is the central tendency of variance. The writer will first, find the mean score of the sample, and then find the deviation. And to know the significant differences researcher used SPSS 16.0 for window.