CHAPTER III

RESEARCH METHODOLOGY

Chapter III elaborates the research methodology of the present study. The main points of this chapter involve the research design employed in the study, the site selection, the data sources, and the methods and procedures of the data collection and the data analysis.

3.1 Research Design

The research design used in the present study is driven by the aforementioned research questions that will be answered throughout the study. This study aims at investigating how English teachers construct multiple-choice test items, what difficulties they face, and the quality of the multiple-choice test items constructed by the English teachers. Hence, based on the nature of the research questions, both qualitative and quantitative methods were used for finding the answers.

Specifically, this study can be labeled as a descriptive study. Following Shields and Rangarjan (2013), descriptive study is employed to describe characteristics of a population or phenomenon being studied. Thus, this kind of research commonly requires the provision of careful and detailed description of the population or phenomenon (Mackey & Gass, 2005; Cresswell, 2012). In this point, this study attempts to gain a detailed description of English teacher-made test construction practices and the quality of the test items.

Descriptive research design is powerful in documenting and establishing a rich description and attaining comprehensive summarization of a phenomenon or specific experience of an individual or group of individuals (Marshall & Rosman, 1999; Lambert & Lambert, 2012). Thus, the purpose of this study is to describe, not to produce theory from the data generated (Lambert & Lambert, 2012).

3.2 Site Selection

As has been stated earlier in Chapter I, the present study attempted to investigate the steps taken by English teachers in multiple-choice test construction, the difficulties they face, and the quality of the test items constructed by English teachers in Indonesian context. Nevertheless, it would not be feasible to examine all English teachers in the Republic of Indonesia. Hence, this study selected only one city in Jawa Barat province, Indonesia, which was Sukabumi city for several reasons. Firstly, since the location of the city is easy to reach, it is efficient in time and cost to conduct the study in the city. Secondly, based on several prior observations, there are a lot of English teachers with various backgrounds and characteristics in the city. In the purpose of gaining richer data from English teachers with various backgrounds and characteristics, the city is considered suitable as the site of this study.

3.3 Participants

The participants of the study were selected using purposive sampling method. Purposive sampling is the process of selecting the sample which is considered as representative of a population (Gay, Mills, & Airasian, 2009). In addition, Merriam (1998) defined purposive sampling as the process of selecting a sample from which one can learn optimally to discover, understand, and obtain insight. Moreover, Malik and Hamied (2014) stated that purposive sampling is where the sample is selected based on who they are and what they know. Hence, being selected using purposive sampling method, the participant is expected to be able to be appropriate sources in order to gain rich data.

This study involved 6 English teachers as the participants. The English teachers were currently teaching in 3 different public senior high schools in the city being selected as the research site. The participants were selected based on their gender, age, grade level, teaching experience, educational background, and score of UKG (Uji Kompetensi Guru) Indonesia. It was intended to gain various and rich data. Also, according to Zhang and Burry-Stock (2003) and Alkharusi

45

(2011), such variables are influencing in teachers' assessment skills. The demography of the participants in the study is presented as follows:

Table 3. The Demography of the Participants

	Gender	Age	Grade Level	Teaching Experience	Educational Background	Score of UKG
Teacher 1	Female	41	X	11 years	S1	69
Teacher 2	Male	40	XI	13 years	S2	73
Teacher 3	Female	37	X	9 years	S1	47
Teacher 4	Female	35	XI	7 years	S2	65
Teacher 5	Male	51	XI	25 years	S2	51
Teacher 6	Female	29	X	5 years	S1	75

3.4 Data Collection

The data of this study were collected through several methods involving questionnaire, interview, and document analysis as a means of triangulation. Triangulation is intended to seek convergence through the use of different data sources and methods (Bowen, 2009). Moreover, by triangulating data, the credibility of a research can be fostered (Eisner, 1991). The detail of each data collection method is described as follows.

3.4.1 Questionnaire

The questionnaire was administered to all participants (6 English teachers) in the purpose to gain the data of the steps taken by the English teachers in preparing and doing test construction. The questionnaire was adapted from Test Construction Skill Inventory (TCSI) developed by Agu, Onyekuba, and Anyichie (2013). Moreover, the questionnaire was also adjusted with the procedures of test construction proposed by Gronlund (1977) and Wragg (2001).

The questionnaire was presented in English and consisted of 26 statements that should be responded by the participants by giving check (√) on the columns representing their choice (See Appendix). The questionnaire used 5-point Likert Scales enabling the participants to choose 'always', 'very often', 'sometimes', 'rarely', and 'never'. The questionnaire was administered in the beginning of the study. After getting access to the participants, the questionnaires were delivered to all participants at the same day. The participants were given 3 days to respond the questionnaire and after that, the completed questionnaires were collected.

3.4.2 Interview

Following Scott and Usher (2011), interviewing is an essential tool to be used by researchers in educational field. Thus, interview was employed as the second data collection method in this study.

Interview is conducted to verify, extend, support, and elaborate the data collection (Cohen, Manion, & Morrison, 2007). In the case of this study, the interview was intended to verify, extend, and enrich the data of the steps taken by the English teachers in preparing and doing test construction which had been collected through the questionnaire. Also, the interview aimed at seeking the difficulties faced by the English teachers in constructing multiple-choice test items.

The interview was set in the form of semi-structured interview that consisted of 15 questions (See Appendix). Semi-structured interview was selected due to its advantage of giving a great deal of flexibility to the interviewer (Nunan, 1992). In general, the questions asked in the interview were divided into two broad categories: one asking about the process in preparing and doing test construction, and the other one asking about the difficulties faced by the teachers in constructing multiple-choice test construction.

The interview was conducted one by one in face to face mode. In order to provide a comfortable condition, the interview was conducted in the participants' native language which was Bahasa Indonesia. Creating a comfortable condition during an interview is essential so that the participants are not hesitant to talk (see Cresswell, 2009). Moreover, the interview was recorded using audio recorder.

3.4.3 Document Analysis

Another data collection method used in this study is document analysis. Following Bowen (2009), document can serve various purposes in a research. In line with the statement, Merriam (1988) stated that document can help researcher uncover meaning, establish understanding, and discover insights relevant to the research undertaken.

The documents used in this study were the latest summative multiple-choice test items constructed by the participants. Since the test items were strictly confidential, the data collection was conducted right after the test was administered to the students. Each set of summative test consisted of 45 and 50 multiple-choice test items with 5 alternatives (See Appendix). Hence, there were 275 multiple-choice test items in total. Moreover, the answer sheets of the students taking the test were also collected in order to be analyzed through item analysis.

According to Angrosino and Mays de Perez (2000), document analysis is typically used to verify the findings taken from other methods. When there is a convergence of data from different sources, the credibility or trustworthiness of the study can be fostered. In case of this study, the document analysis was intended to answer the second research question of the quality of the test items constructed by the participants. The quality of the test items was examined based on the validity analysis, reliability analysis, and the β-index analysis. Also, the item format, item quality, and the proportion of higher-order and lower-order thinking contained in the test items were analyzed. Moreover, the result of the document analysis was also used to verify the findings taken from questionnaires

and interview. The participants' skill and knowledge in test construction along with the difficulties they face in constructing the test could be reflected in the documents. Hence, the document analysis was worth to conduct.

3.5 Data Analysis

The data taken from questionnaire, interview, and document analysis were analyzed using the procedures of both qualitative and quantitative data analysis. In details, the data analysis of the data obtained from each data collection method is presented as follows.

3.5.1 Analysis of Data from Questionnaire

The data obtained from the questionnaire were analyzed descriptively following the process suggested by Patton (1980) which involved bringing order to the data, organizing what is there into patterns and categories, and basic descriptive units. Moreover, the data condensation procedure – "a form of analysis that sharpens, sorts, focuses, discards, and organizes data in such a way that final conclusions can be drawn (Miles and Huberman, 1994) - was also employed.

Firstly, the data were displayed and condensed into categories. The statements in the questionnaire and also the responses were grouped into categories. The 26 statements in the questionnaire were classified into three categories involving preparation stage, construction stage, and evaluation stage. Moreover, the responses were grouped into categories involving 'always', 'very often', 'sometimes', 'rarely', and 'never'.

After the data were grouped into categories, the data were then calculated based on the frequency and presented descriptively in words. The result of the data analysis of the questionnaire was used to answer the first research question of how the participants construct multiple-choice test items by identifying the steps taken by the participant in doing test construction.

3.5.2 Analysis of Data from Interview

The recorded data obtained from the interview were first transcribed in order to make it precise and accurate. During the transcription process, the personal data of the participants were coded to protect their privacy. The data were then analyzed following the framework of qualitative data analysis proposed by Cresswell (2009).

The analysis was started by preparation, organization, sorting, and arranging of the data. This process involved data reduction. After that, the data were displayed and carefully read to get the general meanings of the data. After the general meanings had been obtained, the data were coded and classified into meaning units. Finally, the last step was to draw an interpretation of the whole data and relate the findings to the theoretical background of the study.

3.5.3 Analysis of Data from Document Analysis

As any other methods of data analysis, document analysis requires the examination and interpretation of the data in order to elicit meanings, gain understanding, and develop empirical knowledge (Corbin & Strauss, 2008). The document analysis was intended to answer the second research question of the quality of the test items constructed by the participants. The document analysis was conducted following the procedures of criterion-referenced test item analysis proposed by Brown (1996) that involved item format analysis, item quality analysis, β-index analysis, validity analysis, and reliability analysis. The data were analyzed mainly using quantitative procedures and some qualitative procedures.

Firstly, in order to reveal the degree of validity, reliability, and the β-index of the test items, quantitative analysis was employed. The data were analyzed using Anates V4 and Ms.Excel computer software. The result of the computation was then described in words.

The next step was to conduct item format analysis and item quality analysis of the multiple-choice test items. The data were examined based on a checklist for reviewing multiple-choice test items adopted from Miller, Linn, and Gronlund (2009) (See Appendix). At this stage, the data were examined based on its materials, language, and construct. After the data were analyzed, both test items that had met the criteria of good test items and test items that should be revised were presented and described.

The last aspect to be analyzed was the proportion of higher-order thinking and lower thinking covered in the multiple-choice test items. The analysis was conducted using Bloom Taxonomy as the guidelines to classify higher-order thinking and lower-order thinking test items. The test items were analyzed and classified into the six cognitive domains in Bloom Taxonomy. The amount of test items in each cognitive domain was then calculated and presented in percentages. After that, the result of the analysis was then compared to the table of ideal proportion of higher-order thinking and lower-order thinking in multiple-choice test items proposed by Hamafyelto, Hamman-Tukur, and Hamafyelto (2015). Finally, the interpretation of the result of the data analysis was presented in words.

3.6 Concluding Remark

This chapter has covered the research methodology of the present study. It involved the elaboration of the research design, research site, participants, and the procedures of data collection and data analysis.