CHAPTER III
RESEARCH METHODOLOGY

This chapter describes the methodological framework of this research. The methodological framework is presented in order to provide an understanding regarding the process of conducting the research. This methodological description covers seven subthemes. The subthemes include the research design, the research sites, population, and samples, data collection, data instrument, data procedure and data analysis.

3.1 Research Design

Research design is described as the work plan of a research as it tells how the research is conducted in order to answer the formulated research questions (Fraenkel, Wallen & Hyun, 2012; see also Creswell, 2012). In this case, there are three formulated research questions that define the overall the design of the research. As stated in Chapter I, the research questions are formulated, firstly, to reveal the profile of learner autonomy of the population, secondly to investigate the effectiveness of reading portfolio in developing language learner autonomy, and lastly, to show the students' responses regarding the use of reading portfolio in developing learner autonomy. Therefore, referring to the nature of the research questions, then, a mixed method design which combines both quantitative and qualitative data were used (Creswell, 2012; Malik & Hamied, 2016).

Mixed method is defined as a design that involves quantitative and qualitative research techniques, methods and approaches in a single study (Malik & Hamied, 2016). This combination permits to offer insight that could not be provided by a single method (Creswell, 2012). Accordingly, in this research the data were gathered by quantitatively employing inventory tests (pretest and posttest) (Campbell & Stanley, 1963; Hatch & Farhady, 1982; Heigham & Croker, 2009) and by qualitatively using interviews (King & Horrocks, 2012). Following the nature of the research questions, the inventory tests were used in order to answer the first two research questions meanwhile the interview was conducted in order to answer the last research question respectively (Campbell & Stanley, 1963; Hatch & Farhady, 1982; King & Horrocks, 2012). This also indicates that the quantitative data was collected first which was then followed by the collection of the qualitative data (Malik & Hamied, 2016).

To be more precise, the quantitative method used in this study was the *Nonequivalent Pretest-Posttests Control Group Design* (Campbell & Stanley, 1963; Hatch & Farhady, 1982; McMillan & Schumacher, 2001). The selected research method required the researcher to administer inventory tests in a form of pretest to the population of the study and posttest to two groups of research samples, namely control and experiment groups and an intervention. The test administered to the whole population was intended to reveal the profile of learner autonomy in general and in each aspect. Meanwhile, the test on the selected two groups of research sample was aimed in order to give justification on the comparative causal effect of the intervention or the independent variable (reading portfolio) on the dependent variable (learner autonomy) (Hatch & Farhady, 1982, McMillan & Schumacher, 2001, p. 342).

In addition, the interview was conducted as the qualitative method of data collection (Emilia, 2009; King & Horrocks, 2012). The interview was administered once the programs were completed as it was intended to elicit students' responses about the effect of portfolio on the development of their autonomy especially in reading practice. Qualitative data collection such as interview was selected as it allows for answering the questions on how the subjects of the research behave, feel, and react as they experienced the phenomenon (Heigham & Croker, 2009). Consequently, it was expected that interview could contribute to develop a more in depth understanding of how the experimental intervention, in this case reading portfolio, actually worked to foster the development of learner autonomy (King & Horrocks, 2012). The schema of the research design is as follow.

Figure 3.1 Schema of the Research Design

Pretest	Respondent	Experiment	X+	Posttest	Interview
		Control	X		-
		Phase 1			Phase 2

Time →

Based on the schema of the research design above the test was administered to all of the population so that the general picture of the profile of learner autonomy can be revealed. After that, the respondents (R) were selected from the population either for control group (A) or experimental group (B) (Hatch & Farhady, 1982). The selection of the respondents for experiment and control group in this study did not use the *random assignment* technique, instead, the purposive sampling using *intact group technique* was used (Furqon & Emilia, 2010). This technique signals that the grouping of the respondents followed the grouping as it is formed in the population of the research (Furqon & Emilia, 2010, p. 20). Therefore, a group of sample such as a class of student could be made either as experiment (B) or as control (A) group. This technique can be considered as relatively common to be used in educational setting as often time the researchers cannot artificially create groups for the purpose of experiment only (Furqon & Emilia, 2010).

As the name applies, in *Nonequivalent Pretest-Posttests Control Group Design*, the intervention or treatment (X+), in this case reading portfolio, was given to the experimental group which also signified that the control group was not provided with additional treatment hence it is called X (Campbell & Stanley, 1963; Hatch & Farhady, 1982, Furqon & Emilia, 2010, p. 3). Meanwhile, in order to gain the data needed, pretest and posttest were administered to control and experimental groups. The tests were aimed in order to find out whether the treatment or the independent variable causes a significant difference in regards to the autonomy development between the two groups. The differences themselves were reflected in the pattern of dependent variable specifically in between O_2 and O_4 (Campbell & Stanley, 1963).

Finally, in order to gain general picture of students' responses about the effect of portfolio on their level of autonomy itself qualitative design through interview was given to the students from the experimental group. The interviews were conducted only to the subjects from experimental group so that it is expected that their perception on how the portfolio contributed to the development of learner autonomy itself can be used for the future refinement or betterment of the reading portfolio framework (Little, 2010; King & Horrocks, 2012).

3.2 Setting, Population, and Sample of the Study

The designs of the research required the researcher to find setting, population and sample study so that test, classroom intervention, and investigation of the phenomenon being studied could be carried out (Creswell, 2012). There were several considerations taken in selecting research setting, population and subject of the research. Accordingly, it will be presented as follow.

3.2.1 Setting of the Study

This study was carried out at a vocational high school in West Java. The justification of selecting vocational high school context as the setting of the study was based on Little's (2010) theory on learner autonomy. Little, (2010) states that there is a need to developing learner autonomy in all educational settings including vocational high school level as it is the key to creating intellectual learners who will be ready to face the challenge of fast changing world as well as to reflect on their own country's development. This theory is also in line with the goal of Indonesian curriculum that requires balance competencies of the learners both in terms cognitive and attitudinal aspects as stipulated in educational policy number 65 and 66 year 2013 (*Peraturan Menteri Pendidikan dan Kebudayaan*

Republik Indonesia, 2013). Thus, the setting was concentrated on vocational high school level so that it is expected that the result of the research can be generalized to other similar educational settings (Creswell, 2012).

In addition, the location of the school in which the research was conducted was in West Java. The consideration of selecting West Java as the setting of the research was based on the feasibility and suitability in terms of time, mobility and skill (McMillan and Schumacher, 2001, p. 432; Creswell, 2012). By this it means, that the setting of the research allowed the researcher to be able to manage better in terms of time and cost as it was located in the approachable region to that of researcher's own region. Furthermore, the researcher also had a good access, permission, and support from the school in which the research was held. It also means that the classroom intervention and investigation could be conducted without concerning any acquaintanceship problem so that the trustworthiness of the research could be maintained (McMillan and Schumacher, 2001; Fraekel, Wallen & Hyun, 2007; Creswell, 2012).

3.2.2 Population of the Study

Population is defined as a group to whom the results of the study can be generalized (Fraekel, Wallen & Hyun, 2007; Creswell, 2012). The population of this study was the tenth grade of English language learners in vocational high school level. The selected population enrolled English subject once a week in a 2 x 45 minutes duration per meeting. As the subjects have studied English for more than four years, thus, they were expected to be able to represent distinguishing features of foreign language learners (Little, 1991). The total population of the research can be seen in the following table.

Table 3.1
Population of the Research Study

No.	Class	∑
1	X-A	24
2	X-B	35

3	X-C	25
4	X-D	39
5	X-E	38
6	X-F	32
7	X-G	29
8	X-H	36
9	X-I	39
10	X-J	40
11	X-K	38
12	X-L	39
13	X-M	30
Jumlah		**440**

3.2.3 Samples of the Study

From the total population, representative samples were assigned purposively to the experimental and control group (Furqon & Emilia, 2010, p. 20; Creswell, 2012). The purposive sampling technique was used in order to be able to represent the common characteristic of English language learners in vocational high school level (Campbell & Stanley, 1963). Furthermore, the selection of samples using purposive sampling technique also allowed to assigned research samples that could best represent the population so that adequate information to estimate the total population can be gained (Furqon & Emilia, 2010, p. 20). The same technique was also applied for the purpose of selecting respondents of the interview (Furqon & Emilia, 2010; King & Horrocks, 2012).

As the research design used in order to answer the first two research questions is quantitative in nature, thus, the size of the groups has to be taken into account especially when determining the amount of the test takers and at the stage of selecting the research samples (Campbell & Stanley, 1963; and Creswell, 2012). For the purpose of revealing profile of learner autonomy, then the whole population of the study, i.e. 440 students were involved in the initial inventory test (pretest). Whereas, for the purpose of conducting classroom intervention two classes students were also selected. It is mentioned that the general rule of thumb of determining the sample size for the purpose of experimental research using inferential t-test is that the sample size should at least cover the minimum requirement of 15 samples per-group (Hatch & Farhady, 1982; and Creswell, 2012). From the table of population listed above, all of the intact groups in the population were consisted of more than 25 samples, therefore, it can be concluded that any of the listed groups/classes could be selected as the samples of the study.

In addition, it is also stated that for the purpose of quantitative research, homogeneity of the samples of the study is also critical as it may affect the internal validity of the research itself (Campbell & Stanley, 1963). In this case, Campbell and Stanley (1963, p. 47-48) states that "..the more similar the experimental and control group in their recruitment and the more this similarity is confirmed by the score of pretest," then, the more valid the interpretation of the research becomes. Moreover, it is also stated that the selected sample should not only homogeny in terms of autonomy level but also should be those who are considered as having low degree of autonomy (Campbell & Stanley, 1963, p. 50; Little, 2010). Accordingly, in order to maintain the homogeneity or similarity of the samples study then pretest on learner autonomy was administered to all of the population of the research so that two intact groups that relatively had the same level of autonomy could be taken as the samples of the research. The general representation of the initial autonomy development of the population will be showed in the following table.

Table 3.2
Initial Autonomy Development of the Population

Number	Class	∑	Degree of Autonomy
1	X-A	24	36,51

2	X-B	35	34,76
3	X-C	25	33,45
4	X-D	39	34,87
5	X-E	38	36,84
6	X-F	32	32,86
7	X-G	29	37,01
8	X-H	36	33,53
9	X-I	39	33,38
10	X-J	40	45,15
11	X-K	38	45,20
12	X-L	39	42,04
13	X-M	30	33,00
Total/Average		**440**	**35,9**

From the total groups listed above, two groups, X-F and X-M were taken as the sample of the study. These groups were distributed into two categories namely experimental and control group (Campbell & Stanley, 1963). Class X-F, as the first group, was assigned as the experimental group, meaning that this group received reading portfolio as the classroom treatment, meanwhile class X-M as the second group, was assigned as control group. These groups (X-F and X-M) were selected since the degree of initial autonomy development of the students in these classes could be considered as equal (in approximately 32%). Furthermore, these groups also had a relatively lower level of initial autonomy development compared to other intact groups. The similarities of autonomy development of the sample study were maintained in order to control the validity of the research (Campbell & Stanley, 1963; Hatch & Farhady, 1986; Creswell, 2012). The distribution of samples in each group based on the rationale given above will be depicted in the following table.

Table 3.3

Samples of the Research Study

No	Class	Status	Degree of Autonomy	Number of Students
1	X-F	Experiment	32,86	32
2	X-M	Control	33,00	30
		Total		62

3.3 Materials and Teaching Procedures

3.3.1 Materials

The material that was used throughout the intervention process to the experimental group was the reading portfolio (O'Malley & Pierce, 1996; Kose, 2008; Little, 2009; Little & Perclova, 2010; see Appendix B for the rationale and framework of the reading portfolio). Based on the principles of promotion of learner autonomy, then the reading portfolio implemented during the intervention was consisted of three components, namely; inventory reading task, reflection log, and showcase (Appendix B). There were several texts included in the inventory reading task. The texts consisted of two genres; advertisement and explanation texts. The students were required to bring examples of advertisement and explanation texts as it was intended to challenge students' involvement in reading. In addition, the teacher's ready-made materials (reading texts and practices) were also provided. The texts were selected from several sources (Appendix B)

3.3.2 Teaching Procedures

Once the instrument for the classroom intervention had been adapted and completed, then the next stage of the research study was the implementation of reading portfolio (see Appendix B). At this stage the researcher took a role as teacher-researcher. This indicates that the researcher gave the intervention to the research subjects by herself (see section 3.7.4). The role as teacher-researcher was employed as it allowed the researcher to get more reliable and valid data that were needed

(Creswell, 2012). The implementation of reading portfolio in the experimental group followed the procedures as arranged below (see Appendix B).

Table 3.4
Teaching Procedures

Week	Classwork	Week	Classwork
1	Pretest	6	Task 4
2	Introduction to and training about the reading portfolio and its platform	7	Conference of reading portfolio/ Show case
3	Task 1	8	Posttest
4	Task 2	9	Evaluation of the teaching procedure
5	Task 3	10	Closing

Materials, teaching procedures, the implementation of the teaching procedures and evaluation process will be further depicted in Appendix B.

3.4 Research Instruments

3.4.1 Research Instruments

The data of this study were collected through quantitatively applying inventory test, pretest, and posttest and qualitatively applying interview. The construct of these research instruments will be explained as follow

3.4.1.1 Inventory tests: Pretest and Posttest

As indicated in the research design, in order to gather the data about the profile and the development of learner autonomy, i.e. the first and second research questions, then a pretest and a posttest were administered. Both the pretest and posttest in this research were in a form of questionnaire that consisted of attitudinal measures (Azwar, 2010, p. 105-121; Creswell, 2012). This form of test (questionnaires) and measurement (attitudinal measures) were selected as both of the forms were considered as one of the popular forms of gathering quantitative data that were related to revealing the attributes of the sample study (Azwar, 2011, p. 105; Creswell, 2012, p. 151-155). In this case, the popularity of questionnaires that consisted of attitudinal measures relied in that the researcher can both maintain practicality in terms of administration and achieve reasonable reliability in terms of data collection results (Azwar, 2010; Azwar, 2011, p. 105).

The categories used as a basis for developing specification of the instrument were adapted from Loan and Tin (2016). This form was selected as it provided comprehensive, detailed, and systematic system in linking the three pedagogical principles for learner autonomy development (Little, 1991; 1995; 1999; 2010). Therefore, the questionnaire was consisted of three main dimensions (Azwar, 2011, p 110), namely Learner Involvement, Increase English Use, and Learner reflection (Loan & Tin, 2016, see also Little, 1991; Little, 1995). Although there were several classifications in terms of dimension of autonomy, yet, the distinction between the dimensions themselves is mainly theoretical in nature since in the learning and teaching processes all these aspects are closely interrelated.

In addition, as this present study was focused on the learner autonomy in reading practice as supposed to Loan and Tin (2016) which was centred on language skills, consequently, modification was conducted. Here, the differences themselves were mainly reflected in the operational subscale of the instruments. In this regard, the modifications and the details are as follows:

- The items for Learner Involvement and Target Language Use from the original version were modified into 21 items focusing on the learner autonomy in reading practice.
- The items in “Learner Reflection” were newly developed as they were expected to best capture the effort of reflection in learning conducted

by high school level students as supposed to the original ones which were linguistically considered as more complex in nature as (see Appendix A).

In addition, the test was developed in a form of multiple items measure with *Likert-type* scale (Azwar, 2011, p 105). The selection of 5-point *Likert-type* scale was based on the consideration that the scale was widely accepted as an average interval level of measurement which is in line with common practice in behavioural and educational research (Creswell, 2012). In addition, scaled descriptor was also preferred as it enabled to generate the general picture of autonomy degree/level of the samples study. As the name applies, the *5-point Likert-type* scale was consisted of five alternatives options for each of the descriptor or statement in the questionnaires (Dorney, 2003; Azwar, 2011, p. 106). The score for the options themselves, thus, was ranged starting from 0 to 4 points which also indicated a gradation point from positive to negative or vice versa (Dorney, 2003). The test takers were only required to select one option on the scale that could best represent their attitude about the statements themselves (McMillan & Schumacher, 2001, p. 262). The schema of *the 5-point Likert-type scale* used in the present research will be shown in the following table.

Table 3.5
Scoring Scheme of the Test Instrument

Statement	Options				
	Never true	Usually not true	Somewhat true	Usually true	Always true
Positive (+)	0	1	2	3	4
Negative (-)	4	3	2	1	0

Additionally, the specification of the quantitative research instrument (pretest and posttest) which was adapted and developed from the construct of the theory of learner autonomy will be depicted in the following table.

Table 3.6

Specification of Research Instrument on Learner Autonomy

No	Dimensions/Sub-dimensions and Indicators of Learner Autonomy			
	Dimensions	**Sub-dimensions**	**Indicators**	**No Items**
1	Learner Involvement	Planning	Identifying inventory reading task	27
			Preparing one's own performance	
		Monitoring	Check progress on reading	
			Control progress on reading	
		Evaluating	Evaluate progress on reading	
2	Learner Reflection	Metacognition	Reflect on learning process Reflect on what is learnt from the task	8
		Metalinguistic		
3	Target Language Use	Spoken	Increase spoken English use	10
		Written	Increase written English use	
		Language of Thought	Increase the use of English as the language of thought	

3.4.1.2 Interview

Interview is considered as one of the popular instruments in educational research that typically used to reveal how people feel, react, and behave regarding certain phenomenon (King & Horrocks, 2012). For the purpose of answering the third research question, then interview was used as one of the instruments of the research. In this case, the current research employed *semi structured interview* through *individual, face-to face meeting* (King & Horrocks, 2012). This model was selected as it was able to provide flexibility in the process of interviewing that both the researcher and respondents might need (King & Horrocks, 2012).

To be more precise, the model of instrument allowed the researcher to clarify the findings which were gained from the first phase of the research, namely classroom intervention using reading portfolio to enhance learner autonomy. Reading portfolio, then, was classified as one of the important themes in the interview section. Consequently, the respondents of the interview were limited to the students of the experimental group. In this case, the interview was used in order to get detail exploration of ideas in relation to the students' responses regarding autonomous learning in reading as well as how it was fostered or diminished by the use of reading portfolio (King & Horrocks, 2012).

3.4.2 Validity and Reliability of the Research Instruments

A set of validity and reliability test and analysis were conducted In order to ensure the validity of the instruments.

3.4.2.1 Readability of the Test Instrument

The first step towards ensuring the reliability and validity of the research instrument was by testing the readability of the test instrument. This first step was conducted for the instrument that was used for the purpose of pretest and posttest. In this case, the constructed instrument was trialed with respondents who possessed the same characteristics as the population of the study (Creswell, 2012). The respondents were asked to prove read the items in the instrument in order to ensure that the test items did not consist any bias as well as to prove that the jargons or terminologies used in the test items were well comprehended by the targeted population (Dorney, 2003).

The result of the readability test of the instrument showed that the respondents did not indicate any particular problem of understanding the statements consisted in the test items both in terms of the language and the content of the test. Therefore, it can be concluded that the test items were well comprehended by the targeted population and thus it could be further used as the instrument of the research.

3.4.2.2 Expert Judgment

The next step toward ensuring the reliability and validity of the test instrument was conducting rational validity test or expert judgment of the research instruments (Arikunto, 2006). At this stage, both quantitative instrument (pretest and posttest) and qualitative instrument (interview questions) were analyzed (Arikunto, 2006, p. 168). The rationale test of research instrument was administered by consulting the pre-constructed instruments to knowledgeable expert in the research field (Arikunto, 2006, p. 168). The expert judgment or rational test was focused on three integral aspects of the instruments namely language use and choice, theoretical construct, and theoretical content (McNamara, 2000; Hughes, 2007).

In this case, each item involved in the research instrument was analyzed accordingly. From the analysis, it was revealed that from the total numbers of test items proposed by the researcher, the majority of the items were accepted meanwhile some of them were considered need to be revised. This stage of expert judgment can be considered to be in line with Creswell (2012) as he mentions that the result of the initial analyses from the knowledgeable experts may determine whether the test items proposed by the researcher can be accepted, revised, added or even rejected (see also; Arikunto, 2006). Once the instrument of the research had been thoroughly analyzed and assessed by the knowledgeable experts, then, the instrument was revised accordingly before a pilot test was conducted.

3.4.2.3 Piloting the Test Instrument

The test instrument that had been revised based on the suggestion given by knowledgeable expert was further piloted to language learners that possessed the same characteristics as the population of the study (Arikunto, 2006). In this case, the process of piloting research instrument involved 45 language learners from other vocational high school in West Java. In the stage of piloting the instrument, the alternative choices for each test item were already involved in the research instrument itself. The involvement of alternative choices was considered important since the score of each item that consisted alternative choices was further used to reveal the validity and reliability of the test instruments themselves (Creswell, 2012).

3.4.2.4 Testing Validity of the Test Instrument

After conducting the rational validity test to the knowledgeable experts and piloting the test to the population of the research study, then, the next step was calculating the correlation coefficient of each test item involved in the instrument by using statistical calculation called *Pearson Product Moment* (Arikunto, 2006). The calculation was conducted in order to ensure the validity of each test items involved in the research instrument.

Based on the result of the calculation it was revealed that the number of valid test items was 38 meanwhile the number of invalid test items was 7. The result of the validity test is as follow.

Table 3.7
The Result of Validity Test of the Test Instrument

Status	**Test Items**	∑
Valid Items	1, 2, 3, 4, 5, 7, 8, 9, 11, 12, 13, 14, 16, 17, 18, 19, 20, 22, 23, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 43, 44, 45	38
Invalid Items	6, 10, 15, 21, 24, 41, 42	7
	Total	45

For the purpose of maintaining validity of the test instrument, then, the invalid test items were diminished (Arikunto, 2006). Therefore, the numbers of the test items involved in the research instrument were reduced from what was originally 45 to 35.

To be more precise, the following table will show the specification of the research instrument after conducting the validity test.

Table 3.8
Specification of the Test Instrument Used to Reveal the Development of Learner Autonomy

No	Dimensions/Sub-dimensions and Indicators of Learner Autonomy			
	Dimensions	**Sub-dimensions**	**Indicators**	**No Items**
1	Learner Involvement	Planning	Identifying inventory reading task	21
			Preparing one's own performance	
		Monitoring	Check progress on reading	
			Control progress on reading	
		Evaluating	Evaluate progress on reading	
2	Learner Reflection	Metacognition	Reflect on learning process Reflect on what is learnt from the task	7
		Metalinguistic		
3	Target Language Use	Spoken	Increase spoken English use	10
		Written	Increase written English use	
		Language of Thought	Increase the use of English as the language of thought	

3.4.2.5 Testing Reliability of the Test Instrument

Finally, reliability of the instrument is also considered as another important aspect that needs to be considered when conducting a research study (Creswell, 2012). The reliability of the test instrument signifies the consistency of the test instrument itself to give the same score from time to time (McNamara, 2000; Hughes, 2007). This means that the more reliable a test is, then the more consistent the result of the test would be (McNamara, 2000; Hughes, 2007; Creswell, 2012). Accordingly, in order to maintain the reliability of the test item, then the result of piloted test instrument was calculated by using SPSS scale reliability analysis (Cronbach's a coefficient) (Arikunto, 2006; Creswell, 2012).

The criteria used in order to find out the level of reliability of the instrument itself is as follow (Adapted from Arikunto, 2006, see also McNamara, 2000; Hughes, 2007).

Table 3.9
Reliability Criterion of the Instrument

Criteria of Reliability Coefficient	Category
0.91-1.00	Very high
0.71-0.90	High
0.41-0.71	Moderate
0.21-0.41	Low
< 20	Very low

After conducting the calculation on each dimension/variable of learner autonomy in the research instrument, then the following result was gained.

Table 3.10

Reliability of the Test Instrument

Cronbach's Alpha	N of Items
.819	38

Based on the result of reliability test above, it can be seen that the reliability of the test instrument was about .819 therefore, it can be concluded that the test instrument had a high degree of reliability. In sum, as the instrument is indicated to have a relatively high degree of reliability, then it can be concluded that the test instruments could be further used to reveal the degree of learner autonomy

3.5 Data Collection Procedure

3.5.1 Inventory Tests: Pretest and Posttest

As indicated in the research design, in order to gather the data needed, then, tests were administered (Figure 3.1) (Campbell & Stanley, 1963; Hatch & Farhady, 1982; McMillan & Schumacher, 2001). In the first phase of data collection procedure, pretest (see appendix 1), was administered to all population so that the profile of learner autonomy of the population could be revealed as well as two groups from the population that had relatively similar initial autonomy development could be assigned as control and experimental groups (Hatch & Farhady, 1982; Creswell, 2012). Once the pretest had been administered and that the experimental and control groups had been selected, then it was followed by teaching programs (Appendix B, for the teaching program, portfolio framework and lesson plan). In this matter, the control group received conventional teaching program meanwhile the experimental group received reading portfolio as the form of intervention (Appendix B; Hatch & Farhady, 1982)

After the teaching programs had been completed, then it was followed with the second stage of data collection procedure namely posttest.

Posttest was administered to both control and experimental groups. In this case, the posttest was aimed in order to collect the information regarding the effectiveness of the intervention or treatment (Creswell, 2012; Fraenkel, *et.al.*, 2012). By this it means that the information regarding the effectiveness of the program can be gathered by comparing the depended variable in the posttest results of the experimental and control groups involved in the research (Campbell & Stanley, 1963; Hatch & Farhady, 1982; McMillan & Schumacher, 2001).

3.5.2 Reading Portfolio

The reading portfolios that were made by the students were used in order to reveal the impact of reading portfolio on developing learner autonomy. After the implementation of the reading portfolio, the students were asked to collect the reading portfolios that they have made. For the purpose of data analysis the files of the reading portfolio were copied. The original files were given back to the participants. Collecting document such as the reading portfolios can be very beneficial as it may enable the researcher to reveal how the learner autonomy developed.

3.5.3 Interviews

After statistical evaluations had been conducted, 5 respondents from the experimental group were asked to participate in an interview section. The interviews were conducted in order to ensure the findings from the first phase of the research. The interviews were given in a form of open-ended format through individual face-to face meetings (King & Horrocks, 2012). The time to conduct the interviews varied as they were conducted based on the agreement with the respondents (Creswell, 2012; King & Horrocks, 2012).

Practically, to ease the data collection procedure, the interview sections did followed the standard of interview that was developed previously (King & Horrocks, 2012). However, flexibility was also given during the interview to the respondents themselves so that the natural flow of the conversation between the respondent and the researcher could be maintained (King & Horrocks, 2012). In this regard, the standard format of interview was started by introducing of the researcher own self and also the purpose of interview. The stage of interviews was then followed by giving the respondents series of questions or statements that were considered to be able to lead to the investigation of their initial impression of students' experiences regarding learner autonomy. The next stage of the interviews was focused on exploring the students' responses regarding the use of portfolio in fostering their autonomy development. This involved the students' responses regarding how and what part of reading portfolio that had contributed to the development of learner autonomy itself.

In addition, to make interview process more comprehensive and traceable, then the process of interview was recorded by using recording devices (Fraenkel, Wallen, & Hyun, 2012; King & Horrocks, 2012). Additionally, some field notes were also provided during the interview section so that it could help the researcher to dig information from the respondents more deeply (King & Horrocks, 2012).

3.6 Data Analysis

In this section, data analysis will be discussed. The discussion will involve the quantitative and qualitative data analysis.

3.6.1 Quantitative data analysis

Quantitative data analysis technique was conducted to the data which were gained from inventory tests of the population (pretest) and from posttest of the experimental and control groups (Hatch & Farhady, 1982; Arikunto, 2006; Creswell, 2012). For the purpose of analysis, the computer software *SPSS* (*Statistical Product and Service Solution*) statistics 16.0 for windows was used. The procedures of quantitative data analysis involve; data verification, scoring, classification and analysis. Each of the procedure will be further explain in the following sections.

3.6.1.1 Data Verification

Data verification was the first procedure administered in the process of data analysis (Creswell, 2012; Fraenkel, Wallen, & Hyun, 2012). Data verification was aimed in order to check the completeness of the data gathered. The researcher listed and made sure that all of the inventory tests or pretest and posttest instruments that had been administered to the sample of the research were well verified (Arikunto, 2008; Hughes, 2007). At this point, two aspects were verified, these involve; the completeness of the respondents in filling the identity and the completeness of the respondents in answering the items provided in the test instrument. Verifying the completeness of the gathered quantitative data was considered as paramount since the absence of the respondents in filling the data needed might affect the procedure of quantitative data computation and analysis (Hatch & Farhady, 1982; Arikunto, 2008; Malik & Hamied, 2014).

3.6.1.2 Data Scoring

The next step in data analysis procedure is data scoring (Creswell, 2012). Each of quantitative data gathered from pretest and posttest was scored. At this point, in order to score the data gathered, then the measurement scale was used (Hatch & Farhady, 1982). Measurement scale itself is defined as a conduct of reference that is used to determine the length of interval of the measurement so that the result of the measurement itself can be served quantitatively (Sugiyono, 2006). In this case, the measurement was arranged in a form of data interval that had five point lengths starting from 0 up to 4 points. The scoring pattern itself will be depicted in the table below.

Table 3.12
Scoring Pattern of the Test Items

Statement	Options				
	Never true	Usually not true	Somewhat true	Usually true	Always true
Positive (+)	4	3	2	1	0
Negative (-)	0	1	2	3	4

To ease the presentation the true scores of the students were presented in a form of percentage point. This means that to get the true score of a student, for instance, then the actual score of the student itself was divided by the ideal score which was then multiplied by 100. To be more precise, the formula used is as follow (Sugiyono, 2006).

$$\text{True Score} = \frac{Actual\ Score}{Ideal\ Score} \times 100$$

$$\text{Ideal Score} = \text{k x N}_{\text{Max}}$$

Notational System:
k : Number of Items
N_{Max} : Maximum score on each test item

3.6.1.3 Data classification

Once all of the quantitative data had been scored, then the data were classified into three categories, namely high, moderate and low category of learner autonomy (Hatch & Farhady, 1986; Arikunto, 2006, p. 168). In this case, the categorizations (high, moderate, and low) were applied both to the score of learner autonomy in general and in every aspect in particular (Arikunto, 2006, p. 168). Meanwhile, the classification of the individual score of the population and the sample of the research was specifically conducted so that the information regarding the profile of learner autonomy and the effectiveness of the reading portfolio in developing learner autonomy can be clearly depicted. Accordingly, for the purpose of data classification, measure of central tendency and presentation were used. The result of the classification of the category of learner autonomy both in general and in every aspect itself will be shown in the following table

Table 3.13
Criteria for Classification of Learner Autonomy

No.	Aspect	Criteria: Raw Data	Criteria: Percentage Point	Category
1	Learner Autonomy in General	≥ 102	≥ 67	High
		51 – 101	34 – 66	Moderate
		≤ 50	≤ 33	Low
2	1st Aspect: Active Involvement	>57	≥ 67	High
		29-56	34 – 66	Moderate
		<28	≤ 33	Low
3	2nd Aspect: Reflection	>22	≥ 67	High
		11-21	34 – 66	Moderate
		<10	≤ 33	Low
4	3rd Aspect: Target Language Use	>25	≥ 67	High
		13-24	34 – 66	Moderate
		<23	≤ 33	Low

Meanwhile, the specification of the category for the degree of learner autonomy reflected from the criteria used above will be depicted in the following table.

Table 3.14
Categories of the degree of learner autonomy

Category	Qualification
High	High category of autonomy in reading practice signifies that that the students generally have reached the optimum capacity to carry active involvement, reflection and using the target language properly in the process of foreign language learning,

	especially in reading practice. Precisely, the high category of autonomy indicates that the students are certain of how to get involved in the learning process i.e. they understand on what to do in order to plan, monitor and evaluate their reading task properly (Little, 1991; Cotteral, 2008). They are also able to reflect on the process and product of their reading as well as able to improve it (Young & Yoke, 2001; Benson, 2007; Griffiths, 2008). Finally, this category also signals that the students have no difficulty in using target language optimally in reading practice either as spoken, written or as a language of thought (Little, 2009).
Moderate	Moderate category of autonomy in reading practice signifies that that the students generally have not yet reached the optimum capacity to carry active involvement, reflection and using the target language properly in the process of foreign language learning, especially in reading practice. Precisely, the moderate category of autonomy indicates that the students are still uncertain of how to get involved in the learning process i.e. they do not certain on what to do in order to plan, monitor and evaluate their reading task properly (Little, 1991; Cotteral, 2008). They also do not certain on how to reflect on the process and product of their reading, nor do they know how to improve it (Young & Yoke, 2001; Benson, 2007; Griffiths, 2008). This category also signified that the students still have difficulties in using target language optimally in reading practice either as spoken, written or as a language of thought (Little, 2009).
Low	Low category of autonomy in reading practice signifies that that the students' capacity to carry active involvement, reflection and using the target language properly in the process of foreign language learning, especially in reading practice is at stake. Precisely, the low category of autonomy signals that the

	students never get involved in the learning process i.e. they do not certain on what to do in order to plan, monitor and evaluate their reading task properly (Little, 1991; Cotteral, 2008). They also never reflect on the process and product of their reading (Young & Yoke, 2001; Benson, 2007; Griffiths, 2008). This category also signified that the students have difficulties in using target language optimally in reading practice either as spoken, written or as a language of thought (Little, 2009).

3.6.1.4 Analysis of the Effectiveness of Reading Portfolio

In order to answer the second of the three research questions i.e. to reject or accept the hypothesis, then, the result of the pretest and posttest scores was tested using normality, homogeneity, and non-parametric independent t test.

- **Normality Test**

Normality test was conducted in order to find out whether the data gathered by means of pretest and posttest of experimental and control groups were normally distributed or not. The normality test was conducted by using SPSS 16.0 for windows with the statistical computation namely *Kolmogorov-Smirnov* or *Shapiro-Wilk* (Arikunto, 2006; Sugiyono, 2006). The level of significance used in this computation was 5% (0.05) (Hatch & Farhady, 1986). The hypothesis are:

Table 3.15
Hypothesis Testing of Normality Test

Hypothesis Testing
H_o= Pretest and posttest data of the control and experimental groups were normally distributed
H_1= Pretest and posttest data of the control and experimental groups were not normally distributed

Meanwhile the criteria used for the purpose of computation were:

Table 3.16

Criteria for Hypothesis Testing of Normality Test

Criteria
a) If sig. ≥ 0,05 then Ho is accepted
b) If sig. ≥ 0,05 then Ho is rejected

The data from normality test indicated that there was one of the data which was not normally distributed, accordingly, the nonparametric Mann-Whitney test was taken in order to conduct analysis of t test as supposed to homogeneity variance (See Appendix D section 4.2).

- **Homogeneity test**

Homogeneity test was conducted in order to find out on whether the experiment and the control groups are assumed to have equal variance or not so that the researcher could be certain or feel confident that the groups are from the same population or not (Champbell & Stanley, 1963). The homogeneity of variance test was conducted by using statistical computation namely *Levene's Test* with 5% significance (alpha .05). In this case, the hypothesis used in homogeneity test was:

Table 3.17

Hypothesis Testing of Homogeneity Test

Hypothesis Testing
H_o= Pretest data from control and experimental groups are homogenous
H_1= Pretest data from control and experimental groups are not homogenous.

Meanwhile the criteria used were:

Table 3.18

Criteria for Hypothesis Testing of Homogeneity Test

Criteria

If sig. ≥ 0.05 then Ho is accepted
If sig. ≥ 0.05 then Ho is rejected

The results of the computation of homogeneity test both for pretest and posttest of experimental and control group showed that both data were homogeneous (See Appendix D, Section 4.2).

- **Nonparametric Statistical Test**

As there was one of the data gained from pretest which did not fulfill the normality assumption, then the nonparametric t test (Champbell & Stanley, 1963; Hatch & Farhady, 1986) called Mann-Whitney was used in order to infer the effectiveness of the reading portfolio. Similar to the normality and homogeneity test, the 5% significance level was used as the basis of accepting or rejecting the hypothesis.

Meanwhile, the hypothesis used in the equal variances test is as follow:

Table 3.19
Hypothesis Testing for Equal Variance

Hypothesis
$H_0 : \mu_{pre} = \mu_{prk}$ $H_1 : \mu_{pre} \neq \mu_{prk}$

The notational system of the hypothesis above is:

Table 3.20
Notational System of Hypothesis Testing for Equal Variance

Notational system
μ_{pre} : average of pretest and posttest of the experimental group. μ_{prk} : average of pretest and posttest of the control group.

Meanwhile the criteria used in this procedure are:

Table 3.21

Criteria for Hypothesis Testing

Criteria
a. If sig ≥ then Ho is accepted
b. If sig < 0.05 then Ho is rejected.

The independent sample *t* test was used to compare the progress of the learner autonomy in both groups. It was calculated by comparing means scores which were obtained from the two groups. The result of the *t-test* can be used as an indication to determine which one of the experimental and the control groups that made more significant progress and eventually answer the research question whether reading portfolio is effective to foster the development of learner autonomy or not (Campbell & Stanley, 1963; Hatch & Farhady, 1982; McMillan & Schumacher, 2001). The result of *t* test both for pre and posttest will be depicted in the following chapter.

3.6.2 Qualitative Data Analysis: Interview and Document Analysis

The interview was conducted to the students from experimental group after the intervention (reading portfolio) and test had been completed. The interview was conducted to five participants. At this point, the data that were gained from interviews were in a form of recording that consisted of conversation between interviewer and interviewees. Accordingly, in order to ease the researcher in interpreting the recording data, then, it was transcribed accordingly (King & Horrocks, 2012). In the process of transcription, the participants' names and identities were coded or in this case were replaced by pseudo names so that the participants' privacy and confidentiality could be protected (King & Horrocks, 2012). After that, the transcribed data were further analyzed by systematically looking through it, clustering, or grouping similar idea and labeling them (King & Horrocks, 2012). The aforementioned data analyses were critical as they were aimed to make sense the transcribed data. Accordingly, the final results of the analyses were sorted so that the data had two main categories of information; these involved the students' responses on the impact of reading portfolio on developing learner autonomy in general and the students' responses on the impact of reading portfolio on developing learner autonomy in every aspect in particular.

Document analysis was also conducted in order to support the findings on the impact of reading portfolio on the development of learner autonomy. There were 32 reading portfolio files collected from the research participants. The collected documents were kept and analyzed accordingly. In this case, in order to ease the researcher in interpreting documents, then the documents were clustered and labeled (King & Horrocks, 2012). The results of the document analyses were sorted into two main categories of information. The first category of information is focused on revealing the indicators that could signify the impact of reading portfolio in developing learner autonomy in general. Meanwhile, the second category of information is focused on revealing the evidence of the development of learner autonomy in each aspect (learner involvement, learner reflection, and target language use). In addition, the examples of the information consisted in the documents (reading portfolio) were also provided in the finding and discussion (Chapter 4).

3.7 Research Ethics

3.7.1 Informed consent

Prior to conducting the procedure of data collection, an informed written consent was distributed to every participant of the study. The consent used in the present research was adapted from Furqon and Emilia (2010) (See also King & Horrocks, 2012). In the consent itself the participants were informed regarding the purpose of the study, the confidentiality, as well as the right and duty of the participants themselves. The consents were provided in Bahasa so as to avoid any misunderstanding in between participants and researcher. In this case, the willingness of the participant to participate in the research was signaled by them signing the distributed consent (King & Horrocks, 2012).

3.7.2 Anonymity and Confidentiality of the Respondents

Other means of keeping ethical conduct of the research is by protecting the identity of the participant as well as the information given by them. The protection was given in a form of anonymity of the research participant (Furqon and Emilia, 2010; Creswell, 2012; King & Horrocks, 2012). The participants were only recognized based on their initials, student numbers and pseudo names both in the stage of quantitative and qualitative data collections, analyses and presentations. Any information that may harm the participants was also protected.

3.7.3 Level of Risk

This present research was arranged as to pose as minimum of risk as possible both to the participants and other parties involved in the research. Consequently, the classroom intervention was administered at the participants' own school and at the usual time assigned for English subject (Furqon and Emilia, 2010; King & Horrocks, 2012). In addition, the materials given to the participants were also aligned with the materials that were expected to be given to them. Thus, the level of altering the teaching and learning process as assigned by the curriculum was kept at minimum (Furqon and Emilia, 2010). This effort could be achieved by the support and guidance of the initial English teacher who was responsible for the control and experimental groups. The recording processes of both in terms of note and voice recording were conducted respectfully.

3.7.4 Researcher's Role

The researcher took a role as teacher-researcher or practitioner researcher. Johnson (2005) and Alexakos (2015) define the term teacher-researcher or practitioner-researcher as a person who conducts research on his/her own practice. In accordance, Cochran and Lytle (1993) also claim that the term teacher-research actually refers to a systematic and intentional inquiry carried out by teachers. Several models of applying the roles as teacher-researchers have been established. The models are varied depending on the research approach adapted by the teacher-researcher. The researcher goal was to generate and develop learner autonomy in the classroom. This could be achieved by increasing the student's involvement, reflection, and use of target language in the classroom. Therefore, the researcher decided that one of the ways to achieve the goal, i.e. answering the research question, was to use both quantitative and qualitative approaches.

One of the benefits of having a role as "practitioner researcher" lies in the fact that the researcher had more access to and control over the teaching and learning options and possibilities. Accordingly, the researcher used the options and possibilities available to set up the procedure of the classroom experiment and interview. In this case, with a help and guidance of a teacher-mentor, the researcher was able to identify the needs and the autonomy profiles of the students, select the experimental and control groups, set up the intervention programs, monitor the development processes of learner autonomy, determine the impact of the experiment on the development of learner autonomy, and evaluate the outcomes and the changes that the students experienced.

Although the role as practitioner researcher had been able to give benefits for the research, yet there are also some shortcomings recognized. These shortcomings can lead to the research bias. In this case, the first possible shortcoming relates with the subjectivity of the researcher. In this research, several things such as the duration of the experimental procedure, teaching materials, and the expected learning outcomes were set up by the researcher and they were adjusted with the school curriculum. The adjustment process was carried out so that the teaching and learning process and product could not only fulfil the convention of the research approach but also meet the expectation of the curriculum. Although adjustment was carried out, yet the researcher's subjectivity during the process of the selecting the materials and during the interpretation of the learning outcome is unavoidable.

Another possible shortcoming or bias is in a form of influence of the researcher on the participants. It is stated that the influence of the researcher and the research on the participants is inevitable. Thus, it cannot be claimed that the participants in the research were not in some ways influenced by the role that the researcher took especially at the process of enacting the teaching and learning practices.

Although shortcomings exist yet there are also several ways which were conducted in order to cope with them. One of the ways to overcome the research bias is by conducting reflection. Therefore, as a form of reflection, after completing the research the researcher involved the students and the teacher-mentor in a discussion. Together the researcher and each of the students talked about how the experimental strategies worked for them. Meanwhile, together the researcher and the teacher-mentor also tried to reflect on the possible success and failure of the research program.

3.8 Concluding Remark

In conclusion, this chapter has focused on the research method that was specifically intended to answer the research questions. The research method used in this study is served as a mirror on how the researcher conducted the research.