

## CHAPTER III

### RESEARCH METHOD

In this chapter concluding; place and time of study, research type, research design, population and sample, instrument of the study, instrument validity, instrument reliability, data collecting procedures, and data analysis procedures.

#### A. Place and Time of Study

The study was conducted at SMP Negeri 1 Palangka Raya which is located on A. Yani Street. The writer starts the research from December until January.

#### B. Research Type

The writer used quantitative design to analyze the data obtained. Related to the research design, Aliaga and Gunderson in Daniel stated that “quantitative research is explaining phenomena by collecting numerical data that are analyzed using based methods”.<sup>1</sup>

#### C. Research Design

The writer used quasi experimental as a design in this study. Quasi experimental design is a form of experimental research used extensively in the social sciences and psychology.<sup>2</sup> The design is made since the study works to measure the effectiveness of team pair solo technique on speaking performance score of the eight graders of SMP Negeri 1 Palangka Raya.

This study used non-random. There are two groups in this model, control group and experimental group. Both of groups are given pre-test to measure the score of students before the treatment is given (Y1 and Y2).

---

<sup>1</sup> Daniel Muihs, *Doing Quantitative Research in Education*, London: Sage Publications, 2004, p.1

<sup>2</sup> Saiful Mujab, *The Effectiveness of Using English Song Video Toward Students' Speaking Score at the Second Semester Students of English Departement of The State Islamic College of Palangka Raya*, 2010, p. 38.

The treatment is given for experimental group (X). Post-test is given for both of groups to measure the score of students after the treatment is given (Y1 and Y2). In this experimental, team pair solo is used as a technique in teaching speaking for experimental group.

**Table 1.3 The Scheme of Quasi Experimental Design**

Subject	Per-test	Treatment	Post-test
A	Y1	X	Y2
B	Y1	-	Y2

Where:

A : experimental group

B : control group

X : treatment

Y2 : post-test

Y1 : pre-test

## **D. Population and Sample**

### **1. Population**

According to Ary, population is defined as all members of any well-defined class of people, events, or object.<sup>3</sup> The study was conducted at SMP Negeri 1 of Palangka Raya. The population of the study was taken in eight classes which contains of 255 students.

**Table 1.4 The Number of Eight Graders of SMP Negeri 1 Palangka Raya.**

---

<sup>3</sup>*Ibid*,p. 138

No	Classes	The Number of Students
1	VIII – 1	31
2	VIII – 2	32
3	VIII – 3	32
4	VIII – 4	32
5	VIII – 5	31
6	VIII – 6	32
7	VIII – 7	32
8	VIII – 8	33
	Total	255

## 2. Sample

The small group that is observed is called a sample.<sup>4</sup>The writer used cluster sampling to take the sample. Cluster sampling refers to groups or chunk of elements that would heterogeneity among members which each group are chosen for study.<sup>5</sup> By cluster sampling, the writer selects specific number of classes and tests all the students in those selected class, where the class has been determined by the teacher. The classes are:

**Table 1.5 The Number of Sample**

---

<sup>4</sup>*Ibid*, p. 148

<sup>5</sup>Sabarun, *Population and Sampling*, Unpublished Material for Writing IV:Palangka Raya, p. 2

No	Classes	Groups	Number of the students
1	VIII-3	Experimental	32
2	VIII-2	Control	32
Total			64

### E. Instrument of the Study

According to J.B. Heaton there are some production tests. They are reading aloud, oral drills, using picture for assessing oral production.<sup>6</sup> In this study, the oral test will be used to measure the students' speaking performance score. Dr Alan Wildeman in Joughin stated that oral assessment as providing powerful moments of articulation when the student is able to express his or her knowledge in a relationship with a trusted senior member of the discipline they are studying (private, communication).<sup>7</sup> Joughin stated that oral assessment lends itself to the assessment of different kinds of content; can utilize instruction between student and examiners require students to process their ideas for an audience and may use multiple modes of communication to support the oral component of the assessment.<sup>8</sup> The test is oral test that consist of two questions. Every student needs 2-4 minutes to answer the questions. For posttest, the experiment group was given the same test with the control group according the material in syllabus.

The mayor data in this study are the students' speaking performance score those were taken from pretest and posttest. The pretest was used to measure the students' speaking ability before the treatment is applied by the writer. The treatment was applied in experiment group. The test was scored by two testers. Because the test is oral test, the writer divided the

---

<sup>6</sup> J.B. Heaton, *Writing English Language test*, United State of America, Logman Group UK Limited, 1988, p.98.

<sup>7</sup> Gordon Joughin, *A Short Guide to Oral Assessment*: University of Wollongong, p. 8 (Online on February 2015)

<sup>8</sup> *Ibid*, p. 8

score into four criteria, which are the scores of pronunciation, grammar, vocabulary, and fluency. Each criteria, then was rated into four aspects of scoring, it was based on O'Malley J.M.<sup>9</sup>

No	Language Elements	Score	Description
1	Fluency	5	Able to speak without any hesitation or break in a sentence
		4	Hesitates or pauses at a few places in a few sentence but fluency is not seriously affected
		3	Hesitation is shown in almost every sentence. Ends sentence half-way and then repeats it.
		2	Hesitation a lot that message is not very clear. A lot of repetition of sentence.
		1	Shows a lot of hesitation that speech is not clear
2	Grammatical accuracy	5	Grammar is correct
		4	A few grammatical errors which affect the meaning.
		3	One or two major errors which affect the meaning.
		2	Almost every sentence contains grammatical errors.
3	Pronunciation	1	A lot of grammatical errors are made that meaning of messages is not clear.
		5	All words are clearly and correctly pronounced.
		4	One or two words are incorrectly pronounced but meaning is affected.
		3	A few words are incorrectly pronounced but meaning is affected.
		2	One or two word which are mispronounced because meaning to be affected.
1	Most of words are incorrectly pronounced.		
4	Vocabulary	5	Uses extensive vocabulary
		4	Uses varied vocabulary
		3	Uses adequate vocabulary
		2	Uses limited vocabulary
		1	Uses functional vocabulary

<sup>9</sup> O'Malley, J.M & Pierce, L.V, 1996. *Authentic Assessment for English Language Learner : Practical Approach for Teacher*. Massachuassetts: Addison-Wesley

## F. Research Instruments Validity

Validity is defined as the degree to which evidence and theory support the interpretations of test scores entailed proposed uses of tests.<sup>10</sup> There are two types of validation, they are content validity and face validity.

### 1. Face Validity

Face validity is a term sometime used in connection with test's content.<sup>11</sup> Face validity of test is when the test is indeed testing what it claims to test. For face validity of the test items as follow:

- a) The test is about daily activity.
- b) The evaluation by speaking test based on scoring system.
- c) The language of items uses English.
- d) The test items are suitable with syllabus of English for eight graders of SMP Negeri 1 Palangka Raya.

### 2. Content Validity

According to Heaton, content validity is "Kinds of validity depends on a careful analysis of the being test and the particular course objective. The test should be as constructed as contain a representative sample of the course, the relationship between the test item and the course objectives always being apparent".<sup>12</sup>

---

<sup>10</sup> Donald Ary, *Introduction to Research in Education, Eight Edition*, p. 225

<sup>11</sup> J. B. Heaton, *Writing English Language test*, England: Longman, 1974, p. 228

<sup>12</sup> *Ibid*, p. 153

A comparison of test specification and test content is the basic for judgment for content validity. The test was made based on the course objectives in the syllabus of eight graders of SMP Negeri 1 Palangka Raya. The writer used inter-rater method. Inter-rater method is two raters who score the students' speaking performance to get the score composition as possible. The writer used product moment correlation as the formula to calculate the validity from the test result.<sup>13</sup>

$$r_{xy} = \frac{N \sum XY - (\sum X) (\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}}$$

Where:

$r_{xy}$  : Index correlation number "r" product moment

N : Number of cases

$\sum X$  : Total value of score X

$\sum Y$  : Total value of score Y

$\sum XY$  : Multiplication result between score X and Y

Interpretation:

$r_{xy} > t_{\text{table}} = \text{Valid}$

$r_{xy} < t_{\text{table}} = \text{Not Valid}$

Ridwan stated the criteria of interpretation of validity.<sup>14</sup>

0.800-1.000 = Very High Validity

0.600-0.799 = High Validity

0.400-0.599 = Fair Validity

0.200-0.399 = Poor Validity

---

<sup>13</sup> Anas Sudjino, *Pengantar Statistik Pendidikan*, Jakarta: PT Raja Grafindo Persada, 1997. P. 193

<sup>14</sup> Riduwan, *Metode dan Teknik Menyusun Thesis*, Bandung: Alfabeta, 2007 P 110

## G. Research Instruments Reliability

Reliability is a necessary characteristic of any good test for it to be valid at all. A test must first be reliable as a measuring instrument. It is the degree of consistency with which it measures whatever it is measuring.<sup>15</sup> In this study the writer uses reliability in terms of inter-rater reliability. Inter-rater reliability estimates the reliability of two scores which are gained from two tasters for the same subjects of the test.<sup>16</sup> In this study the test was done by two testers, the score was high reliability.

To obtain inter-rater reliability, the score of two raters are correlated using SPSS program. Then the writer gets the interpretation of coefficients correlation, whether they belong to high, moderate, or positive weak negative inter rater reliability category.

Calculation result of  $r$  is compared with  $r_{table}$  by 5% degree of significance with  $df = N-2$ . If  $r$  was higher than  $r_{table}$  so it meant reliable and if  $r$  is lower than  $r_{table}$  so it meant unreliable. In this case, the writer applies the coefficient correlation and the interpretation of inter-rater reliability proposed by Winkle et al as shown in table 3.3.<sup>17</sup> The writer uses Alpha as a formula to measure the reliability of oral test with the criteria;

$r_{11} > r_{table}$  = Reliable

$r_{11} < r_{table}$  = Not Reliable

---

<sup>15</sup> Donald Ary, Lucy Cheser Jacobs, Chir Sorensen, Asghar Razavieh, *Introduction to Research in Education*, Wadsworth: USA, 2010, p. 236.

<sup>16</sup> Riduwan, *Metode dan Teknik Menyusun Thesis*, Bandung: Alfabeta. 2004. P. 229.

<sup>17</sup> Hopkin, CD and Richard, L.A, *Classroom Measurement and Evaluation*. Illinois: F.E. Peacock Publisher, Inc, 1990, p.295.



To know the reliability of the instrument, the value of  $r_{11}$  is interpreted based on the qualification of reliability as follows:<sup>18</sup>

0.800-1.000	= Very High Reliability
0.600-0.799	= High Reliability
0.400-0.599	= Fair Reliability
0.200-0.399	= Poor Reliability
0.000-0.199	= Very Poor Reliability

## **H. Data Collecting Procedures**

In collecting the data the writer was applied six procedures, they are:

1. The writer observed the data:
  - a. number of class
  - b. number of students
2. The writer decided the class into experiment group and control group.
3. The writer gave pretest to experiment group and control group.
4. The writer gave treatment to the experiment group by using team pair solo for three times.
5. The writer taught the control group without using team pair solo for three times, but using technique that used to be used by the teacher, it was direct method.
6. The writer gave posttest to the experimental group and control group.

## **I. Data Analysis Procedures**

---

<sup>18</sup> Riduwan, *Metode dan Teknik Menyusun Thesis*, Bandung: Alfabeta. 2004. P. 113

1. The writer gave score to the students' pre-test result.
2. The writer gave score to the students' post-test result.
3. The writer calculated the mean, median, and modus of students score, standard deviation and standard error of experiment group and control group.
  - a. Mean

$$Mx = \frac{\sum fx}{N}$$

Where:

Mx : mean

Fx : total result product between each score with frequency.

N : number of case

- b. Median

$$Mdn = 1 + \frac{1/2N - fkb}{fi} \times i$$

Where:

Mdn : median

N : number of case

fkb : cumulative frequency located in under interval contain median

fi : authentic frequency (frequency of score contain median)

i : interval class

c. Modus

$$Mo = 1 + \frac{fa}{fa+fb} \times i$$

Where:

Mo : modus

fa : frequency located in above interval contain modus

fb : frequency located in under interval contain modus

i : interval class

d. Standard deviation

$$SD = \sqrt{\frac{\sum FX^2}{N}}$$

Where:

SD : standard deviation

i : interval

N : number of students

e. Standard error

$$Sem = \frac{sd}{\sqrt{n-1}}$$

Where:

Sem : standard error

Sd : standard deviation

N : number of students

4. The writer calculated normality and homogeneity.

a. Normality

It used to know the normality of the data that is going to be analyzed whether both groups have normal distributions or not.

Chi square is used here:<sup>19</sup>

$$X^2 = \sum \left[ \frac{(f_o - f_h)^2}{f_h} \right]$$

Where:

$X^2$  : Chi square

$f_o$  : frequency from observation

$f_h$  : expected frequency

Calculation result of  $x^2$  is compared with  $x$  table by 5% degree of significance.

If  $x^2$  is lower than  $x$  table so the distribution list is normal.

#### b. Homogeneity

It is used to know whether experimental group and control group, that are decided, come from population that has relatively same variant or not. The formula is:<sup>20</sup>

$$F = \frac{\text{Bigger Variant}}{\text{Smaller Variant}}$$

Where:

F : frequency

The hypotheses in homogeneity:

$F_{\text{value}} \leq f_{\text{table}}$ , means both of variants are homogeneity.

---

<sup>19</sup>Sudjana, *Metode Statistika*, Bandung: Tarsito, 1996. P. 273

<sup>20</sup>*Ibid*, p. 280

$F_{\text{value}} \geq f_{\text{table}}$ , both of variants are homogeneity.

If calculation result of F is lower than F table by 5% degree of significance so  $H_0$  is accepted, it is mean both groups have same variant.

5. The writer calculated the data by using t-test the hypothesis of the study, whether the using team pair solo technique gives effect to eight grade students' speaking performance score of SMP Negeri 1 Palangka Raya. To examine the hypothesis, the writer uses t-test formula as follows:

$$t_o = \frac{M1 - M2}{SE_{m1 - m2}}$$

Were:

To : Test observation

M1-M2 : The difference of two means

$SE_{m1 - m2}$  : The standard error of the difference between two means.

To know the hypothesis is accepted or rejected the writer uses the criterion:

If  $T_{\text{test}} \geq t_{\text{table}}$ , it means  $H_a$  is accepted and  $H_0$  is rejected

If  $T_{\text{test}} \leq t_{\text{table}}$ , it means  $H_a$  is rejected and  $H_0$  is accepted

6. The writer interpreted the result of t-test. Previously, the writer account the degrees of freedom (df) with formula:

$$df = (N1 + N2 - 2)$$

Where:

N1 : number of subject experiment group

N2 : number of subject control group

2 : number of variable

The writer uses the level of significant at 5 %. If the result of t-test is higher than t-table, it means  $H_a$  is acceptable but the result of t-test is lower than t-table, it means  $H_o$  is acceptable.<sup>21</sup>

7. The writer made the conclusion of data analysis that obtain.
8. In addition, the writer uses SPSS 21.0 program to compare the data.
9. Discussing and conclusion the result of data analysis.

---

<sup>21</sup>M.Nurcholik, " *The Effectiveness of scrabble in increasing vocabulary achievement toward the students of MTs Darul Ulum of Palangka Raya, 2010*, the state islamic college, p. 22-23