

# CHAPTER I

## INTRODUCTION

### A. Background

Teachers' work is teaching and doing assessments as part of their responsibility of learning process of their pupils. Analogically, it is like what a chef does. He does not only have to do the cooking, but he also needs to test the dish he has prepared. Without tasting his/her cook, a chef will not know the quality of the serving, i.e., the quality of his professional work. It is also the case with a teacher. Without conducting an assessment, a teacher will never know the quality of his/her teaching.

Pramono (2014:12) argues that a test is a tool with a series of tasks that students must do to measure their abilities. When students learn a language, a language test is a tool that provides information that is relevant to several concerns in language teaching (Hughes, 2003:8). According to Bachman and Palmer (1996: 8), the purpose of language testing is for measuring language proficiency, discovering students' success in achieving the objectives of a study course, diagnosing students' strengths and weaknesses in learning, and assisting in students' placement. Similarly, Harmer (2007: 379 – 381) explains four main reasons for testing. They are placement tests, diagnostic tests, progress or achievement tests, and proficiency tests. A placement test is for placing new students in the right class in a school. A diagnostic test exposes learner difficulties, gaps in their knowledge, and skill deficiencies during a course.

Progress or achievement test measures learners' language and skill progress concerning the syllabus they have been following, while proficiency test gives a general picture of a student's knowledge and ability.

Then, do all teachers design a test? In practice, teachers rarely take tests independently; most of them could not design one (Osna et al.;2016, Nazaruddin; 2017). Hughes (2011; 1) found a lack of quality in language testing since no accurate measurement is available to do it. Quansah, Amoako, and Ankomah (2019) revealed that the tests designed by teachers are not qualified. This problem is found in teachers at any level, despite their professional certifications and years of learning experience. They simply lack knowledge in implementing their tasks. The knowledge is not there to support their duties in deciding the objective of teaching, developing the specifications table, and assessing student learning outcomes.

Based on the studies above, it is evidence that tests designed by teachers are not reliable. The questions then arose: does the problem with the test item exist in Banyumas regency? If there are, what kinds of problems are they? Furthermore, what factors are behind the problems? To answer those questions, the writer designs a study to analyze the items of question in *Ujian Sekolah Berstandar Nasional*, USBN, or National Standardized School Examination (NSSE). The items are used to measure the students' academic competence before they graduate from school. It is a final examination for measuring their achievement related to the syllabus they have followed. Its result will determine the students' learning graduation as it is written in the Minister of Education and Culture number 81A

year 2013 about the implementation of the 2013 curriculum that students are declared to be graduates, only if they pass this final examination.

Thus, USBN is one of the educational evaluations. Educational evaluation is essential as it indicates instructional achievement in all nations (Tangsakul et al., 2017). In other words, education evaluation reflects the education quality in each country. The judgment of the quality is determined by the result of the evaluation in each school. Thus, information on school learning outcomes is needed. Therefore, USBN is undertaken as a final step of students' learning at school, and it plays a crucial part in reflecting the quality of education. It is in line with the Education and Culture Ministry/Mendikbud (2017) statement that USBN is an assessment of students' learning outcomes conducted by the educational unit as a process of students' learning outcomes information gathering.

As a consequence, USBN must be prepared well by students and also teachers with their different roles. The students should read the materials very well. Their teachers have to design qualified tests. The test items in the school examination must be designed by a teacher on duty in each school.

For English subjects, the USBN is a 120 minutes test consisting of 45 multiple-choice questions and five essay questions comprising two parts, Listening and reading. The listening section includes ten questions designed by MGMP (a local English teachers association at the regency level). Meanwhile, the other 40 questions are written by teachers at the school. These questions are of two kinds, 35 multiple-choice items and five essay questions.

Each kind of test item poses different challenges. It is easier to measure multiple-choice but comparatively is difficult to design (Case and Donahue, 2008). Moreover, Khan et al. (2013) argue that creating a single best question is challenging and time-consuming, even for those formally trained in multiple-choice question designing. Suwartono (2019) strengthened the idea that multiple-choice items seem too easy for test-takers, but the perception may lead to serious threats. Indeed, preparing the multiple-choice items requires more considerations in addition to validity. These emphasize that writing multiple-choice questions is not easy. In writing the USBN test questions, teachers have to follow the Educational Assessment Center/*Puspendik*. Unfortunately, the teachers in charge usually do not get any designing test training. This raises the question of whether the multiple-choice test items have met the guidance of good writing test items.

Some researchers have analyzed multiple-choice item questions. Gajjar et al. (2014) evaluated multiple-choice questions item and developed a pool of valid items by assessing with difficulty index (DIF I), discrimination index (DI), and distractor efficiency (DE). It was conducted in a medical school in Ahmedabad, emphasizing the selection of quality MCQs that honestly assess knowledge and distinguish students from different backgrounds and skills correctly.

For the academic quality of multiple-choice items, Rao, etc. (2016) had conducted a study on an item analysis of multiple-choice questions in the department of Pathology in India. They collected, summarized, and used information from students' responses. The Difficulty index (P) and discrimination

index (D) are used as the parameters to evaluate the standard of multiple-choice questions.

Toksöz and Ertunç (2017) had conducted an item analysis of multiple-choice items. They analyzed 453 students' responses in terms of item facility, item discrimination, and distractor efficiency by using frequency showing the distribution of the responses of preparatory students. Similar studies have also been done in the Indonesian context. Agustiana et al. (2018) examined the quality of multiple-choice item analysis using SPSS Version 22 and Anates program version 4. Setyawarno (2020) investigated the analysis of multiple-choice items using the *Item* (Item and Test Analysis) computer application. He analyzed the quality of multiple-choice test items based on the classic theory approach.

Those studies have analyzed the multiple-choice item quality based on the students' responses as test-takers; none analyzes the point based on good criteria of multiple-choice item designing. English teachers need skills on how to write good multiple-choice items. Indonesian English teachers have to gain good skills in designing multiple-choice items based on the government's criteria. The government in charge of test designing is *Pusat Penilaian Pendidikan* or *Puspendik* (Assessment Centre of the Indonesian Ministry of Education and Culture). The study about an analysis of multiple-choice test items based on government criteria has not been conducted yet. Based on these reasons, the writer designs research to fill the gap, entitled Multiple-Choice Test Items of English National Standardized School Examination: Revealing Problems and Factors Causing Them, by examining the test item questions focusing on the teachers'

competence in making the multiple-choice test questions for USBN. This research focuses on the problem raised in multiple-choice test items since there should be nothing wrong with the USBN test.

### **B. Statements of Problems**

The problems proposed in this study are:

1. What problems are found in the USBN multiple-choice test items designed by English teachers?
2. What are the impacts of the problems on the quality of USBN multiple-choice test items designed by English teachers?
3. What factors are responsible for the problems?

### **C. Purposes**

This study aims to investigate problems in the USBN multiple-choice test items designed by English teachers, covering the following points:

1. To identify the problems in USBN multiple-choice items designed by English teachers in Banyumas regency.
2. To describe the impact of the problems on the quality of USBN multiple-choice test items designed by English teachers in Banyumas regency.
3. To identify the factors responsible for the problems.

## **D. Significance, Scope, and Definitions**

### **1. Significance of the study**

The study's findings will benefit English teachers as enrichment knowledge for understanding the test blueprint and designing a good writing test. It is also expected to improve the proficiency of English teachers in designing tests. Since teachers have the role of schools' main factor, this study will be useful information about teachers' ability to write the multiple-choice test items. The information can be a consideration for improving teachers' competence in designing tests by conducting in-house training (IHT) at school.

English Teachers Association (Musyawarah Guru Mata Pelajaran/MGMP Bahasa Inggris) gets advantages from this study because information about the analysis of multiple-choice test items designed by teachers will mirror teachers' ability in designing the multiple-choice test items. The information can be a consideration in mapping out some activities associated with teachers' competence upgrading program.

This study can be an asset to the government since the information of this research result can lead the government in making some considerations about teachers' needs in their practice to evaluate students. It is expected that the government will hold some teachers' competence upgrading programs needed by teachers.

## 2. Scope and Definitions

### 1. Multiple-choice items

In providing scope and definition of multiple-choice items, the writer refers to definitions given by Fulcher (2010) and because teachers in Banyumas regency must submit to the rules of writing the questions set by the Development Directorate of Senior High School, General Directorate of Elementary and Secondary School, the Ministry of Education and Culture. Consequently, in this study, the writer will also refer to the theory of designing test written in *Panduan Penulisan Soal SMA/SMK/MA/MAK* published by *Pusat Penilaian Pendidikan, Badan Penelitian dan Pengembangan, Kementerian Pendidikan dan Kebudayaan Indonesia*. Both of them stated that multiple-choice items consist of a stem, which gives the prompt. Then options follow. One of these options is the answer, which is the correct response. The other choices are distractors or wrong answers.

### 2. USBN

In delivering a scope and definition of USBN, the writer adopts the Minister of Education and Culture Regulation number 4 the year 2018 concerning Assessment of Learning Outcomes by the Education Unit and Assessment of Learning outcomes by the government. It is stated that the National Standardized School Examination, abbreviated as

USBN is an activity to measure student competency performance. The Education Unit carries it out by referring to Graduates' Competency Standards to obtain recognition for learning achievement (Chapter 1, article 1, paragraph 5)

