CHAPTER III

RESEARCH METHODOLOGY

This chapter provides a detailed description of the methodology used to answer the research question and reach the purposes of this study. There are eight sections that are elaborated; a restatement of the purposes and the research questions, the research design, the research site, the participants, the data collection, the data analysis, the research procedure and timeline, and the summary.

As it has been mentioned in chapter I, the present study has three purposes. To seek the answers to the research questions, the study intends to investigate the impact of oral and written feedback on Indonesian EFL students' writing, the decrease numbers of errors and students' responses to feedback on writing. With reference to the aforementioned purposes, the present study addresses three research questions as follows:

1. What is the impact of oral and written feedback followed by each two subcategories (direct and indirect feedback) on Indonesian EFL student' writing?
2. Does the feedback provision decrease the numbers of grammatical errors from writing task I to writing task II?
3. What are the students' responses to feedback on writing?

3.1 Research Design

Corresponding to the aforementioned research questions dan purposes, a mixed methods design was employed to meet the research purposes and answer the research questions by combining both qualitative and qualitative method. According to Creswell (2012, p. 535), a mixed methods design can be employed when a research has both quantitative and qualitative data that afford a better apprehension of research problem instead of either type. Creswell also asserts that this design can strengthen both quantitative and qualitative data. Furthermore, this design is needed when either a qualitative or quantitative study is not enough to confront the problem of research. Concerning the mixed methods design, the

present study was categorized as an explanatory sequential mixed methods design or called as a two-phase model (Creswell, 2012, p. 542). The purpose underlining this is that "the researcher uses qualitative data to help explain or build on the initial quantitative results" which the data were analyzed separately in two phases and quantitative data in the first phase were used to plan the qualitative follow-up in the second phase (Malik & Hamied, 2014, p. 278-279). In this design, there is a priority for quantitative data collection and analysis by collecting it first in the sequence and the qualitative data can be used to elaborate the quantitative results through in-depth qualitative exploration in the second phase (Creswell, 2012, p. 542-543). In this study, to meet the answer of the third research question, the results of the questionnaire in the form of numerical data were calculated to find out the frequency of students' responses in the form of a percentage. Further, students' responses were explored through a semi-structured interview to obtain more detail information to support the result of the questionnaire.

Furthermore, the quantitative method was employed to seek the answer to the first and the second research question. This was employed as it covered several characteristics of the quantitative method. Firstly, the use of number dominates the data collection step and the statistical analysis and terminology become the part of the quantitative study (Dornyei, 2007, p. 32; Creswell, 2012, p. 13). Secondly, it deals with a certain focus and condenses information from a large number of particular events to search for a general description or it investigates causal hypotheses (Jolley & Mitchell, 2007) and the results of the study can be generalized (Creswell, 2012, p. 306; Cohen et al., 2007, p. 272). Thirdly, the statistical analysis is used to compare groups or to relate variables and to interpret results "by comparing them with prior prediction and past research". Besides, it is written by using standard, fixed structure and criteria of evaluation (Creswell, 2012, p. 13).

This study could be categorized as a true experiment. Specifically, Creswell (2012, p. 308) asserts that the true experiment equates the groups through random assignment. In addition, in a true experiment, "the researcher assigns the students randomly to the different conditions of the experimental variable" or to groups or different groups and distributes any variability of them

between or among the groups in an experiment. In a random assignment, the students are as the representative of the population. By randomization, the researcher gives control to extraneous characteristics of the participants that could affect the result of the study. Cohen et al. (2007, p. 275) also assert that a true experiment has one or more control and experimental groups and it provides pre-test to "ensure parity" and post-test to discover the effect on the dependent variable. This study involved two experimental groups and one control group of the students in the eleventh grade of senior high school. There were two writing tasks. In each of the tasks, the students were asked to write a recount text. The design of the present study is presented as follows (Hatch & Farhady, 1982; Hatch & Lazaration, 1991)

$$G_1 \text{ (random)} -T_1-X-T_2$$
$$G_2 \text{ (random)} -T_1-0-T_2$$

Where:
    G_1 is experimental group
    G_2 is control group
    T_1 is pre-test
    X is treatment
    T_2 post-test

3.2 Variable

According to Creswell (2012, p. 114), a variable is "a characteristic or attribute of an individual or an organization that a researcher can measure or observe". Further, Creswell asserts that variable is the key to direct a researcher to gather data of study to meet the purposes. Malik and Hamied (2014, p. 73) mention that a variable means "anything that can vary in measure" and there are independent and dependent variables. Independent variables are the variables that can cause effects on outcomes or dependent variables and be manipulated by the researcher. These can be found as factors, treatments, predictors, and determinants. Concerning dependent variables, they are defined as attributes that are influenced by the independent variables. Dependent variables can be measured and observed as the effects or outcomes, for instance, achievement scores or learning, but cannot be manipulated (Creswell, 2012; Malik & Hamied, 2014).

This study classified the variables into both independent variable and dependent variable. Firstly, the independent variable of the present study was feedback from the teacher, which consisted of oral feedback and written feedback. That oral and written feedback were followed by direct feedback and indirect feedback. Thus, feedback from the teacher became the manipulated variable. In other words, oral feedback and written feedback with each subcategory (indirect and direct feedback) became the feedback treatment. Secondly, the dependent variables of this study were students' writing scores from writing tasks and grammatical errors. The dependent variable such as writing grammatical errors was measured through errors existing in drafts and writing scores were measured based on the ESL composition profiles from Hughey (1992 cited in Williams, J. 2005)

3.3 Research site

This study was carried out at one senior high school in Bandung, West Java, Indonesia. The selected research site was based on two reasons. First, the senior high school is where English students are introduced to the texts including recount text. Secondly, accessibility to conduct the study in the targeted school also became the consideration. This consideration is also supported by Silverman (2005) who agreed with this. When choosing a research site, it should have the accessibility of a setting showing such an interesting phenomenon to the researchers, providing precise data expediently and briskly (Silverman, 2005). In other words, the research site was the most fully accessible for conducting the study. The researcher was permitted and supported to conduct this study at the research site.

3.4 Population and Sample

The population of the present study was the eleventh-grade students of one senior high school in Bandung. The eleventh-grade students were chosen by considering two reasons. First, the present study focused on students' writing, specifically in writing a recount text that was taught in the eleventh grade at a senior high school level and the students were assumed to have sufficient knowledge about the topic
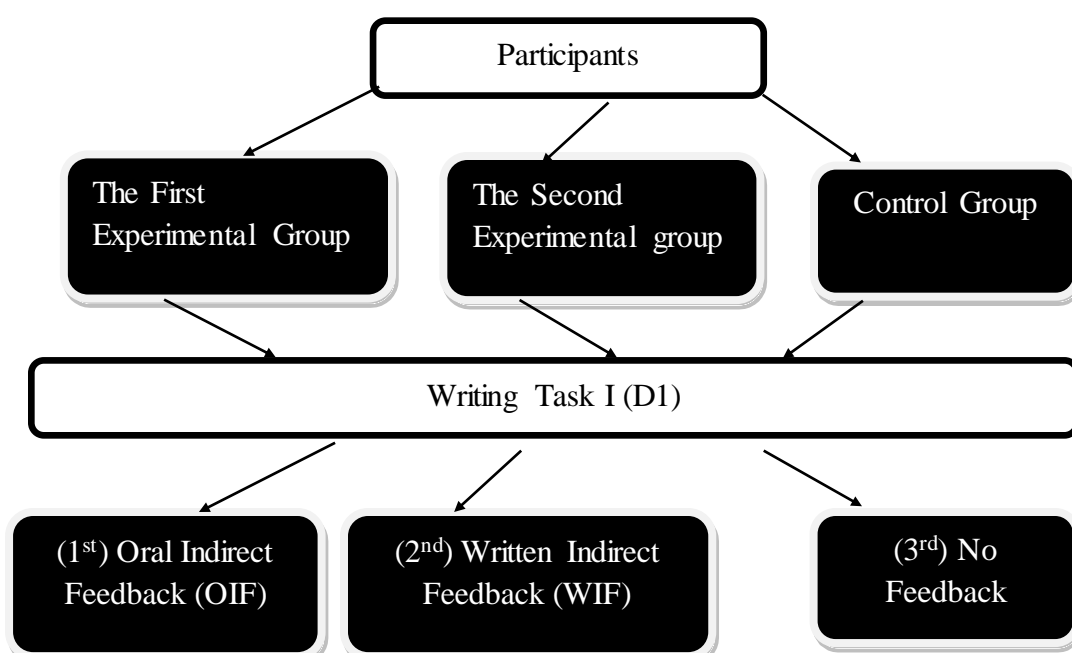
to be able to produce it as they received the lesson of recount text in the previous semester. In this sense, Gebril (2005 cited in Aljarrah, 2016) emphasizes topic familiarity which can determine the quality of the composition product. Secondly, the eleventh-grade students do not focus themselves on the national examination (UN) so that the present study would not disturb their preparation.

In the present study, the teacher-researcher conducted the feedback treatment to six classes as the research sample. In regard to the effectiveness, as there were nine classes, a cluster random sampling was employed. According to Gall, Gall, & Borg (2003), cluster sampling can be used "when it is more feasible to select groups of individuals rather than individuals from a defined population" or to select groups or cluster of subjects rather than individuals (Fraenkel et al., 2011). Therefore, among the classes in the research site, there were six classes as each group contains two classes (two experimental groups and one control group) in this study by employing cluster random sampling regarding the effectiveness when dealing with a larger number of groups (Fraenkel et al., 2011; Gall, et al., 2003). Each class approximately consists of 30 students and the researcher took 23 students from each class as the sample to avoid the absence of the students. The classes in that school in the eleventh grade were equal as the school didn't make any difference of classes. In other words, it was possible to do cluster random sampling. Thus, there were 138 students who participated in the present study as the sample of the population.
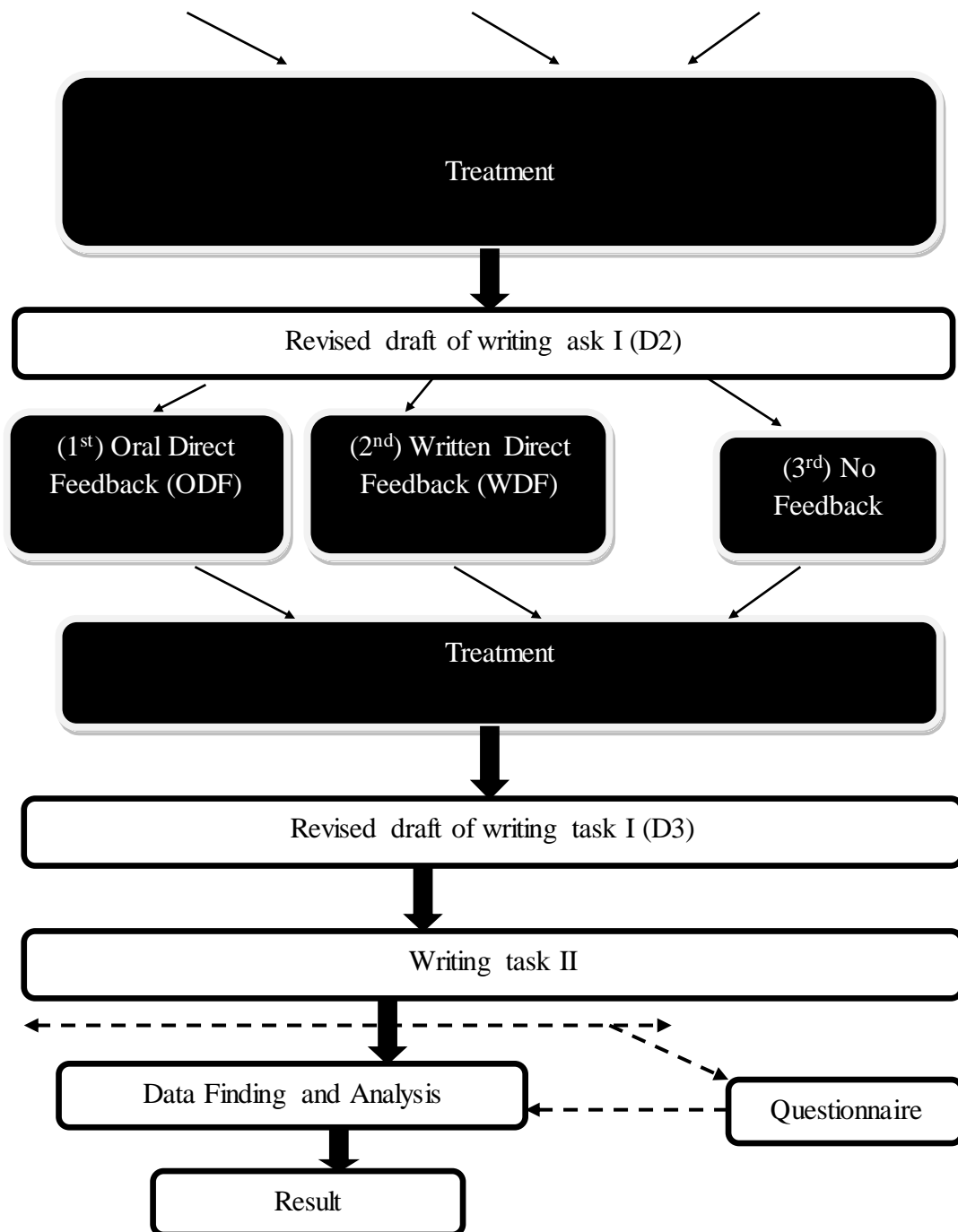
```
                        ┌──────────────┐
                        │ Participants │
                        └──────────────┘
         ┌──────────────────┬──────────────────┐
┌─────────────────┐ ┌─────────────────┐ ┌───────────────┐
│ The First       │ │ The Second      │ │ Control Group │
│ Experimental    │ │ Experimental    │ │               │
│ Group           │ │ group           │ │               │
└─────────────────┘ └─────────────────┘ └───────────────┘
         └──────────────────┴──────────────────┘
                  ┌──────────────────────┐
                  │ Writing Task I (D1)   │
                  └──────────────────────┘
         ┌──────────────────┬──────────────────┐
┌─────────────────┐ ┌─────────────────┐ ┌───────────────┐
│ (1st) Oral      │ │ (2nd) Written   │ │ (3rd) No      │
│ Indirect        │ │ Indirect        │ │ Feedback      │
│ Feedback (OIF)  │ │ Feedback (WIF)  │ │               │
└─────────────────┘ └─────────────────┘ └───────────────┘
```

Figure 3.1 Study Design

3.5 Data Collection Techniques

During the collection data process, there were three instruments used in this study; writing tasks, questionnaire, and interview. The following sections describe the data collection process, including procedures and also the research schedule.

3.5.1 Instrumentation

There were three instruments which were used to collect the data to answer the research questions. These comprised writing tasks, questionnaire, and interview. The forthcoming section explains the detail of each instrument.

3.5.1.1 Writing tasks

The first instrument was the writing task. The writing tasks were intended to collect the data in order to answer the first and second research question. They were used to examine the impacts of the feedback provision which was provided by the teacher-researcher. Furthermore, the writing tasks conducted were in the form of recount text. To measure scores as the result of writing task from the students, an ESL composition profile was utilized in the present study. This ESL composition profile comprises five criteria that are content, organization, vocabulary, language use, and mechanics. The maximum total score is 100 points that are the sum of these five components. There were two independent raters that rated the students' drafts. The first rater has a doctorate in English language teaching and has experienced in years in teaching English. He is a lecturer at one public university in Bandung. The second rater has graduated from the master degree in English education and has experienced in years in teaching English. She is a language instructor in the language center of one public university and as a lecturer in one private university in Bandung. Technically, the two raters did rating the texts independently and they were not told which students belong to which group. The two independent raters rated 552 students' writings.

3.5.1.2 Questionnaire

The questionnaire was used to collect the data to answer the third research question by giving a set of written statements to the students. The questionnaire was only distributed to the experimental groups which were provided with feedback on their writing. It intended to measure the students' responses to feedback. The researcher employed this kind of data collection instrument to effectively use the time. The questionnaire items were in the forms of closed items. According to Nunan and Bailey (2009), a closed item is one in which the

range of possible responses is determined by the researcher and the respondents select from or evaluate the option provided.

The items on the questionnaire were adapted from Rowe and Wood (2008), Najah Asad Al Mohammedi (2016), Seker and Dincer (2014), Sarah (2016), and Ismail (2011). The consideration to adapt the items was due to the similarity of the investigated variables. Besides, it was so related to the basic theory of the questionnaire establishment in the present study. Therefore, the questionnaire was considered to be utilized appropriately as one of the instruments to answer the third research question.

To avoid misunderstanding and give ease to the students when accomplishing questionnaire and understanding the aim of each item, the questionnaire was translated into *Bahasa Indonesia* before administering to the respondents. The questionnaire consisted of 35 close-ended items related to writing, feedback on writing, oral feedback, written feedback, indirect feedback, direct feedback, students' feeling toward feedback and students' preference for feedback. The items in the questionnaire were distributed into some aspects presented in the specified table

Table 3.1 Framework of Questionnaire

No	Categories	Item number
1	writing	1,2,3,4
2	feedback on writing	5,6,7,8
3	written feedback	9,10,11,12
4	oral feedback	13,14,15,16
5	indirect feedback	17,18,19,20
6	direct feedback	21,22,23,24
7	students' feeling toward feedback	25,26,27,28
8	students' preference for feedback	29,30,31,32,33,34,35
	Total	35 items

In the present study, the closed items on the questionnaire were applied to the format of *Likert* scale. *Likert* scale is defined as "a measurement which asks individuals to check their levels of agreement with various statements about an attitude object, i.e. strongly agree, agree, uncertain, disagree and strongly disagree" (Gal et al., 2003). As stated by Linn and Gronlund (1995), it is quite easy to conduct and score in measuring perception or attitudes by using a *Likert*

scale. There are four options for *Likert* scale. The options are Strongly Agree, Agree, Disagree and Strongly Disagree. The questionnaire was prepared in a table of *Likert* scale format. The options were strongly agreed (SA), agree (A), disagree (D) and strongly disagree (SD). The students were required to give a checklist (✓) that was related to the items

No	Item	Strongly agree (SA)	Agree (A)	Disagree (D)	Strongly disagree (SD)

The number of respondents was changed into percentages and converted with the criterion based on the following table

Table 3.2 R% (Percentage of Respondents) Criterion

No	R%	Criterion
1	0	None
2	1-25	Small number
3	26-49	Nearly half of
4	50	Half of
5	51-79	More than half of
6	80-99	Almost all of
7	100	All of

3.5.1.3 The Interview

The interview was conducted to confirm and explore more about students' responses to feedback. It was also conducted to support the findings from the questionnaire. The interview was held only for the students in the experimental groups in the form of a semi-structured interview, specifically four students from each class who agreed to participate voluntarily in the interview.

The questions in the interview were constructed reflected on the items of the questionnaire to obtain and explore detail possible information to support the data from the questionnaire. There were 7 basic aspects for questions (available in the appendix) determined beforehand and followed by other questions which depended upon the responses of students during the interview. The interview was conducted in *Bahasa Indonesia* and English as the interviewees were provided choices to use either *Bahasa Indonesia* or English during the interview so they could feel comfortable. Sixteen students from the experimental groups were

considered enough to represent the numbers of participants in the experimental groups. The interview was conducted after the stages of feedback treatment.

3.5.2 Procedure

3.5.2.1 Conducting Pilot Study

A pilot study had been conducted to assure the instrument and procedure. The pilot study had been done with 15 students in the eleventh grade who did not belong to both experimental and control groups but still in the same population. In the pilot study, there were 5 students receiving oral feedback, 5 students receiving written feedback and 5 students receiving no feedback. In doing so, the students were asked to write a recount text on a given topic from the teacher-researcher. The students were asked to write a paragraph or a maximum 250 words during 30 minutes. As a result, the students' texts collected were analyzed and scored by using the ESL Composition Profile from Hughey (1992 in Williams, J. 2005). The profile categorizes the texts of students into four criteria, very poor (34-46), fair to poor (47-67), good to average (68-85), and excellent to very good (86-100) (see Purnawarman, 2011). Moreover, the grammatical errors were calculated to see that whether or not the feedback provision can decrease the errors in students' writing during the pilot study.

3.5.2.2 Conducting Actual Study

The students were assigned into three groups: two experimental groups and one control group. Each group engaged in writing tasks for paragraph writing. The students were asked to write one paragraph of recount text, their holiday experience, as writing task I approximately 250 words. After doing the writing task, they gave back to the teacher-researcher as their first draft (D1).

In the treatment session, the first experimental group received indirect feedback in the form of oral feedback and the second group received written indirect feedback meanwhile the control group received no feedback. Further, the students were asked to do revision based on feedback provided in their previous draft and collected to the teacher-researcher as their revised draft (D2). For the control group, the students also were asked to revise their texts based on their self-

correction without receiving any feedback from the teacher-researcher. Instead, they were directed at what they were going to do with their draft as the teacher-researcher gave oral instruction. The teacher-researcher analyzed the students' revised drafts (D2) concerning errors in the students' writing and gave to the independent raters to be scored. In the third stage, the first experimental group received direct feedback in the form of oral feedback and the second experimental group received feedback in the form of written direct feedback. The third group received no feedback, but they were also asked to revise their texts based on their self-correction without any feedback provided by the teacher-researcher. The students in the experimental groups and control group were asked to do revision based on feedback provided on the previous draft and collected to the teacher as their revised draft (D3). Lastly, in the fourth stage, the teacher-researcher did not provide any feedback as the students received writing task II as their final draft at the end of the stage of the present study.

More specifically, the first experimental group received feedback in the form of oral differed from the second experimental group. Technically, the teacher-researcher chose five drafts to be involved in delivering oral feedback in front of the classroom as a whole which was followed by a one-to-one interaction as a follow-up activity. The five students' drafts were considered as the representative of the texts possessing the major errors that she needed to present and highlight to cover the errors found. The chosen students' writing drafts which were displayed were unmentioned for the name of the writer or anonymous. Thus, the students were told to focus on the errors displayed through LCD projector and what the teacher-researcher explained concerning those errors, even though those were not their own writing as they were expected to learn. At the beginning of the process of providing oral feedback, she explained the instruction that what they were going to do. More specifically, in the second stage, the teacher-researcher at first started to explain that their writing drafts from the previous meeting have been analyzed for the errors, the copy of their writing would be returned to them and they were going to receive feedback of their writing. She began with one writing and explained what errors were found by only pointing out where those were located without providing the correct form. During one-to-one interaction as

the follow-up activity, they were allowed to discuss and clarify their writing, but not to obtain the correct form of their errors directly. The similar procedure was applied to the other four writing samples. Having finished providing feedback, the students were asked to revise their writing by rewriting not making a new story on the new worksheet given without any correct form as help from the teacher-researcher. While making the revised draft, they were expected to correct and revise with themselves. In this state, While the students were doing the revision, she waited until all have finished. Having finished the revision, they handed over their revised draft to her. Further, in the third stage, the similar procedure was applied, but the teacher–researcher provided the correct form of the errors indicated. During one-to-one interaction as the follow-up activity, they were allowed to discuss and clarify their writing until they got the point of revision. In this state, she highlighted what errors were found by pointing out where those were located and followed by the correct forms so that they only needed to rewrite in the new worksheet as the revised draft. Having finished revising their writing, they collected it to the teacher-researcher.

The treatment session in the present study consisted of four stages. In the first stage, the teacher-researcher conducted the writing task by firstly introducing the task to assure the students having proportional background knowledge on the topic in question. Afterward, they were asked to do writing task I and also told that their writing would be scored in the matter of several aspects based on the rubric and calculated for the grammatical errors. The next meeting, as the second stage, the students received the provision of feedback and were expected to make a revision based on feedback provided in the previous writing. In doing so, all the three groups were asked to rewrite their text by revising all errors addressed by feedback provided by the teacher-researcher. In the third stage, again the students received the feedback provision and they were asked to make a revision based on feedback provided in their previous draft. In the fourth stage, the students received writing task II. They were asked to make one new paragraph writing, another new holiday experience differed from writing task I, still in the form of recount text. They did not receive any feedback and were not asked to make a revision as it was the last stage. Administering the questionnaire was conducted after all stages. The

students from experimental groups were told to fill out the questionnaire about their responses to feedback on writing. The questionnaire was administered to the experimental groups only or 92 students. It was followed by the interview to obtain more information.

The present study collected the students' drafts of writing task I, revised drafts until writing task II. All the writing tasks, 552 pieces of the students' writing in the total number, produced by the 138 students were collected from the experimental groups and the control group.

The following table provides the schedule of the study, which depicted the study conducting from July to September

Table 3.3 The Schedule of the Study

Step	Group	Time
Pilot Study		July-August 2017
• Writing task I		26 July 2017
• Feedback I		27 July 2017
• Feedback II	pilot group	1st August 2017
• Writing task II		2nd August 2017
Writing task I (D1)	experimental group 1	3rd &8 August 2017
Writing task I (D1)	experimental group 2	3rd & 4 August 2017
Writing task I (D1)	control group	4 & 8 August 2017
Feedback I (Revision/ D2)	experimental group 1	10 & 15 August 2017
Feedback I (Revision/ D2)	experimental group 2	7 & 10 August 2017
Feedback I (Revision/D2)	control group	11 & 28 August 2017
Feedback II (Revision/D3)	experimental group 1	24 & 29 August 2017
Feedback II (Revision/D3)	experimental group 2	14 & 24 August 2017
Feedback II (Revision/D3)	control group	25Aug.& 17 Sept.2017
Writing task II	experimental group 1	31 Aug. & 5 Sept. 2017
Writing task II	experimental group 2	7 & 11 Sept. 2017
Writing task II	control group	8 & 17 Sept. 2017
Questionnaire and Interview		31 August, 5,7,8,11, & 17 September 2017

3.6 Data Analysis Techniques

3.6.1 Pilot Study Data Analysis

Prior to conducting the research, the researcher did a pilot study. This pilot study sought to test the reliability and validity of the instrument and procedure used in this study. When the pilot study has been finished, the results of students' writing drafts were scored based on the ESL Composition Profile by two independent raters. Further, the grammatical errors were calculated to find out the reducing numbers of errors indicating the grammatical accuracy of students' writing as the result of pilot procedural and instruction. Having conducted the pilot study, some constructive information and temporary conclusion were considered.

3.6.2 Actual Study Data Analysis

The writing drafts written by the students in all three groups from each stage in the present study were scored based on the criteria of Scoring Guidelines for ESL Composition (Hughey, 1992 in Williams, J. 2005) by two independent raters. The scoring guideline comprises several criteria; content, organization, vocabulary, language use and mechanics. The score of each criterion has a different range. First, content ranges from 30 as the maximum score to 13 as the minimum score. Second, organization ranges from 20 as excellent to 7 as very poor. Third, vocabulary ranges from 20 as excellent to 7 as very poor. Fourth, language use ranges from 25 as excellent to 10 very poor. At last, mechanics range from 5 as excellent to 2 as very poor. The scoring guidelines can be seen in the appendix.

The scores were derived from two independent raters. Prior to doing the rating, both independent raters were given a clear and detailed explanation about the rating scale guideline. When writing drafts were obtained from students, those drafts were immediately given to the raters. Technically, when students were done, for instance, with writing task I, they collected their draft to the teacher-researcher and those drafts thereafter were handed to the raters to be scored. The similar step was applied to each stage until the last stages of the present study. In doing so, the researcher had several considerations. First, the researcher believed that handing over students' drafts for each stage would not be too demanding for the raters since there were 138 drafts in one stage and there were four stages.

Secondly, writing scores as the data would be obtained along with the stages accomplished in the study. Thus, it was expected to procure those data as scheduled. At last, the raters were expected to have sufficient time to rate the drafts by giving a pack of drafts for each stage instead of the whole total drafts when accomplished from all stages to the raters. Additionally, during the rating process, students' writing drafts were not mentioned in which group they belonged to in order the raters could maintain the objectivity of the rating process or avoid any intervention.

Moreover, the independent raters were the persons in charge to maintain the essence of equality in rating students' writing. The raters considered to pick-and-drop the drafts when rating. In other words, the raters spent adequately much time to rate the drafts aiming to keep the equivalence of the score. Furthermore, it was necessary to ensure the reliability of two raters on rating students' writing drafts. To assure it, an inter-rater reliability test was performed. This inter-rater reliability test was done first before conducting further statistical computation. The Pearson product moment correlation was performed to estimate the reliability of the students' scores. The inter-rater reliability coefficient obtained for the scores of the students' drafts and indicated positive correlation. In other words, the scores produced by two independent raters were consistent (not significantly different).

Further, the students' writing scores were computed and analyzed through t-test and analysis of variance (ANOVA). The t-test is "an excellent statistical procedure to use in comparing two means" (Hatch & Farhady, 1982, p. 119). To compare the groups with themselves on two occasions and to find out whether or not the groups made any significant changes from each writing task, paired sample t-test was performed. Further, in this study, to find out the writing of students differed as a result of feedback when compared to no feedback, one-way ANOVA was conducted. The analysis of variance (ANOVA), according to Gall et al., (2003, p. 307), refers to "a statistical procedure that compares the amount of between-groups variance in individuals' scores with the amount of within-groups variance". Other than aforementioned statistical computations, two post-hoc tests were conducted when the value obtained from one-way ANOVA indicated

differences in the mean scores of groups. This study utilized Tukey post-hoc test and Games-Howell post-hoc test to compare the mean scores of the groups to reveal that the significant difference lay.

3.6.2.1 Writing task I, Revised Drafts and Writing task II data analysis

Writing task I and writing task II were given to the experimental and control group. The present study employed a hypothesis with alpha 0.05 level. The data collected from writing task I, revised drafts and writing task II were calculated and compared using SPSS program for Windows. Subsequently, the result was utilized to uncover the impact of the feedback provision on students' writing. In other words, the researcher intended to find a causative relationship between the independent variable and the dependent variable.

Some assumptions needed to be fulfilled in analyzing the data before conducting further statistical computation. Those were the normality of distribution test and the homogeneity of the data

Firstly, the normal distribution test was performed to investigate whether or not the distribution of pre-test and post-test scores in groups were normally distributed. Kolmogorov-Smirnov test formula in SPSS program was used to analyze the normality of distribution. The steps were the following. First, stating the hypothesis and setting the alpha level at 0.05 (two-tailed)

H_0: the score of the experimental group and the control group are normally distributed

H_1: the score of the experimental group and the control group are not normally distributed

Further, analyzing the normality distribution using Kolmogorov-Smirnov test formula in SPSS program. Next, comparing the Asymp Sig. (probability) with the level of significance to test the hypothesis. If the Asymp. Sig was more than the level significance (0.05), the null hypothesis was accepted and the data were normally distributed.

Secondly, the variance homogeneity test was conducted to examine whether or not the score was homogenous variance. This statistical calculation employed ANOVA Lavene test formula in SPSS program. The steps were the

following. First, stating the hypothesis and setting the alpha level at 0.05 (two-tailed)

H_0: the score of the experimental group and the control group are homogenous

H_1: the score of the experimental group and the control group are not homogenous

Further, analyzing the normality distribution using Lavene formula in SPSS program. Next, comparing the Asymp Sig. (probability) with the level of significance to test the hypothesis. If the Asymp. Sig was more than the level significance (0.05), the null hypothesis was accepted and the data were homogenous.

In order to find the changes of students' writing as the improvement, paired sample t-test was performed. Paired sample t-test was performed to find the differences between the previous drafts and the drafts after the feedback treatment in each group. The steps were the following. First, stating the hypothesis and setting the alpha level at 0.05 (two-tailed)

H_0: there is no significant difference between students' writing score in pre-test and post-test

H_1: there is a significant difference between students' writing score in pre-test and post-test

Further, calculating T-test score using SPSS program. Next, comparing t-obtained and t-critical. If t-obtained > t-critical, it means that the hypothesis was rejected, there is a significant difference between the scores before and after treatment. In contrast, If t-obtained < t-critical, it means that the hypothesis was not rejected; there is no significant difference between the score before and after treatment.

3.6.3 Questionnaire Data Analysis

The data obtained from the questionnaire were analyzed quantitatively by using descriptive statistics in terms of mean, frequency and percentage to know the students' perception of feedback on their writing. To analyze the result of the questionnaire, the scores were in the form of *Likert* scale that measures the extent to which a student agreed or disagreed with the items. The score was 1 to 5. The

scale was 1 as "strongly agree", 2 as "disagree", 3 as" agree", and 4 as "strongly agree (Dornyei, 2010; Creswell, 2008). For both positive and negative statements in the questionnaire were scored as follows

	SD	D	A	SA
negative statement	4	3	2	1
positive statement	1	2	3	4

The data from the questionnaire were analyzed and interpreted based on the frequency (f_o) of the answers from students. Thus, the percentage formula used to analyze the data collected from the questionnaire is following

$$P = \frac{\text{Total respondent answering an item } (f_o)}{\text{Total respondent } (n)} \times 100\%$$

P = Percentage (%)

f_o = Total respondents who answer an item

n = Total respondents

(Sudjana, 1984)

3.6.4 Interview Data Analysis

The analysis the data from the interview was done through some steps. Firstly, transcribing the interview data from students' interviews after obtaining and previously reading the data. Secondly, the data were categorized to make comparison between the aspects to develop theoretical concepts. In this step, it was done also by matching the data from different participants and associating to the third research question. Thirdly, similar categories were grouped and sorted to each aspect corresponding with the aspects stated in the questionnaire. Fourthly, interpreting was employed to understand the data to get findings in order to answer the third research question as mentioned earlier in chapter one.

3.7 Validity and Reliability

The instruments should go through the validity and reliability test to assure the validity and reliability. Prior to administering the questionnaire to the research sample, for content validity, the questionnaire items were initially reviewed each item by a lecturer from one public university who have a doctorate degree and years of teaching experience in *Pendidikan Bahasa Indonesia* study since the items in the questionnaire were all translated into *Bahasa Indonesia*. The consideration to translate the items into *Bahasa Indonesia* was that students were expected to easily comprehend the items or statements. Likewise, the interview questions were also reviewed by the same lecturer prior to conducting the interview.

The questionnaire was administered to students in the pilot study. There were 36 students in the pilot study, which were not involved in the research sample, but they belonged to the same population that were asked to fulfill the questionnaire. To calculate the validity and reliability of the questionnaire with *Likert* scale, the result of testing was calculated by Alpha-formula and SPSS 20 for Windows 8 program was run. Previously, there were 43 items on the questionnaire. The score of validity for each item is r_{count} that can be known from the corrected item-total correlation table from the SPSS data output. Each item is valid if $r_{count} > r_{table}$. The questionnaire was administered to 36 respondents with the level of significant 5% and the r_{table} is 0.329. The result of computation revealed 8 invalid items as the value of r_{count} is lower than r_{table}. Therefore, there were only 35 items that would be utilized for the actual study.

In regard to the reliability, the calculation of Alpha Cronbach was used to analyze the reliability of the instrument. The Alpha Cronbach of the computation result of the questionnaire was 0.735 A research instrument has high reliability if the coefficient of Alpha Cronbach $\geq 0, 6$. Since the Alpha Cronbach of the questionnaire was higher than 0.6, it proved that the construction of questionnaire was reliable.

3.8 Concluding Remarks

This chapter has presented the methodology used in the present study. The function of this chapter is as a guide to determine the research design, research site, population and sample, data collection technique, instrumentation, and data analysis executed in the present study. Research finding and discussion are provided in the forthcoming chapter.