

CHAPTER III

RESEARCH METHODOLOGY

A. Research Design

The design of this research was an experimental research. According to Creswell “experimental is you test an idea (or practice or procedure) to determine whether it influences an outcome or dependent variable”.¹ Quasi-experimental research was applied to this research, a quasi-experimental design is one that looks a bit like an experimental design but lacks the key ingredient, random assignment. In a quasi-experimental design, the researcher did not randomly assign participants to compare groups, because random assignment was not feasible. To improve a quasi-experimental design, the researcher could match the comparison groups on characteristics related to the dependent variable. In this research, researchers’ used pre-test and post-test nonequivalent control group.

This design had several important characteristics. First, the design has pretest and posttest measured for all participants. Second, the design called for two groups, one got some programs or treatment and one which did not (termed the "program" and "comparison" groups respectively). Third, the two groups were nonequivalent; it was expected that they might differ prior to the study. Often, nonequivalent groups are simply two intact groups which are convenient to the researcher, in these cases, the research

¹Jhon W. Creswell, *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, Pearson Education, New Jersey, 2008, p. 299

took two groups as a sample of this study: experimental and control groups. Experimental group means the students who were given treatment by using Bouncing Stories strategy, while the control group was students who were not given the treatment. Before performing the treatment, the author gave some pre-test for both classes. Then the writer gave the treatment for experimental class. At the end of treatment the writer gave a post- test for both of the sample groups.

Table III.1
Research Design

Group	Pre-test	Treatment	Post –test
Experimental	O ₁	X	O ₂
Control	O ₃	-	O ₄

Where :

O₁ : Pre-test for experimental group and control group

X : Receiving particular treatment

- : Without particular treatment

O₂ : Post-test for experimental group and control group

A. Time and the Location of the Research

1. Time of the Research

This research was conducted from July to September 2013.

2. Location of the Research

The location of this research was SMAN 001 Kampar Utara.

B. Subject and the Object of the Research

1. Subject of the Research

The subject of the research was the students of the second year at SMAN 001 Kampar Utara.

2. Object of the Research

Object of the research was the students' writing ability in descriptive paragraph.

C. Population and Sample of the Research

1. Population of the Research

The population of the research was the students of the second year at SMAN 001 Kampar Utara. They consisted of four classes. The number of students was 85 students. It can be seen in the following table.

Table III.2
The Population of the Research

NO	Class	Total population
1	XI IPA 1	22 students
2	XI IPA 2	23 students
3	XI IPS 1	20 students
4	XI IPS 2	20 students
	Total	85students

2. Sample of the Research

The population above is large enough to be taken as sample of the research. Gay and Airisian suggest the researcher can use cluster sampling randomly. Cluster sampling randomly selects group, not individuals. All

the members of selected groups have similar characteristics.² So the writer selected two classes as sample those were XI IPS 1 as experimental class and XI IPS 2 as control class. The samples of this research were 40 students, 20 students for experimental class and 20 students for control class. It can be shown in the table below:

Table III.3
The Sample of the Research

No	Class	Total Population
1	XI IPS 1	20 students
2	XI IPS 2	20 students
	Total	40 students

D. Technique of Collecting Data

The data were collected by giving the students writing assignment in writing descriptive paragraph.

1. Test

The writer collected the data by using test in order to find out the students' ability in writing descriptive paragraph. The writer applied Bouncing Stories strategy to respondents and respondent made descriptive paragraph by using this strategy. The tests were divided into two types.

a. Pre-test

Pre-test was conducted at the beginning of the process. Both control and experimental classes did the pretest. The purpose of

²L.R. Gay, *Educational Research Competencies for Analysis and Application*, Sixth Ed, prentice Hall, New Jersey, 2000, p. 129

this pre-test was to find out the students' writing ability in descriptive paragraph. The result of this pretest was compared to the result of the post test; therefore, the writer could make a judgment based on the students' performance before and after the treatment.

b. Post-test

Post-test was conducted to get the data at the end of the process of implementing the method; it means the data were collected after the treatment of Bouncing Stories strategy implemented.

In order to make the judgment of the students' writing, Heaton theory shown in the following tables. The criteria of writing are content, vocabulary, language use, organization, and mechanic. Then, the typical scale of each component has a set of qualities (level) to be rated in a series of possible rating. Heaton's theories of the rating are as follows³:

Table III.4
Content

Range	Level	Criteria
27 – 30	Very good – Excellent	Knowledge, substantive etc.
22 – 26	Average – Good	Some knowledge of subject, adequate range, etc
17 – 21	Poor – Fair	Limited knowledge of subject, little substance, etc
13 – 16	Very poor	Does not show knowledge of subject, non-substantive, etc

³J.B. Heaton, *Writing English Language Test*, Longman, New York, 1988, p. 132

Table III.5
Organization

Range	Level	Criteria
18 – 20	Very good – Excellent	Fluent expression, ideas clearly stated, etc
14 – 17	Average – Good	Somewhat choppy, loosely organized but main ideas stand out
10 – 13	Poor – Fair	Non fluent, confusing ideas or disconnected
7 – 9	Very poor	Does not communicate, no organization

Table III.6
Vocabulary

Range	Level	Criteria
18 – 20	Excellent to very good	Sophisticated range – effective word/idiom choice and usage –etc
14 – 17	Good to average	Adequate range- occasional errors of word/idiom form, choice, usage but meaning not obscured
10 – 13	Fair to poor	Limited range- frequent errors of word/idiom form, choice usage- etc
7- 9	Very poor	Essentially translation- little knowledge of English vocabulary

Table III.7
Language Use

Range	Level	Criteria
22 – 25	Very good– Excellent	Effective complex construction- etc
19 – 21	Average– Good	Effective but simple construction- etc
11- 17	Poor – Fair	Major problems in simple complex/construction-etc
5 – 10	Very poor	Virtually no mastery of sentence construction rules-etc

**Table III.8
Mechanic**

Range	Level	Criteria
5	Excellent to very good	Demonstrates mastery of construction –etc
4	Good to average	Occasional errors of spelling, punctuation, capitalization-etc
3	Fair to poor	Frequent errors of spelling, punctuation, capitalization-etc
2	Very poor	No mastery of convention dominated by errors of spelling, punctuatuion, capitalization, paragraphing- etc

**Table III.9
Assessment Aspect of Writing Descriptive Paragraph**

No	Aspects Assessed	score				
		0	1	2	3	4
1	Content					
2	Organization					
	a. Identification					
	b. Description					
3	Vocabulary					
4	Grammatical Feature					
	a. Present Tense					
	b. Adjectives					
	c. Adverb					
	d. Specific Participants					
5	Mechanic					
	a. Spelling					
	b. Punctuation					
Total						
Maximum score				20		

Explanation of score:

1. = incompetent
2. = competent enough
3. = competent
4. = very competent

$$\text{Final Score} = \frac{\text{totalscore}}{\text{maximumscore}} \times 80^4$$

Table III.10
Classification of Students Score

Score	Category
80-100	Very good
66-79	Good
56-65	Enough
40-55	Less
30-39	Fail

(Taken from Suharsimi Arikunto)⁵

F. The Validity and the Reliability of the Test

According to Hughes,⁶ a test is valid if it measures accurately what it is intended to measure. In line to Hughes, Gay,⁷ said that validity is the appropriateness of the interpretation made from the test score. Furthermore, Gay says that there are three kinds of validity. They are

⁴ Arthur Huges, *Testing for Language Teaching; Second Teaching*, Cmbridge University Press, United Kingdom, 2003, p. 104

⁵Suharsimi Arikunto, *Dasar-dasar Evaluasi Pendidikan*, Bumi Aksara, Jakarta, 2009, p. 24

⁶Arthur Hughes, *Op. Cit.*, p. 22

⁷L.R. Gay and Peter Airasian, *Educational Research Competencies for Analysis and Application*, 6th Ed, United State of America: Prentice-Hall Inc, 2000, p. 161

content validity, criterion-related validity, and construct validity. All of them have different use and function.

Content Validity is used to compare the content of the test to the domain being measured. Gay also states that there is no formula used in this kind of validity and there is no way how to express it quantitatively.⁸ Content validity just focuses on how well the items represent the intended area. In addition, Hadari Nawawi states that this kind of validity is also said as a curricular validity.⁹ It means that the content of the curriculum of a course must be mastered by the students becomes the standard in determining the validity. To determine the validity using such as validity is by referring to the material given to the students based on the curriculum.

Based on the explanation above, the writer used the content validity to measure whether the test was valid or not in this research. In other words, the tests given to the students were based on the material that they had learned.

According to Gay, “reliability is the degree to which the test consistently measures whatever it is measuring”¹⁰. Furthermore he says that to find the reliability of the test such as essay tests, short-answer tests, performance and product tests, and projective test, we are concerned with inter-judge or intra-judge reliability. The inter-judge reliability is also known as an inter - scorer, inter-rater, or inter-observer reliability.

⁸*Ibid.*,

⁹Hadari Nawawi and M. Martini Hadari, *Instrumen Penelitian Bidang Sosial*, Gajah Mada University Press, Pontianak, 2006, p. 181-182.

¹⁰L.R. Gay and Peter Airasian, *Op. Cit.*, p. 175

In this research, the writer used inter-judge (inter-rater) reliability. It means that the score of the test was evaluated by more than one person. In this research, the students' writing scores were evaluated by two raters.

Table III. 11
Table Categories of Reliability

No	Reliability	Level of Reliability
1	0.0-0.20	Low
2	0.21-0.40	Sufficient
3	0.41-0.70	High
4	0.71-1.0	Very high

(Taken from Tinambunan in Meltiawati in Zelly)¹¹

Table III. 12
Correlation

Correlations			
		rater1	rater2
rater1	Pearson Correlation	1	0.838**
	Sig. (2-tailed)		.000
	N	20	20
rater2	Pearson Correlation	0.838**	1
	Sig. (2-tailed)	.000	
	N	20	20

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

Based on the table above, it can be seen that the inter-rater reliability of this research was 0.838 categorized into very high level. That r calculation is 0.838 correlated to r table, df=38 at level 5% and 1%. Because df=38 was not found from the r table, so the researcher took df=40 to be correlated either at level 5% and 1%. At level 5% at r table 0.304 and at level 1% r table 0.393. Thus, the r table observation is obtained higher than r table, either at level 5% and 1%. So the researcher

¹¹Zelly Putriani, *The Correlation Between Reported Speech Mastery and Speaking Ability of the Students of SMKN 1 Pekanbaru* (Pekanbaru Unplished, 2011), p. 35

concluded there is a significant correlated between score of rater 1 and score of rater 2.

G. The Technique of Analyzing Data

In analyzing the data, the researcher used the post-test score of the experimental group and control group. In order to analyze the students' writing ability in descriptive paragraph, the researcher used passing standard of English lesson at SMAN 001 Kampar Utara (KKM) that was 65 for students' writing ability in descriptive paragraph, it means for those who get score ≤ 65 , they do not pass KKM standard, while for those who get score ≥ 65 they pass KKM standard.

In this research, the data were analyzed by using statistical method, to find whether or not there was significant difference of mean sample from data variable compared¹².

The data were analyzed by using SPSS 16.0 version computer program. After computing t-test, it was necessary to obtain degree of freedom used to determine whether the t-score was significant or not.

The t-obtained value is consulted with the value of t_{table} . Statically, the hypotheses are:

$$H_a: t_o > t_{table}$$

$$H_a: t_o < t_{table}$$

¹²Hartono, *Statistik untuk Penelitian*, Pustaka Belajar, Yogyakarta, 2004, p. 178

H_a is accepted if $t_o > t\text{-table}$ or there is a significant difference on students' writing ability in descriptive paragraph between those who are taught by using Bouncing Stories strategy and those who are not.

H_o is accepted if $t_o < t\text{-table}$ or there is no significant difference on students' writing ability in descriptive paragraph between those who are taught by using Bouncing Stories strategy and those who are not.