

CHAPTER III

METHODOLOGY

This chapter consists of five subchapters, which are: research designs, research subject and research qualification, data collecting technique, instrument validity and reliability, and data analysis method. Research designs contain type of research conducted in this study and the limitations of this study which function to make this study more focused. Then research subject and research qualification explains the subject of this study—which is about the grade, classes and students that were investigated—and how the program was executed with the execution schedule in it. For data collecting technique, here the research informs the ways the data were gathered. The instruments used and the scoring criteria applied in this research are also explained in this subchapter.

In instrument validity and reliability, it is explained how the research attempted to make the instruments used become valid and reliable, so they could be applied in the research. In the last subchapter, which is about data analysis method, the research informs how the data were analyzed, how they are presented in this thesis and how the analysis was done and concluded.

A. Research Designs

This research was conducted in qualitative study. It tried to research how portfolio assessment could make improvement on students' writing performance. The portfolio assessment program itself was applied in two classes chosen from the six classes of Grade VIII at a Public Junior High School.

The data were first taken from pre-test and post-test in the two classes. The other data were taken from the documentation of portfolio assessment results and observation along the program. All the data are quantitative and qualitative. The quantitative data are the scores resulted from the tests and portfolio assessments, while the qualitative data are the words in students' works, correction, feedback, and observation result. All the data were analyzed by using qualitative method to describe the findings and results of this study. After the analysis was done, the research concluded the findings and results.

This research only studies writing portfolio assessment for Grade VIII of Junior High School. The portfolio assessment program of this research, which was applied in two classes of Grade VIII, was done integratedly with the teaching program of the classes as a completion of lesson program, so this study did not burden the English teacher for another extra lesson session. The portfolio assessment itself was focused on writing skill.

The research only took two classes as the object of study, because of the limited time of research scheduled by UT institution. This research was supposed to be conducted in two months, and after that the thesis writing should be done in the following two months.

To make clear how the study was conducted, the following is diagram of the research procedure:

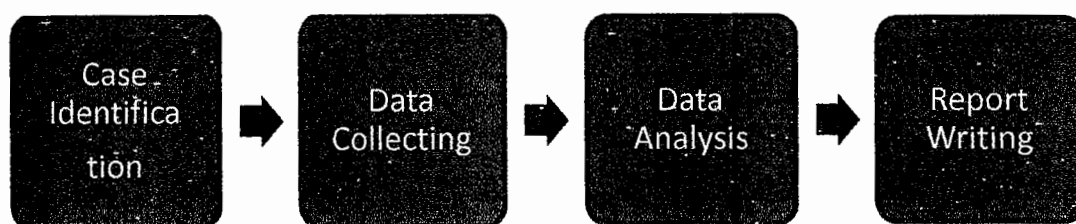


Figure 3
Diagram of Research Procedure

Based on the diagram above, it can be described that this research was started from *case identification*; some findings from the previous researchers were studied to find the gap could be investigated. After that, the gap was developed into research problems with study case research. To learn more about the research problems that could be found in the reality, the research did *data collecting* that it was expected there would be some findings in the investigation. These findings would support the analysis later that was carried on in *data analysis* step. Finally, after the analysis was done, a *report writing* was done to describe and to explain the whole research had been conducted.

B. Research Subject and Research Qualification

In this study, the research took two classes as the object of research and each of the classes had pre-test of writing, portfolio assessment program on writing (8 meetings), and post-test of writing. It means the research took two treatment classes, and there was no treatment and no test for the other classes of Grade VIII in the school. Each of the students' test results and works in the portfolio assessment was documented in a portfolio folder. The school has six classes for Grade VIII, each of which contains around 32 students.

To make the program run well, the research had a meeting with the English teacher of Grade VIII to make a collaboration of executing the program. The research cooperated with her to discuss how she would participate in the program and what she should prepare for the program. After the meeting, the program was executed in the school.

Before the portfolio assessment program was applied, the students were informed about the program first; that they would have pre- and post-tests, and six portfolio assessments. Beside that, they were also informed that this program was integrated with the English lesson program they were learning in the semester in Grade VIII, so what they would do in the portfolio assessments was taught in class by the teacher and the research as well. The teacher, supported by the research, then explained how the program would be held based on the schedule. Every time a portfolio assessment was going to be held, the students would be asked by the research to learn the lesson material related to the portfolio assessment topic going to be done by them, at home, before the portfolio assessment was executed. By doing so, it was expected they would be ready for the portfolio assesment. After a portfolio assessment was done, they were also informed about the feedback that was expected to help them improve their writing performance.

The research program of the two classes was divided into 10 meetings with pre-test and post-test and was held in five weeks. The table of the meetings can be seen in Table 1 (Appendices). The research program was scheduled as seen in the table to make it more organized and to control the research so that the research activities would run as it had been expected and would not be out of the research topic and objectives.

C. Data Collecting Technique

The instruments used in this study were tests, documentation, and observation. The tests were writing pre-test and post-test in the two treatment

classes. The data collected from this instrument were the tests scores of the students and their writing works. The test scores of post-test were compared to the scores of pre-test to find the improvement of students' writing performance in the program.

Then for documentation, the research documented portfolio assessment results in the program as it was scheduled in the previous subchapter. The data gained from this documentation were the scores of portfolio assessment of the students and the qualitative data in words about students' works, correction, and feedback in the program.

For the last instrument, observation, qualitative data were gained through observation of the portfolio assessment program when it was being executed. It was done in every meeting held in the program. During the program activity, the research did observation of the students participated in class. This observation was to know how the students did the test or portfolio assessment, their attitude and reaction toward the research program activity being executed in class.

The research was helped by the English teacher of the two classes to collect the data. The pre-test, post-test, and portfolio assessments were all done by the research with the teacher's help. The research also designed the tests and portfolio assessments, corrected them together with the teacher, so the inter-rater reliability of the instruments could be applied. For the feedback, the research helped the teacher to make written feedback and then later the research informed it to the students.

As it was described in the schedule of the program, the writing tests were done twice (pre-test and post-test), the portfolio assessments were done six times,

and there were twice feedback sessions so the students would know their mistakes in their works and their progress. There were some aspects to be scored in the writing tests and portfolio assessments and later will be explained in the analysis in Chapter IV. Here are the aspects to be scored as the indicators of students' capability (based on writing scoring rubric made by Wang & Liao (2008) in Uçar & Yazıcı, 2016):

Table 2
Aspects of Scoring Adapted from Wang & Liao (2008)

No.	Name of students	Aspects of scoring					Total Score
		Focus	Elaboration/Content	Organization	Convention	Vocabulary	
1.							
2.							
3.							

The research used this scoring table to assess the students' writing work and the aspects in the table represent each aspect of the work. Later the total score of each aspect was averaged wholly for each class, then the research averaged the total score of the two classes (described in Chapter IV).

Here is the explanation of the aspects of scoring above:

- Focus indicates how students wrote based on the topic, not out of the topic.
- Elaboration/content indicates how students used details of writing that match the topic.
- Organization indicates how the work was organized.
- Convention indicates spelling and grammar aspects.

- Vocabulary indicates how students used suitable words in the writing work.

In each of the tests or assessments, students' work was scored by the teacher and the research, then they were given feedback based on the aspects of scoring above so the students' strengths and weaknesses in writing would be known. They were informed about these in detail in feedback sessions.

For collecting the data, the research cooperated with the headmaster (to ask permission to have research in the school) and the English teacher of Grade VIII of the school. The research first explained the research program to them and then planned together the writing pre-test, post-test, and portfolio assessments with them to gather the data. After the data were collected, then a comparison between pre-test and post-test scores would be done by the research. The pre-test and post-test were arranged as follow:

Table 3
Pre-Test and Post-Test Materials

ASPECTS	PRE-TEST	POST-TEST
Kind of test	Writing essay	Writing essay
Topic	<ul style="list-style-type: none"> • Daily activities • The most impressive experience 	<ul style="list-style-type: none"> • Hobbies • Unpleasant experience
Instruction	Students are asked to write their daily activities (30 minutes) and to tell the most impressive experience (30 minutes).	Students are asked to write about their hobbies (30 minutes) and to retell an unpleasant experience (30 minutes).

For the comparison of pre-test and post-test, the research made an improvement checklist. At the end of the program the research should check the students' improvement by filling it as follows:

Table 4
Improvement Checklist

No.	Name	SCORES		Improvement (√)
		Pre-test	Post-test	
1.				
2.				
3.				

The table above would show whether the portfolio assessment program succeeded or not.

D. Instrument Validity and Reliability

The tests in this research consist of eight tests: pre-test, post-test, and six kinds of portfolio assessment. All of the tests and portfolio assessments were designed by the research. To make them valid, the research did instrument validity. It consists of two kinds: content-related validity and construct validity.

Content-related validity means that the tests designed in this research were made based on the English lesson module used by the Grade VIII students as the object studied by the research, therefore the contents of the tests were suitable for them. By utilizing the module, it was expected that they were familiar with the topics.

On the other side, construct validity means the research made a construct in a blue print design for each of the tests and portfolio assessments that contained competency standard, basic competence, indicators, writing test card, and scoring rubric, then all the blue print designs were validated by an expert. Before validation, the research contacted and asked the expert to help the research

validate the tests and portfolio assessments. The expert or the validator was a lecturer of English Education Department of a university with S3 degree in English Education major. He validated all the tests and portfolio assessments used in this research.

All the tests validated were in writing form. For pre-test and post-test, the type of the tests were the same, but the content of each was different. Both pre-test and post-test consisted of two instructions: the first instruction was to make a paragraph which required present tense form; the second instruction was to make a paragraph which required past tense form. For portfolio assessments, the researcher designed a test with only one instruction which required students to write in a paragraph (except in *Making Invitation Card*). Each of the paragraph in all the tests should contain seventy – a hundred words with duration of 20 – 30 minutes depending on the difficulty level of the test.

On the other hand, the scoring rubric that was adapted from Wang and Liao (2008) and modified by the research in the score range and the score criteria description to make the scoring easier and to make the scoring result more understandable for students, since the students should also know their scores and feedback. The short description of the scoring rubric is as follows:

Table 5
Scoring Rubric Adapted from Wang & Liao (2008)

CRITERIA	SCORE	DESCRIPTION
Focus Elaboration/Content	80 – 100	Excellent to Very Good
Organization	70 – 79	Good
Conventions (spelling and grammar aspects)	60 – 69	Fair
Vocabulary	50 – 59	Inadequate
	< 50	Innacceptable

Adapted and modified from Wang & Liao (2008)

For the construct validation, the research attempted the instrument construct to be validated by the expert had been contacted before. In the part of aspects of scoring for instrument validation, the research arranged it with eight criteria and five score levels which can be seen in the following table:



Table 6
Aspects of Scroing for Instrument Validation

No.	TEST PROMPT OF WRITING	SCORE					COMMENT
		1	2	3	4	5	
1.	The instructions are clear in the writing test						
2.	There is an explanation of how the writing test will be assessed						
3.	The instructions in the writing test use simple sentences						
4.	The proportional usage of Indonesian language in the instructions						
5.	The usage of English in the instructions						
6.	The instructions in the writing test are easy to understand by the students						
7.	The grammar aspect in the test script is correct.						
8.	The spelling aspect in the test script is correct.						

Scores meaning:

1 = poor

2 = fair

3 = good

4 = very good

5 = excellent

The validator corrected some mistakes made by the research, then the instruments were revised by the research. After that they were corrected again by the expert until finally there were no more mistakes found and he validated all of them.

After the validation step, the study did instrument reliability. The reliability was done by having inter-rater reliability which means the research involved two correctors/raters (the research and the Grade VIII English teacher) for all the tests and portfolio assessments. The research became the first rater, the English teacher of Grade VIII became the second rater.

Both raters had the same essays to correct. After the correction, the raters discussed the result to see if the rates of each rater were too different between each other or not. In the discussion, the raters also wanted to see if the scoring was too lenient or too strict. By having the same perception between the two, the reliability of the tests would be considered good.

Every time the scoring was finished in each test or assessment, the research made the average scores that were counted from the scores resulted by each rater. The average scores then would be analyzed by the research after the program in the school was done completely.

E. Data Analysis Method

The researcher did qualitative method to analyze the data. After the data were collected, to analyze them, the research did data reduction, data display, and conclusion drawing.

In data reduction, some steps were done; they were summarizing, selecting, focusing, finding pattern, and removing unnecessary data. The research

summarized the scores gained in the tests by making scores average, then selected the scores summaries to make the analysis focused. Beside the scores summaries, there were also words data in forms of students' test answers, correction in words, feedback, and observation data. The research focused on the important data that support the research problems. From the summaries of scores and words data, it could be found the pattern of the research problems; which consisted of the benefit of portfolio assessment, the problems faced by students, and the solutions of the problems. All the data that did not support the analysis of the research problems were removed; for example, the research removed the data of the students who joined pre-test but did not in post-test, because their scores could not be compared and analyzed by then.

After the data reduction, the research then displayed the data. The display is in commentaries, description, diagrams, categories relationships, chart, and tables as seen in this thesis. As an example of data display, the number of students who were involved in the research was listed and arranged in tables for the scores data, based on the writing proficiency level. Then the presentation of the average scores data is described in a bar chart. The example is as follows:

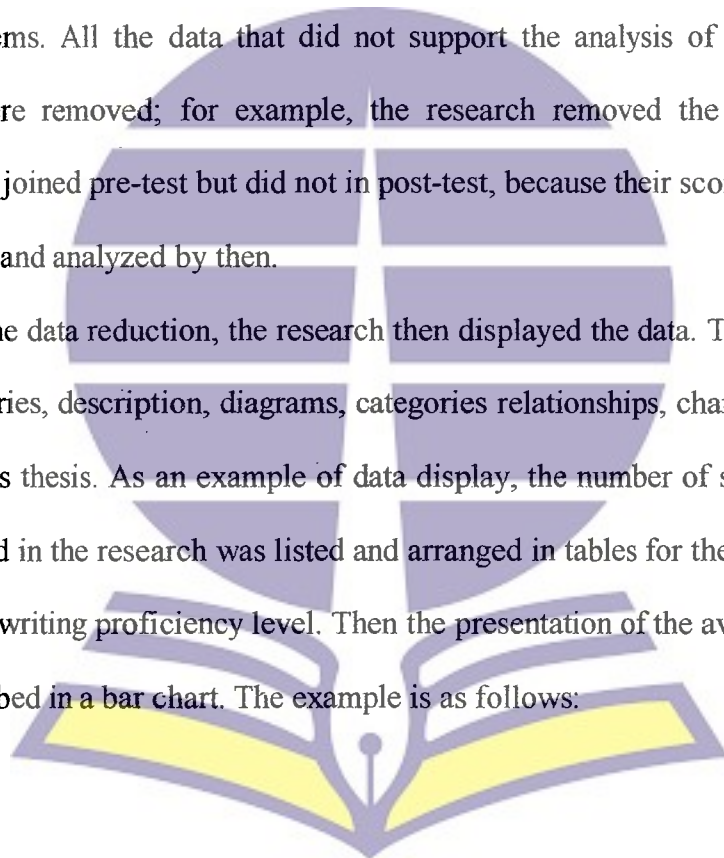


Table 7
Pre-test and Post-test Scores Example

Test	Pre-test	Post-test
Content	84	83
Conventions	80	81
Focus	85	87
Organization	78	78
Vocabulary	75	70

The bar chart of the above data is below:

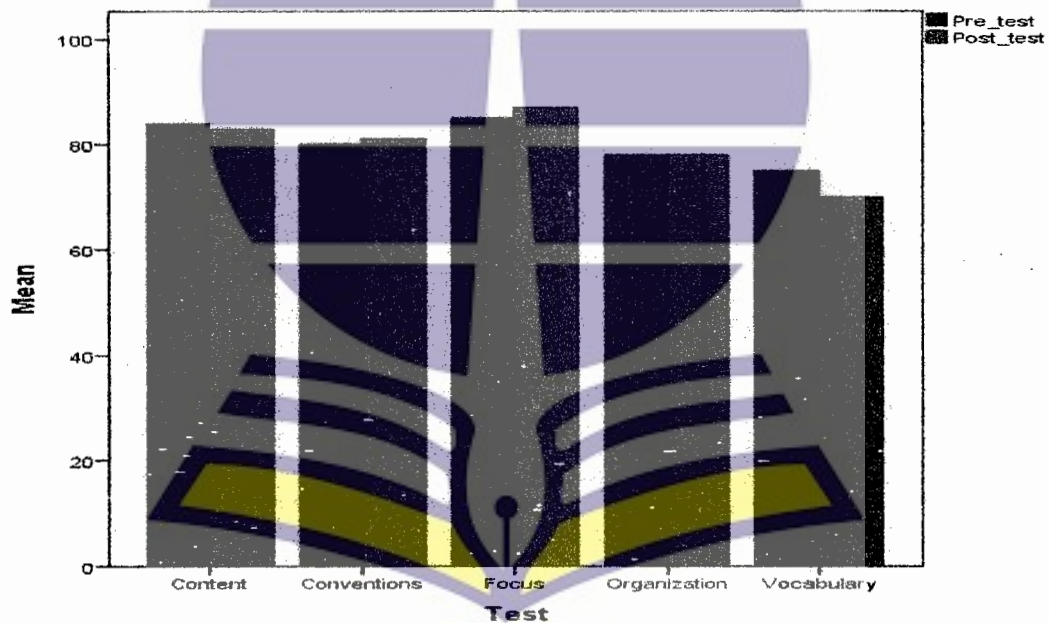


Figure 4
Bar Chart of Pre-test and Post-test

The last step of analysis was drawing conclusion. From the data analysis that was done by referring to the data display, then the research concluded the analysis

based on the research problems, evidences, findings comparison, and relational analysis. By then conclusions were drawn.

