

CHAPTER III

RESEARCH METHODOLOGY

3.1. Introduction

There are four parts which will be discussed in this chapter. They are the research design, the research site and the participants of the research, the data collection which will also describe the procedure of the implementation of the extensive listening, and and the last one is the data analysis.

3.2. Research Design

The research design used in this study is quantitative study and employed an experimental research. According to Fraenkel and Wallen (2008), experimental research is a type of research that directly attempts to influence a particular variable, and when properly applied, it is the best type for testing hypothesis about cause-and-effect relationships. In an experimental study, researchers look at the effect(s) of at least one independent variable on one or more dependent variable. In this research, the independent variable would be extensive listening, and the dependent variable would be listening skill.

Experimental research can also be used to test our hypothesis. According to Burns (1994), hypothesis is an educated guess which is suggested for the purpose of being tested. In this study, the hypothesis stated is the Null Hypothesis which states that there is no significant difference between the experimental group that implement the extensive listening procedure and the control group that does not implement the extensive listening procedure.

The design used in this study was quasi-experimental design. Quasi-experimental design is similar to true-experimental design. The difference is that

the quasi-experimental design does not include the use of random assignment (Christensen, 1941; Fraenkel & Wallen, 2008; Creswell, 2009). In other words, the participants are chosen purposefully to fulfill the needs of the study.

The quasi-experimental design is divided into two categories, they are the nonequivalent before-after design and the simulated before-after design (Christensen, 1941). The research design which is used in this study is the nonequivalent before-after design. Another name is used by Creswell (2009) concerning this type of research design. It is called *pre- and post-test design*. It is used considering the small numbers of samples (< 30) in one group so it is impossible to do the random assignment of subjects. The table of the design is as follows.

Table 3.1.
The design of pre- and post-test

Pre- and post-test design time →

Select experimental group	Pre-test	Treatment	Post-test
Select control group	Pre-test	No treatment	Post-test

In this design, there are two groups which are selected purposefully. The experimental group and the control group. The experimental group is the group who receive a treatment after a pre-test is administered to both groups. And after the treatment, a post-test is given also to both groups to see whether the treatment has a positive effect or not (Creswell, 2009).

3.3. Research Site and Participants

The site of the study was one of an Islamic private high school in Bandung Barat. It was chosen because it was easy to get access to the school. The participants of this study are 41 students of the eleventh grade. They were chosen because they are going to have national examination the following year and since

there is also listening test in the national examination, they need to be prepared and also prepare themselves to face the examination. The students are from 2 classes of Science class and social science class. Since there is no difference in the curriculum for those classes, the two classes are suitable for this study. Their average ability in English is also similar so the two classes will be treated as two groups. The Science Class, which consists of 20 students, was the experimental group. The Social class, which was chosen as the control group, consists of 24 students but three of the students were absent during the pre-test and the post-test, so there are only 21 students who were chosen as the participants of the study from this group.

3.4. Data Collection

In collecting the data of the study, which used the pre- and post-test design, there were three kinds of instruments which were given to the participants of this study. They were the pre-test, the post-test and the questionnaires.

3.4.1. Instrumentation

Instruments are tools or devices for measuring, observing, or documenting quantitative data (Fraenkel & Wallen, 2008:110; Creswell, 2012:151). They are used to measure achievement, assess individual ability, observe behaviour, develop a psychological profile of an individual, or interview a person. In this study, the instruments used for collecting the data are tests and questionnaire. The tests were used for measuring the students' achievement of their listening skill and the questionnaire was for observing the students' behaviour and perception during the experimental study.

3.4.1.1. Pre-test

The pre-test was given to the participants of the study. It aimed to measure some attribute or characteristics of the participants in an experiment before they receive a treatment (Creswell, 2012). There are two other purposes of giving a pre-test, Linn & Gronlund (1995) mention that testing and assessment may be given at the beginning of an instructional segment to determine whether students have the prerequisite skills needed for the instruction (to determine readiness) or to determine students' placement or modification of instruction (Linn & Gronlund, 1995). In this study, the test was given to measure the students' ability in listening comprehension before the students receive the treatment which was a set of instructions to do extensive listening. The test consisted of listening test which was adapted and modified from a book titled 'New English Files' and also from the listening section of National Examination year 2005. There were 25 items of the listening test which are based on four dialogues and one monologue. The items were in multiple choice form with one stem and five options.

There were some considerations why the multiple choice form were used in the tests. According to Linn and Gronlund (1995: 197), the multiple choice form is flexible and can be used to measure various learning outcomes in any level of understanding. It is also applicable in the measurement of various phases of achievement. Oller (1979) also adds that the advantages of multiple choice test are that it is more convenient in scoring and administrating the result of the test. these considerations have become the reasons why multiple choice form was used as the tool for measuring the students' understanding of the listening materials.

In the test, the students were presented with a stimulus dialogues and monologues and then they were asked to respond to a set of comprehension questions. This is also known as communicative stimulus – response tasks (Brown, 2004: 132). The following is the example of the dialogue and multiple-choice comprehension items in the test.

Table 3.2

The example of the dialogue with the comprehension items

The students hear:

Listen to the dialogue and choose the best answer!

(Michael is calling Susan on the phone)

Susan : Hello.
Michael : Hi, Susan. It's me, Michael.
Susan : Oh, hi, Michael. How are you doing?
Michael : Oh, not bad. Say, are you doing anything on Saturday night?
Susan : No, nothing special. Why?
Michael : Well, do you feel like going to that new disco?
Susan : Oh, that's a terrific idea.
Michael : Great. What about having a pizza first?
Susan : Sure, why not? Where do you want to meet?
Michael : Why don't I pick you up at your house?
Susan : OK. What time?
Michael : Is seven o'clock OK?
Susan : Fine. Well, see you at seven.
Michael : Right. So long for now.
Susan : Bye.

The students read:

1. Why is Michael calling Susan on the phone?
 - a. He wants to say Hi
 - b. He wants to have a pizza
 - c. He wants to invite Susan to go out
 - d. He wants to go to see a concert
 - e. He wants to go to a new disco
2. Where are they going to meet?
 - a. At Susan's house
 - b. At the new disco
 - c. At the cinema
 - d. At Michael's house
 - e. At the restaurant
3. Which statement is TRUE?
 - a. Susan invites Michael to go to a new disco

- b. Susan asks Michael to pick her up at her house
- c. They are going to have a pizza before they go to the disco
- d. They are going to have a pizza after they go to the disco
- e. They are going to go on Friday night

The items in the pre-test consisted of 25 items which were based on five dialogues and one monologue. The topics of the dialogues for the test are described in the following table.

Table 3.3.

The topics of the dialogues for the test

Number of items	Topic of the dialogue and monologue
1 to 4	Dialogue between Natasha and Darren about their first meeting. Taken from 'New English File'
5 to 7	Dialogue about Michael who invites Susan to go out. Taken from 'Person to Person'
8 to 12	Dialogue between Simon and his doctor about Simon's problem. taken from 'New English File'
13 to 17	Dialogue between Mark and Allie who said goodbye in the airport. Taken from 'New English File'
18 to 21	Dialogue between Charlotte and Rob who talk about their plan for holiday and about Charlotte's anger to Rob. Taken from 'New English File'
22 to 25	Monologue about Albert Einstein. Taken from 'National Examination (2005)'

3.4.1.2. Post-test

After the pre-test, the participants received a treatment which consisted of instructions and guidance to do the extensive listening procedure. Here, the

students were given a set of listening materials which they took home and listened to on their own free time.

Beside the listening materials from the teacher, the students were also encouraged to listen to other listening materials by their own choice. The listening material can be English songs, stories, films, news, etc. The length of the treatment was two months, and the students' progress and achievement was checked weekly during the treatment.

After two months of treatment, the students were given a post-test. The main interest of a post-test is in measuring the extent to which the intended learning outcomes have been achieved (Linn & Gronlund, 1995). In other words, this post-test was given to measure the improvement of the students' listening skill. The test items in the post-test was the same as the test item in the pre-test.

3.4.1.3. Questionnaire

The other instrument used in this study is questionnaire. Brown (2001), as cited in Dornyei (2009), proposes a definition for questionnaires as follows:

“questionnaires are any written instruments that present respondents with series of questions or statements to which they are to react either by writing out their answers or selecting from among existing answers”

(Brown, 2001:6)

There are three types of data about the respondents which are produced from questionnaires, they are: factual, behavioral, and attitudinal which also concerns attitudes, opinions, beliefs, interests, and values (Dornyei, 2009:5).

In this study there were only two types of data which were attempted to be found, behavioral and attitudinal. The attitudinal questions only concern attitudes, opinions, and interests. These questionnaires were intended to answer the second research question, which was about the students' perceptions of the implementation of the extensive listening. The students were given a set of questionnaires related to the topic and were asked to give their own opinion about the extensive listening instructions which they had received before.

The type of the questionnaire used was the close ended questions in which the respondents were provided with ready-made response options to choose from (Dornyei, 2009:26). Dornyei (2009) furthermore adds that the advantage of this type of questionnaire is that their coding and tabulation is straightforward and leaves no room for rater subjectivity.

The questionnaire consisted of 20 statements and were divided into four parts. The first part, which included statements number 1, 4, 6, 7, and 12 was about the students' interest in English and listening comprehension. The second part, statements number 2, 9, 13, 14, 16, and 20, asked about the students' opinion of the implementation of extensive listening. The third part, which was asking about the students' attitudes toward English and listening comprehension, included statements number 3, 5, 8, 10, and 19. The last part which included statements number 11, 15, 17, and 18 was asking about the students' behavioral in studying English and listening comprehension. In answering the questionnaire the students had to choose one of the options provided based on their opinion and interest. The options were 'strongly disagree', 'disagree', 'slightly disagree', 'slightly agree', 'agree', and 'strongly agree'. The questionnaire can be seen in the appendix.

In order to avoid misunderstanding and misinterpretation, the questionnaire was written in Bahasa Indonesia and the instruction to the students was also given in Bahasa Indonesia.

3.4.2. Procedure

This study was conducted not only to answer the research questions but also to test the hypotheses. Before testing the hypotheses, the measurement of variables or the test items and the questionnaires have to be checked whether they are valid and reliable (Hatch & Lazaraton, 1991). The reliability of a test is defined as the extent to which the result can be considered consistent or stable

when administered under similar condition (Brown 1988; Hatch & Lazaraton, 1991).

3.4.2.1. The reliability and validity of the test

To estimate the reliability, according to Brown (1988), there are three ways which can be chosen to conduct. The first one is *test-retest reliability* which is usually estimated by administering the same test twice to one group of subjects and then calculating a correlation coefficient between the pairs of scores from two administrations. The second one is *Equivalent forms reliability* which is estimated by administering two equivalent tests to one group of subject and calculating the correlation coefficient between the scores on the form. The third one is *Internal consistency reliability* which can be estimated in a number of ways. One of the easiest methods to understand is *the split half method*.

There are three steps in this method of measuring reliability of the test. The first step is scoring the even-numbered and the odd-numbered separately. The second one is calculating the correlation between the scores and the last one is adjusting the result of the calculation using a formula called the Spearman-Brown prophecy formula for full test reliability (Brown, 1988: 99).

There is also another way to estimate the reliability. It is estimated without using statistical procedure and formula but only by using the ANATEST. In this study, the ANATEST was used instead of the statistical procedure mentioned above. It was used not only for estimating the reliability of the test but also for analysing whether the test items for the try out and pilot study were good enough. And then the result of the analysis was used as a guideline for making the tests for the pre test and the post test.

Besides estimating the reliability, the tests used in a study should also be valid. According to Brown (1988), test validity can be defined as the degree in which a test measures what it claims to be measuring (see also Brown, 2004: 22; Linn & Gronlund, 1995: 47). Fraenkel & Wallen (2008) furthermore add that

validity refers to the appropriateness, meaningfulness, correctness, and usefulness of the inferences a researcher makes.

There are three kinds of validity, content validity, construct validity, and criterion-related validity. Content validity refers to the content and the format of the instrument, the content and format of the test must be consistent with the definition of the variable and the sample of subject to be measured (Fraenkel & Wallen, 2008: 148). In this study, the listening test used can be considered valid because the content and the format of the test are appropriate for measuring the students' listening ability and the test items have also been approved by experts, i.e. the supervisors, who were involved in the process of the study.

Construct validity refers to the nature of the psychological construct or characteristic being measured by the instrument (Fraenkel & Wallen, 2008: 148). One of the ways of establishing the construct validity of the test, according to Brown (1988), is to set up an experiment to demonstrate that a given test is indeed testing the construct that it claims to be testing. In this study, the construct validity was established by doing the try-out and the pilot study to the students who are the participants of the study.

Another kinds of validity which has to be established is the criterion-related validity. It refers to the relationship between scores obtained using the instrument and scores obtained using one or more other instruments or measures (Fraenkel & Wallen, 2008: 148; Linn & Gronlund, 1995: 50).

3.4.2.2. The reliability and validity of the questionnaire

The reliability and the validity of the questionnaire, which was also the instrument in this study, should also be established. They are two key concepts in measurement theory which refer to the psychometric properties of the measurement techniques and the data obtained by them (Dornyei, 2003: 110). He furthermore adds the definition of the reliability of a psychometric instrument which refers to the extent to which scores on the instrument are free from errors of

measurement, and validity is the extent to which a psychometric instrument measures what it has been designed to measure.

To estimate the reliability of the questionnaire, one of the statistical procedures which can be used in the SPSS is by estimating *the Cronbach Alpha coefficient*. Dörnyei (2003: 112) mentions that questionnaires can be considered as reliable when the result of the analysis exceeds 0.70. In this study, from the analysis using the SPSS it was found that the value of the Cronbach Alpha of the questionnaire used was 0.72. This means that the questionnaire used in this study can be considered reliable.

The validity of the questionnaire should also be taken into account. Which means that the content of the questionnaire needs to cover the study issues that have been specified (Gray, 2004). Foddy (1993) as cited in Gray (2003), suggested a series of steps that must be followed in establishing the validity of the questionnaire.

- The researcher has to be clear about the information required and encode this accurately into a question.
- The respondent must interpret the question in a way the researcher intended.
- The respondent must construct an answer that contains information that the researcher has requested.
- The researcher must interpret the answer as the respondent had intended to be interpreted.

Foddy (1993) in Gray (2003)

In this study, all the steps stated above had been conducted. The statements in the questionnaire aimed to find information about the students' opinion about listening in general and extensive listening, and the students' attitude towards the implementation of extensive listening. From the reliability analysis, it was found that the students did not get difficulties in understanding the questions and they were able to give reliable answer. The last step which was about analysing and interpreting the students' answer was also conducted thoroughly. The result and analysis of the questionnaire can be seen in Chapter IV.

3.4.2.3. Listening Activities in Experimental and Control Group

This study aimed to measure the effect of implementing extensive listening procedure on the students' listening skill. It implemented one of the extensive listening procedure proposed by Vandergrift and Goh (2012), which is the Facilitated Independent Listening. The listening procedure was implemented or administered in the experimental group, where students were given extra assignment to practice their listening skills outside the classroom in their own free time. The control group also practiced their listening skill by using intensive listening which was only conducted in the classroom.

The difference between the activities in the experimental group and the control group can be seen in the following table which is about the sample of teaching procedure in both groups.

Table 3.4
The sample of teaching procedure

	Experimental group (implementing the extensive listening)	Control group (conventional technique / selective listening)
	Pre activities <ul style="list-style-type: none"> • The teacher greets the students • The teacher tells the objectives of the lesson and explains the activity that the students will do • The teacher introduces a topic that is going to be discussed in the lesson • The teacher and the students discuss the topic of the lesson. 	Pre activities <ul style="list-style-type: none"> • The teacher greets the students • The teacher tells the objectives of the lesson and explains the activity that the students will do. • The teacher introduces a topic that is going to be discussed in the lesson • The teacher and the students discuss the topic of the lesson.
	Whilst activity <ul style="list-style-type: none"> • The teacher plays a recording of a dialogue. • The students listen to the dialogue. • The teacher asks the students what the dialogue is about and then discusses the dialogue with the students. • The teacher explains about the function and the expression stated in the dialogue 	Whilst activities <ul style="list-style-type: none"> • The teacher plays a recording of a dialogue. • The students listen to the dialogue. • The teacher gives some questions related to the dialogue and asks the students to answer the questions. • The teacher asks the students to listen to the dialogue again to check their answers. • The teacher and the students discuss the dialogue and the answers to the questions of the dialogue.

		<ul style="list-style-type: none"> The teacher explains about the function and the expression stated in the dialogue.
	<p>Post activity</p> <ul style="list-style-type: none"> The teacher gives some recordings of dialogues and asks the students to listen to other dialogues at home. The teacher asks the students to visit some websites to find some listening materials that they can listen to and study. The teacher reviews the materials that have been discussed. The teacher reminds and encourages the students to do the extensive listening at home. 	<p>Post activities</p> <ul style="list-style-type: none"> The teacher reviews the materials that have been discussed. The teacher gives homework or assignment related to the teaching materials.

From the table, it can be seen that the difference between the experimental and the control group is that there were extra assignment for the students to do the extensive listening outside the classroom. The assignments included listening to the listening materials given by the teacher at home, visiting some websites to find listening materials and practicing the listening materials they found in their free time.

3.5. Data Analysis

The section will discuss the statements of the hypothesis of the study and then the procedure used in analysing the data of this study. It also includes the explanations of data analysis during the try out, the pilot study, the pre-test and post-test, and also the analysis of the questionnaires.

Before analyzing the data, the hypotheses of the study were needed to be formulated. The hypotheses stated here are the Null hypothesis and the alternative hypotheses.

H_0 : there is no significant difference between the means of the experimental group that implement the extensive listening and the control group that do not implement the extensive listening.

H₁ : there is significant difference between the means of the experimental group that implement the extensive listening and the control group that do not implement the extensive listening.

The procedure of collecting and analyzing the data during the try-out, the pilot study, the pre-test, the post-test and data analysis of the questionnaires are as follow.

- The try out

The purpose of analysing the result of the try out is to find appropriate test items for the pilot study and the pre-test and post-test. There were 30 items in the test which were divided into five parts. The division of the items is described in the following table.

Table 3.5.
The types of items used in the try out

Items 1 to 10	Listening to the difference and choosing the one that they hear.
Items 11 to 15	Listening to questions and choosing the best response for the questions.
Items 16 to 20	Listening to the difference and choosing the one that they hear.
Items 21 to 25	Listening to questions and choosing the best response for the questions.
Items 26 to 30	Listening to two monologues and choosing the best answer for the comprehension questions given.

The test items in the try out were in multiple choice form, with one stem and five options. The result of the test was scored and analysed by using ANATEST. There are some advantages of item analysis. The first one is that item-analysis data provides a basis for efficient class discussion of the test result. It also provides a basis for remedial work of the items of the test, for the general

improvement of classroom instruction and also for increased skill in test construction (Linn and Gronlund, 1995).

ANATES which was used for item analysis reveals that the items of the test is too easy, moderate, or too difficult. It also tells whether the distractor or the options of the test is very good, good enough, or even bad and need to be changed. (the result of the analysis can be seen in appendix 12).

The construction of the test items in the try out was based on list of micro- and macroskills of listening proposed by Brown (2004), and the conversational listening taxonomy suggested by Richards (1983) as cited by Buck (2001) stated in Chapter II.

- The Pilot Study

The pilot study was conducted after administering the try out and analyzing the result of the try out. This was done to prepare the students to take the pre-test before the experiment and also to check whether the items used in the tests were valid and reliable. The items of the test of the pilot study were the revision and changes from the items in the try out. The revision and changes were needed because after conducting the item-analysis, it was found that there were test items in the try out which are considered too easy and too difficult, and there were also some options which are considered as bad distracters. The result of the pilot study was also analysed by using the ANATEST and the result of the item-analysis was used as a guideline in making revision and changes in the item test for the pre-test.

- The Pre-test

The pre-test was conducted before the treatment in the experimental group which is the implementation of extensive listening. It was given to both the experimental group and also the control group. The test items for the pre-test consists of 25 items, which were divided into five parts based on the dialogues

that the students have to listen. The result of the pre-test was scored manually and then the mean scores were analyzed by using the SPSS. The analysis used were the Kolmogorov-Smirnov test which was used to analyze the normality of the scores, and then the paired sample t-test and the independent sample t-test which were used to find the differences between the means and also the significance of the difference between the mean scores. The result of the tests can be seen in the appendices.

- The Post-test

The post-test was conducted after the implementation of the extensive listening in the experimental group. The items of the test were the same with the items in the pre-test. The result of the tests were then scored, and the mean scores were calculated and also analyzed by using the SPSS. The difference of the mean scores between the experimental and control group were analyzed by using the paired samples t-test. The difference and the significance between the two groups were also analyzed by using the independent samples t-test. These were done to find out the effect of the treatment on the experimental group.

- The Questionnaire

The type of the questionnaire used was close-ended question, with the Likert scale as the rating scale. The questionnaire consists of 20 items and were divided into four groups to measure the students' interest, opinion, attitude and behavior. With close-ended items, coding is usually straightforward: each predetermined response option is assign a number (Dornyei: 2009.p: 85). There were six options in this questionnaires, strongly disagree, disagree, slightly disagree, slightly agree, agree and strongly agree. The coding for each option is: 1 for 'strongly disagree', 2 for 'disagree', 3 for 'slightly disagree', 4 for 'slightly agree', 5 for 'agree', and 6 for 'strongly agree'.

The result of the coding was then analyzed by using the SPSS and the analysis used was the Cronbach's Alpha which was used to measure the reability of the questionnaire. To calculate the mean scores for each item of the questionnaire, the number of the option was calculated with the number of

frequency and then divided by the number of students (N=20). By analyzing this mean scores, the students' perception about the implementation of the extensive listening can be found out.

3.6. Conclusion of the Chapter

This chapter has discussed the methodology used in this study. It includes the elaboration of the research design which is the quasi experimental design, the research site and participants, the procedure used in collecting the data, and the procedure in analyzing the data. The type of quasi-experimental used in this study is the pre- and post test design, which seeks to find out the difference and the significance of the difference between two groups (the experimental and the control group). In collecting the data, this study used three kinds of instruments, which are the pre-test, the post-test and the questionnaire. The collected data are then analysed by using SPSS (statistical packages in the social science).